

# Development of a School-Based Treatment Program for Middle School Youth With ADHD

Steven W. Evans  
Joshua Langberg  
Veronica Raggi  
Jessica Allen  
Elizabeth C. Buvinger  
*James Madison University*

The authors conduct an evaluation of a middle school-based treatment program for youth with ADHD during early stages of treatment development. The studies focus on interpreting outcome trends in preliminary data and identifying assessment issues that will be important to consider when conducting a clinical trial. Parent reports indicate that the majority of students benefit from improvements in academic, social, and overall functioning. Although teachers report beneficial effects for the majority of the participants in the program, there is little agreement about individual students. Measurement problems are associated with understanding normal change during a school year, accounting for normal behavior changes in December and May, and considerable disagreement between teachers. Suggestions to guide future work in this area are provided. (*J. of Att. Dis.* 2005; 9(1), 343-353)

**Keywords:** ADHD; treatment; school

Recent epidemiological data suggest a prevalence rate between 2% and 4% for adolescents with ADHD in the general population (Barkley, 1998). Similar to children with ADHD, presenting problems for adolescents with ADHD typically include academic difficulties, discipline problems at school and home, and conflict with peers. Adolescents tend to manifest these problems with more serious consequences than children do, including school dropout, substance use, and legal problems. Between 25% and 75% of adolescents with ADHD meet diagnostic criteria for other disruptive behavior disorders (Barkley), resulting in a high rate of comorbidity that makes working with a wide range of problems an inevitable part of treating them. Therefore, adolescents with ADHD are at a high risk for serious impairment, and this requires a comprehensive, multimodal approach that addresses impairment across multiple contextual domains (Faraone, Biederman, & Monuteaux, 2002).

There are numerous studies of stimulant treatment for children with ADHD (Swanson, McBurnett, Christian, & Wigal, 1995); however, in a review, we found only 20 controlled studies of the use of stimulants with adoles-

cents (Evans et al., 2001; Findling, Short, & Manos, 2001; Smith, Waschbusch, Willoughby, & Evans, 2000). Overall, the state of the literature on stimulant treatment with children and adolescents suggests that stimulants are efficacious but produce improvement short of normalization of functioning. Problems related to sustained adherence to medication regimens, resistance by some families and youth to taking medication, presence of a significant number of nonresponders, and the fact that stimulants do not normalize behavior in many areas support the need to

Authors' Note: This study was supported in part by the Alvin V. Baird Attention & Learning Disabilities Center, the Mosier Fellowship at James Madison University, and a grant from the Virginia Tobacco Settlement Foundation to the first author. We are grateful to the adolescents, their parents, and the teachers and administrators in the Rockingham County Schools. We would like to thank Zewelani Serpell and Brandon Schultz for their helpful comments on drafts of this article and Lyn Hart for her contributions.

Address correspondence to Steven W. Evans, Alvin V. Baird Attention & Learning Disabilities Center, MSC 9013, James Madison University, Harrisonburg, VA 22807; e-mail: [evanssw@jmu.edu](mailto:evanssw@jmu.edu).

develop and evaluate psychosocial interventions for adolescents with ADHD.

Because the vast majority of the literature on the treatment of ADHD has focused on elementary school-aged children, the temptation exists to base treatments for adolescents with ADHD on studies of children. Although some studies suggest that medication treatment may indeed generalize from children to adolescents (e.g., Smith, Pelham, Gnagy, & Yudell, 1998), the continuity of psychosocial treatment is largely untested. The independence, social influences, ability to defy adult directives, and dramatically different school environments are a few of the major differences between the age groups that influence the feasibility and effectiveness of treatments. As a result, investigations into family, academic, and behavioral interventions with this population are important, and some studies have begun to address this deficiency in the literature.

Family psychosocial intervention is frequently recommended for adolescents with ADHD (Robin, 1998), and two studies have examined this type of treatment (Barkley, Edwards, Laneri, Fletcher, & Metevia, 2001; Barkley, Guevremont, Anastopoulos, & Fletcher, 1992). In the first (Barkley et al., 1992), three types of family-based treatments were compared, and all three performed about equally and produced statistically significant improvements on a variety of parent and adolescent rating scales. However, few participants exhibited clinically significant improvement, and the authors concluded that more potent treatments were needed than 8 to 10 1-hr sessions. Subsequently, Barkley et al. (2001) increased the intensity (doubling the number of sessions) of the family treatment and again reported statistically significant improvements. Tests of clinical significance measured by a reliable change index and movement into the normal range revealed that only a minority of the participants demonstrated clinically significant improvement. Additional treatment modalities may need to be provided with family treatment to achieve the desired benefit.

Classroom-based behavioral interventions also appear to hold promise for adolescents with ADHD. In particular, one study evaluated the effects of functional assessment followed by classroom-based interventions with middle school students with ADHD (Ervin, DuPaul, Kern, & Friman, 1998). In this case study, the authors completed a functional assessment of target behaviors of two middle school-aged youth who exhibited high rates of off-task behavior. Following the assessment, functional hypotheses were developed and recommendations offered to teachers. After teachers implemented the rec-

ommended interventions, the investigators found large improvements in on-task behavior.

In addition to behavioral interventions targeting on-task and disruptive behavior within the classroom, there is some literature to suggest that academic interventions that manipulate academic instruction and materials and target both behavioral and academic outcomes may be effective for adolescents with ADHD. In one study, adolescents were taught to take structured notes while listening to a lecture-format history class (Evans, Pelham, & Grudberg, 1995). Following this note-taking training, the adolescents showed improvement on measures of comprehension and increased levels of on-task behavior on days when they took notes compared to days when they simply listened to the lectures. The effect size (ES) for the note-taking intervention was in the moderate range. Other academic accommodations have been suggested for college students with ADHD, including modifications of required courses, test accommodations, note taking, audio recordings of written material, and organization skills (e.g., Richard, 1995). Evaluation of the benefit of these strategies has yet to be completed with college students or adolescents.

Overall, the literature on family treatment, behavioral interventions, and academic techniques suggests that none of these strategies are likely to be sufficient on their own. In fact, they each target different areas of impairment, suggesting that a combination of these treatments may be necessary to address the broad range of impairment characteristic of adolescents with ADHD. Because problems at school tend to be a priority for many parents of these youth (Robin, 1990), providing treatment at school may be the best opportunity to effectively intervene. This rationale led to the development of the Challenging Horizons Program (CHP) by the first author in 1999. The CHP is a school-based treatment program for middle school youth with ADHD. Parents refer their children to the program (frequently in response to suggestions by teachers and school administrators), at which point a full educational and diagnostic evaluation is administered. The CHP combines family, academic, social, and behavioral interventions into an after-school treatment program that operates 3 days per week for 2 hr and 15 min per day for the entire academic year. The treatment manual for the CHP includes both group and individual academic and social interventions. Specific descriptions of these interventions have been reported elsewhere (Evans, Dowling, & Brown, in press). A pilot study reported initial results for seven middle school students enrolled in the CHP (Evans, Axelrod, & Langberg,

2004). The report indicated that participants in the CHP demonstrated improvement in social and academic functioning, with most ESs in the moderate to large range. Preevaluations and postevaluations revealed that using the index for reliable change (Jacobson & Truax, 1991), participants demonstrated clinically significant improvement on 46% of the pre-post parent and teacher ratings of ADHD symptoms and functioning. Although encouraging, interpretation was difficult as there was no control group and the sample size was very small ( $N = 7$ ). In addition, there was considerable variability in the teacher report data, raising questions about outcome measurement techniques.

Recent reports have indicated that teacher rating data from secondary school teachers are not as reliable or valid as data from elementary school teachers (Evans, Allen, Moore, & Strauss, in press; Molina, Pelham, Blumenthal, & Galiszewski, 1998; Simpson, 1991). These differences between elementary and secondary school teachers are not surprising given the large differences between elementary and secondary school environments. Elementary school teachers observe their students in a wide range of academic and social situations, including class work, walks to the restrooms, and play activities. Because elementary school teachers are seen as the focal point for their assigned children, other teachers, staff, and administrators keep them updated about events outside of the classroom, such as bus behavior, progress in physical education class, and behavior in the lunchroom. In addition, parents tend to communicate more with elementary school teachers than with secondary school teachers. As a result, elementary school teachers develop considerable breadth and depth to their knowledge of the 20 to 30 students with whom they spend approximately 6 hr per day. On the other hand, secondary school teachers are frequently assigned more than 100 children and observe their behavior for approximately 1 hr per day. They are not the focal point of their assigned students' education, so it is common that they do not know about problems in settings outside of their classroom. Finally, their correspondence with most children's parents is often minimal to none. As a result, the value of secondary school teachers as sources of information about students is certain to be diminished compared to that of elementary school teachers.

Nevertheless, it is critical to obtain data about school functioning when measuring the effects of treatment because many referral problems center on the school setting. Although a parent report offers one potential source for these data, there are concerns about using parents as the sole source of data because many parents report that

their knowledge of their children's school functioning decreases greatly once their children begin middle school. Previous reports have indicated that parent reports are not a suitable replacement for teacher reports, even for elementary school-aged children (Mitsis, McKay, Schulz, Newcorn, & Halperin, 2000). Nevertheless, very few outcome studies with adolescents use secondary school teacher report as a dependent measure.

The purpose of this study is to address two of the questions involved in the early development phases of a new treatment (Rounsaville, Carroll, & Onken, 2001; Weisz, Southam-Gerow, Gordis, & Connor-Smith, 2003). First, early studies are needed to obtain estimates of the potential benefits of the CHP as an emerging treatment model for middle school youth with ADHD. Second, it is important to identify methods for measuring outcomes with secondary school-based treatments for youth with disruptive behavior disorders. This is especially critical for this population given the assessment problems reviewed in the previous two paragraphs. Although case studies from the first year of this development process were reported elsewhere (Evans et al., 2004), this article reports data from the preliminary evaluations of the CHP during its second through fourth year of development.

## Study 1

The purpose of this first study was to obtain ES estimates that could be used when conducting power analyses to establish sample sizes in future studies. To compare treatment effects with normal change, this study used a quasi-experimental design including two groups (treatment and community care). A community care group was used instead of comparing the CHP to another type of psychosocial treatment or to a waitlist control because there are no psychosocial treatments for adolescents with ADHD that have significant empirical support to warrant a comparison (Smith et al., 2000), and the use of a waitlist control receiving no treatment was not practical given the 9-month treatment period. Participants in the community care group were recruited from schools that did not provide the CHP, and they were able to pursue medication or psychosocial treatments available in the community.

## Participants

Participants in both groups were diagnosed with ADHD either combined or inattentive type based on the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text revision; American Psychiatric Association, 2000). Diagnoses were established using the Diagnostic

Interview Schedule for Children (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000), ADHD Rating Scale (ADHD-RS; DuPaul, Power, Anastopoulos, & Reid, 1998), and the Impairment Rating Scale (IRS; Fabiano & Pelham, 2002) administered by a school psychology graduate student trained by the first author. The graduate student also administered the Wechsler Intelligence Scale for Children (WISC-III; Wechsler, 1991) and the Wechsler Individual Achievement Test (The Psychological Corporation, 1992). Parents and two teachers completed the ADHD-RS and IRS for each participant. Inclusion criteria included a minimum WISC-III Full Scale IQ of 80, a diagnosis of ADHD (any subtype), and the absence of bipolar disorder or psychosis. Those children who did not meet these inclusion criteria were not admitted to either group but were given options of alternative services. All parents and youth participating in either group signed appropriate consent and assent forms to participate in the study.

Participants in the treatment group included 13 Caucasian middle school youth (2 female) ages 11 to 14 who attended the school that housed the CHP during the 2000-2001 school year. The mean IQ and achievement scores with standard deviations were as follows: for IQ,  $M = 98.9$ ,  $SD = 17.2$ ; for reading,  $M = 100.4$ ,  $SD = 14.1$ ; for math,  $M = 97.0$ ,  $SD = 19.7$ ; and for written language,  $M = 109.8$ ,  $SD = 13.7$ . Nine of the 13 participants took stimulant medication throughout the year.

The participants in the community care group were recruited through newspaper advertisements. Parents of middle school students attending schools other than the one housing the CHP called research staff in response to the advertisement. Parents whose children had a previous diagnosis of ADHD or presented histories consistent with a diagnosis of ADHD came to the clinic to complete the same evaluation as the participants in the treatment group. The community care group consisted of 14 Caucasian middle school students (4 female) ages 11 to 14. The mean IQ and achievement scores with standard deviations were as follows: for IQ,  $M = 100.8$ ,  $SD = 13.6$ ; for reading,  $M = 94.7$ ,  $SD = 11.7$ ; for math,  $M = 97.1$ ,  $SD = 14.4$ ; and for written language,  $M = 105.5$ ,  $SD = 9.3$ . The groups were equivalent on these measures. Thirteen out of the 14 community care participants took stimulant medication throughout the year.

## Procedures

The participants in the treatment condition attended the program after school 3 days per week during the course of their involvement in the CHP. The program began each day with an individual meeting between the

primary counselors and students to review individual interventions pertaining to organization and the accurate recording of assignments. Goals were tracked and modified during these meetings as needed. Following the individual meetings, everyone participated in the group social skills intervention. Social skills interventions involved the development of social problem-solving skills, learning conversation skills, and recognizing others' perceptions of one's own behavior. This intervention was followed by a recreation period that involved playing a sport and included prompts, reinforcers, and modeling of appropriate social skills by counselors. After returning from the recreation period, the entire group participated in educational interventions followed by a brief individual meeting with the primary counselor. The individual and group educational interventions included instruction in academic strategies, such as note taking (Evans et al., 1995), the development of organizational skills for their assignments and materials, and study skills. Group parent meetings were held monthly to provide parents with information about the interventions being provided in the program and to give them information about ADHD and the treatment options available to them. Individual parent meetings were available to parents wishing assistance with parenting issues. Additional information about the interventions in the CHP is available elsewhere (Evans, Dowling, et al., in press).

## Dependent Measures

The IRS (Fabiano & Pelham, 2002) consists of six questions (seven questions on the parent version) pertaining to school, family, self-esteem, and social functioning. Scores range from 0 to 6, with higher scores indicating greater impairment. A score of 3 or greater is considered impaired. Parent and teacher versions were used as dependent measures to provide data from two perspectives on academic progress, social functioning, and overall impairment. Each student's math and science teachers completed the IRS at the beginning of the child's placement in the CHP and at the end. The interventions in the CHP primarily targeted academic, social, and overall functioning, so these were the only IRS items used in the outcome analyses. Interrater reliability for these three scales for middle school teachers are .37 (academic), .23 (social), and .33 (overall; Evans, Allen, et al., in press). It has also been reported that the parent-completed IRS has good sensitivity (ranging from .53 to .79) and specificity (ranging from .51 to .80) for identifying youth in need of treatment (Raggi et al., 2003).

In addition, grade point average (GPA; calculated based on reading, math, science, and social studies

grades) was a dependent measure. Although grades can be influenced by many factors, they are one of the most frequently used indices of school functioning. In addition, many of the students developed academic goals that were operationally defined in relation to grades. Grades were collected six times a year based on three grading periods per semester.

The ADHD-RS (DuPaul et al., 1998) was administered at the beginning and end of the school year to parents, science teachers, and math teachers. The ADHD-RS provides an overall score and scores on each of two factors: hyperactivity/impulsivity and inattention. Reliability and validity data are reported by the authors in the manual (DuPaul et al., 1998). Internal consistency of the subscales and total score for the Teacher Rating scale were between .88 and .96, and test-retest reliability estimates for the same three scores ranged from .88 to .90. The data for the parent report version were similar (but slightly lower), with internal consistency between .86 and .92 and test-retest between .78 and .86. Agreement between parents and teachers was in the moderate range and between .40 and .45.

## Results

In the original sample, the CHP group consisted of 13 students, and the community care group consisted of 14. Despite persistent effort, however, there were parents and teachers who did not complete all of the ratings scales. Participants with missing data for a certain rater were not included in the relevant analyses. As a result, 23% of parent data, 23% of science teacher data, and 23% of math teacher data from the CHP group were missing in the analyses. This resulted in data for 4 students in the CHP being eliminated from the analyses, as they were missing multiple ratings. In the community care group, science teacher and parent data were complete on all 14 participants; however, math teacher data were missing for 4 participants, restricting analyses on this measure to 10 participants.

Initial analyses involved calculating the within-group ES for each group for the ratings of social, academic, and overall functioning from the IRS (premean minus postmean and pooled standard deviations; see Table 1). Given the small sample size and quasi-experimental design, ES analyses were completed because they can provide a much clearer description of outcomes in these situations than statistical analyses of group means by time. Using Cohen's (1988) definitions of small, medium, and large ES, these data indicate that parents reported moderate beneficial effects for the treatment group, whereas the parents of participants in the community care

**Table 1**  
**Effect Sizes for Primary Dependent Measures**  
**From the Impairment Rating Scale**

Rater or Domain	Treatment	Control
Parent		
Peers	.70	.21
Academics	.76	-.44
Overall	.51	.05
Science teachers		
Peers	-.66	-.04
Academics	.30	-.21
Overall	-.24	-.23
Math teachers		
Peers	-.18	-.28
Academics	.46	.39
Overall	.32	.36

sample reported little benefit or declines. The data from the teachers yielded inconsistent results and do not support any effect of treatment. The teacher data do converge to suggest a small to medium beneficial effect for the treatment group in the area of academics (ES = .30 and .46). Conflicting effects were reported by teachers for the area of academics with the community control sample (ES = -.21 and .39).

The ES calculations provide information on the magnitude of change. In addition, an indication of the effect of treatment may also be estimated by calculating the percentage of the sample that began treatment in the impaired range and ended in the normal range. Examining both indices of change allows for the consideration of the magnitude of change and the extent of clinical recovery. These two indices are independent constructs in the determination of effective outcomes. Using a cutoff of 3 on the IRS (Fabiano & Pelham, 2002) to define *impaired*, these calculations were completed using the same three items on the IRS as reported in the ES analyses above. According to parent ratings on the IRS, 50% of the participants in the CHP group began the year in the impaired range for academic functioning, and 60% of those who began the year in the impaired range finished the year in the normal range. In contrast, parents of participants in the community care group reported that 71% of the participants began the year in the impaired range for academic functioning, and 20% of them finished the year in the normal range. According to parent ratings of social functioning, 80% of the participants in the CHP group began the year in the impaired range, and 38% of them finished the year in the normal range. In contrast, parents of participants in the control group reported that 79% of

the participants began the year in the impaired range for social functioning and that 9% of them finished the year in the normal range. Finally, parents rated 80% of the CHP participants in the impaired range for overall functioning prior to treatment, and 38% of them ended treatment in the normal range. Corresponding numbers for the control group were 93% and 8%. In summary, of those youth in the CHP group who started the year in the impaired range in any of the three domains being assessed, between 38% and 60% (depending on the domain) improved to the normal range in at least one domain compared to the range for those in the community care group, which was only 8% to 20%.

These same analyses were completed using teacher rating data, with the criterion for impairment being that one or both teachers rated the participant in the impaired range prior to treatment and the participant's scores were in the nonimpaired range at the end of treatment on both teachers' ratings. Using these criteria, there were no participants in either group who moved from the impaired to the nonimpaired range according to teacher ratings.

Further analysis of school functioning was conducted using a repeated measures analysis of variance (ANOVA) with group as the independent variable and GPA as the dependent measure. The GPA score was calculated by determining the mean GPA for each semester with Grading Periods 1, 2, and 3 averaged for Semester 1 and Grading Periods 4, 5, and 6 averaged for Semester 2. A Time  $\times$  Group interaction nearing significance was found for semester,  $F(1, 23) = 3.588, p = .071$ . Follow-up  $t$  tests revealed that the CHP and community care groups did not significantly differ on GPA for either semester ( $p > .1$ ). Although there was no significance between groups on GPA, the students in the community care group did significantly decline from the first semester to the second,  $t(1, 13) = 3.02, p = .01$ , whereas there was no decline in the treatment group. In fact, the mean GPA for the treatment group increased slightly, although this finding was not statistically significant. These results indicate that the academic benefit of the CHP may be to prevent a decline in grades that appears to be characteristic of youth with ADHD receiving medication prescribed in the community.

To understand the practical implications of this benefit, the mean number of Ds and Fs received by each student was calculated for each group as a measure of academic progress, as many parents view these grades as unacceptable. Examination of mean numbers of Ds and Fs indicated that the participants in the CHP group averaged 3.27 Ds and Fs and that the students in the community care group averaged approximately 4.1 Ds and Fs during the first semester. The CHP participants' mean number of

Ds and Fs remained approximately the same at 3.18 during the second semester, whereas the community care group's average increased to 5.1. As a result, participants in the community care group were receiving almost two more Ds or Fs during the second semester than participants in the CHP group. Although these findings are encouraging, it should be noted that 55% of the participants in the CHP group still receiving more than 1 D or F, suggesting that further enhancement of the academic interventions in the CHP is warranted.

## Study 2

The results from Study 1 provided evidence that data from parent report and grades suggested beneficial treatment effects for adolescents participating in the CHP. Teacher data revealed a small to medium beneficial ES for progress in academics for the treatment group but inconsistent results for all other items and groups. The purpose of this second study was to address the measurement issue highlighted by the inconsistency of teacher ratings reported in Study 1. In Study 2, parent and teacher data were collected monthly from a treatment group for 2 additional years of the CHP to facilitate examination of trends during the year and to inform the assessment practices for the evaluation of the CHP.

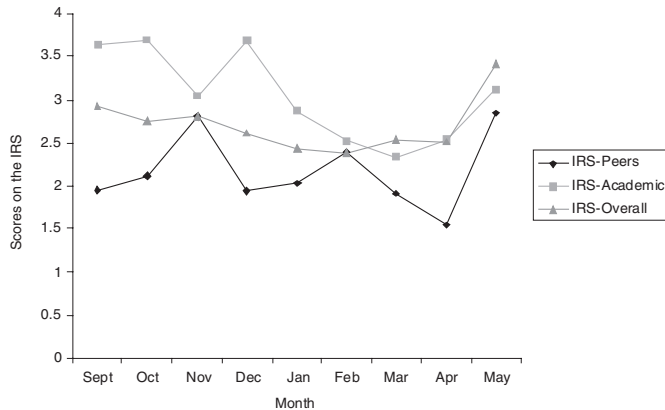
## Participants

Participants in Study 2 were 35 Caucasian middle school students ages 11 to 14 who attended the CHP between fall 2001 and spring 2003. Six of the students involved in the 2001-2002 school year also participated in the 2002-2003 school year and thus contributed data in both studies. The initial evaluation methods were the same as those described in Study 1. The final sample included 6 females, and the mean full-scale IQ for the entire group was 100.4 ( $SD = 11.9$ ). Mean and standard deviations for achievement scores were as follows: for reading,  $M = 98.4, SD = 13.6$ ; for math,  $M = 96.7, SD = 15.5$ ; and for written language,  $M = 94.7, SD = 10.4$ .

## Results

Initial analyses of the data were performed on the 35 participants from Study 2 using the monthly ratings of peer relations, academic progress, and overall impairment on the IRS and the three scores from the ADHD-RS (total, inattention, and hyperactivity/impulsivity). Data used in this examination were limited to those participants who were in the program for all 9 months of the school year, so data trends would not be confounded with

**Figure 1**  
Science Teacher Ratings on the IRS

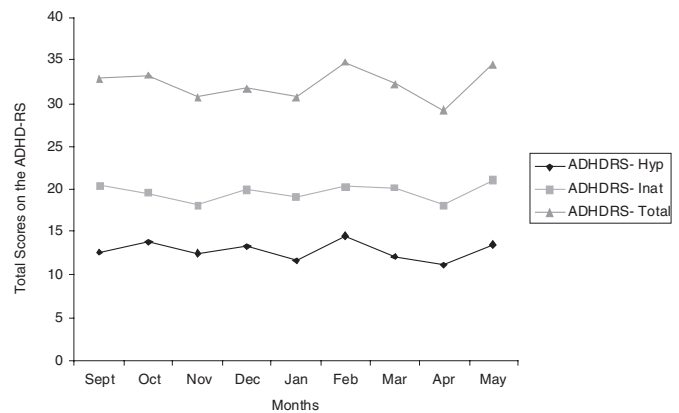


Note: IRS = Impairment Rating Scale.

data from participants who started or stopped during the year. There was a great deal of variability and unique information about individual participants depending on the rater and time of year. It was noted that the rating scale data became notably worse for the final month of the academic year (see Figures 1 and 2). This is consistent with the frequently noted observation that students' level of behavior and academic performance declines in May in anticipation of the end of the school year. As a result, a simple first month–last month comparison is likely to compromise the interpretation of potential treatment effects. It may be more valid to consider March and April ratings when examining the effects of year-long school-based interventions than data from May.

To examine the potential treatment benefits of the CHP for the participants, data were combined to create estimates of impairment and achievement for each participant. Teacher data were averaged for the first 3 months to produce a beginning-of-the-year measure, and data from March and April were averaged to provide an end-of-the-year measure. In addition, the primary analyses focused on the global indicators of impairment (overall score on IRS) and symptoms (total score on ADHD-RS). Focusing on the main indices of symptoms and impairment and pooling data across initial and ending months were done to increase the reliability and validity of the findings and to improve our opportunity to see trends. After eliminating those who participated less than 9 months, 19 participants remained. Of these 19, there were sufficient parent rating data for 15 participants, sufficient math teacher data for 14, and sufficient science teacher rating data for all 19 participants. In this reduced sample, the rates of missing data were 15% of parent data, 5% of science

**Figure 2**  
Parent Ratings on the ADHD-RS



Note: ADHD-RS = ADHD–Rating Scale; Hyp = Hyperactive; Inat = Inattentive.

teacher data, and 2% of math teacher data. When data were missing, the means described above were calculated using only the available data. Missing data were almost entirely attributable to variations in teachers' willingness to complete the scales. Although this results in variations in the participants in the samples being evaluated, this technique was adopted to maximize the sample size for each analysis.

There was no notable change and little month-to-month variability in the total score on the ADHD-RS between raters. Parents tended to rate their children as more impaired than the teachers did, as indicated by the mean scores by rater (for parents,  $M = 32.5$ ; for math teachers,  $M = 18.3$ ; and for science teachers,  $M = 22.1$ ). There was considerable variability in the parent and teacher ratings of overall impairment on the IRS. To identify which of the participants (if any) had improved and whether there was agreement between the raters as to who improved, three within-subject ESs (parent, science teacher, and math teacher) for ratings of impairment (i.e., IRS overall) were calculated. The participants were classified as either improving ( $ES > 0$ ), remaining the same ( $ES = 0$ ), or declining ( $ES < 0$ ), and data based on this classification are reported in Table 2. The ESs indicated that the majority of participants improved, and the mean ES for those who improved was greater than 1.2 for each rater (e.g., 8 of 15 improved according to parent ratings).

There was very little agreement between the raters on whether individual participants improved. Of those participants who had complete data from all three raters, there was agreement on classification (improved, remained the same, or declined) on only 27% of the cases,

**Table 2**  
**Improvement Classification and Effect Sizes**

Rater	<i>n</i>	Improvement			Same	Declined		
		Number Improved	<i>M</i> <i>ES</i>	Range	Number Stayed Same	Number Declined	<i>M</i> <i>ES</i>	Range
Parent	15	8	1.21	0.54 to 1.83	5	2	-2.34	-1.87 to -2.81
Math	14	11	1.40	0.28 to 2.18	2	1	-1.96	-1.96
Science	19	10	1.24	0.43 to 2.18	2	7	-0.84	-0.32 to -1.93
Number/Total		29/48			9/48	10/48		

and these were students who fell in the improved category. In other words, although parents, science teachers, and math teachers did agree that the majority of participants improved, there was disagreement about which of them showed improvement.

## Discussion

These studies were conducted to establish estimates of the effectiveness of the CHP and to identify critical assessment issues that will be helpful to guide methods in future research. Average ESs based on parent ratings of impairment were in the moderate range, whereas ESs based on teacher ratings of academic progress were in the small to moderate range. Restricting analyses of ESs to those who benefited from treatment resulted in ESs ranging from 0.28 to 2.18. Data suggest that those who benefit from the CHP are in the majority of those who participate, but there were a notable number of participants with negative ESs.

The data from this study raised questions about the definition of *effective*. The data from the community care sample revealed a decline in functioning during the course of the school year. As a result, an ES of 0 could represent an important clinical benefit. These data are consistent with frequent parent and teacher reports describing students who begin the year well, followed by a serious and steep decline in behavior and academics at some point in the school year. The amount of time between the beginning of the year and this crash in performance may be one indication of severity. Furthermore, the need for treatment may vary such that dosage of psychosocial treatment may need to be provided in an escalating manner across the academic year. The high dosage of treatment that may be provided at the end of 1 year may not be needed at the beginning of the next; however, it may be likely that the child returns to a point of needing the previous high dosage (or an even higher dose) at some point during the year. Identifying these trends and their interaction with various forms and dosages of treat-

ment will require continued research. The need for appropriate comparison groups in these studies will be critical in determining a definition of *benefit* and defining *degrees of benefit*.

The measurement issues identified in this study will present challenges to those studying changes in functioning at school for adolescents. The most straightforward recommendation from these data is to avoid using data collected in May. Data from the month of May consistently declined from the previous 2 months on almost all parent and teacher rating data. This is consistent with teachers' informal observations of declining behavior as summer vacation approaches. In traditional pre-post designs, May would be a logical target month for assessment, yet using data from May is likely to minimize the potential benefit of an intervention. For similar reasons, it may be wise to avoid using data from December. These data were not collected in this study because of our experience in schools indicating that these data are likely to be outliers. Students only spend a couple of weeks in school in December, and classroom activities frequently vary considerably from typical practice. These modifications to a monthly data collection schedule leave 7 months to obtain data.

Drawing conclusions based on divergent reports from teachers will require particular attention in the design of future studies. Not only was it inconsistent with parent ratings, but it was also inconsistent between teachers. It was not uncommon to have large ESs in opposite directions by different raters (e.g., math vs. science teacher) for the same child, and there was considerable disagreement about which participants improved. Although some of this may be caused by variability in student behavior between classrooms, the characteristics of teacher-student relationships in secondary schools may have limited the value of this information. As described previously, secondary school teachers have more students, see them less often, see them in fewer settings, and are less likely to receive information from others about the students than elementary school teachers. These variations

in teacher-student relationships are likely to reduce the amount of information a secondary school teacher knows about any one student and, as a result, increase the amount of error and variation in their student reports (Evans, Allen, et al., in press). It may be helpful to restrict requests for ratings from secondary school teachers to the area of academics because they are likely to be best able to make evaluations in this area. Future studies may also find it beneficial to combine data across months to increase the reliability and validity of the ratings. For example, we combined data from ratings in September through November to establish an estimate of functioning early in the academic year.

Using data from other sources may also help investigators interested in assessing levels of school functioning. Our use of grades is one example of this. In spite of the fact that grades are a rather molar index of functioning and are subject to many influences, we reported marginally significant treatment benefits on grades. Grades have the advantage of being very interpretable and salient indices of academic functioning. If interventions improve grades, then communities and school administrators will be interested in the services despite the measurement problems associated with using grades as a measure. If grades are used as an outcome measure for studies spanning more than one school, it will be important to determine if the grading scale used in the two schools are equivalent. For example, some schools use 94% as the lowest possible score to receive an A, whereas others use 90%. These variations could compromise the ability to interpret the findings. While requiring more work than simply collecting grades, it may be wise to collect the actual percentages achieved that are used to calculate the grades.

Other alternatives to teacher ratings include attendance and discipline data, as these are also routinely kept by schools. In our previous attempts to use these data, questions pertaining to the reliability of the school records arose and led to a decision to discontinue their use. For example, it has been our experience that many discipline incidents go unrecorded, the criteria used to determine when to send a child to the office (usually the threshold for recording an incident) varies tremendously between teachers, and the likelihood of recording an incident and receiving a consequence is partly a function of how busy the administrator is when the child arrives to the office. When schools use universal interventions, such as Positive Behavioral Support (Horner, Sugai, Todd, & Lewis-Palmer, 2005) and the system described by Molina, Smith, & Pelham (2005), the reliability and validity of this type of data may improve.

These preliminary data are consistent with our previous report (Evans et al., 2004), suggesting that the CHP has potential to be an effective psychosocial treatment program for young adolescents with ADHD. There are several limitations of this study that will have to be addressed before drawing any definitive conclusions. First, random assignment of participants to conditions within school is needed to control for the potential confound in Study 1 of an effect of school. Although community care participants were recruited from multiple schools that did not have the CHP, it is still possible that participants receiving the CHP treatment may have done better than the comparison group for reasons related to the school and unrelated to the CHP. Second, there were too many missing data in these studies. The timely collection of rating scale data from teachers and parents is a labor-expensive process, and new and efficient methods to improve these procedures are needed. Web-based teacher rating scale data collection as well as procedures that request teachers to complete the instruments while the research staff waits may be helpful. Third, there is a need to develop instruments or procedures that improve the validity and reliability of teacher reports. The problems related to these data described in this article make interpretation difficult. It may be that given the changes in context between elementary and secondary schools, there will have to be different versions of teacher assessments. The version for secondary school teachers may need to focus primarily on academic functioning and class disruption and leave assessments of other factors (e.g., peer relations) to other techniques. In addition, there may be other useful sources of information in secondary schools that could provide valuable information about behavior and interpersonal skills. Cafeteria monitors, bus drivers, and administrators who manage school discipline are likely to know students with moderate to severe impairment because of ADHD and their participation in the assessment process may be useful.

From these limitations came many of the recommendations for future research in this area. There is a critical need to develop and evaluate effective psychosocial treatments for adolescents with ADHD, and schools provide an excellent context for this work. In addition, they provide serious challenges to those techniques developed for controlled studies in clinical settings. The treatment development procedures outlined by Weisz et al. (2003) and Rounsaville et al. (2001) emphasize the need to address many of these questions early in the development process. Addressing the issues raised here in the design of subsequent research will increase the likelihood that successful

assessment techniques will be employed and effective treatment and data collection procedures will be used.

## References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Barkley, R. A. (1998). *Attention Deficit Hyperactivity Disorder: A handbook for diagnosis and treatment*. New York: Guilford.
- Barkley, R. A., Edwards, G., Laneri, M., Fletcher, K., & Metevia, L. (2001). The efficacy of problem-solving communication training alone, behavior management training alone, and their combination for parent-adolescent conflict in teenagers with ADHD and ODD. *Journal of Consulting and Clinical Psychology, 69*, 926-941.
- Barkley, R. A., Guevremont, D. C., Anastopoulos, A. D., & Fletcher, K. E. (1992). A comparison of three family therapy programs for treating family conflicts in adolescents with Attention-Deficit Hyperactivity Disorder. *Journal of Consulting and Clinical Psychology, 60*, 450-462.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
- DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale-IV*. New York: Guilford.
- Ervin, R., DuPaul, G., Kern, L., & Friman, P. (1998). Classroom-based functional and adjunctive assessments: Proactive approaches to intervention selection for adolescents with Attention Deficit Hyperactivity Disorder. *Journal of Applied Behavior Analysis, 31*, 65-78.
- Evans, S. W., Allen, J., Moore, S., & Strauss, V. (in press). Measuring symptoms and functioning of youth with ADHD in middle schools. *Journal of Abnormal Child Psychology*.
- Evans, S. W., Axelrod, J. L., & Langberg, J. (2004). Efficacy of a school-based treatment program for middle school youth with ADHD: Pilot data. *Behavior Modification, 28*, 528-547.
- Evans, S. W., Dowling, C., & Brown, R. (in press). Psychosocial treatment of adolescents with attention deficit hyperactivity disorder. In K. McBurnett, L. Pfiffner, R. Schachar, G. R. Elliott, & J. Nigg (Eds.), *Attention Deficit Hyperactivity Disorder*. New York: Marcel Dekker.
- Evans, S. W., Pelham, W. E., & Grudberg, M. V. (1995). The efficacy of notetaking to improve behavior and comprehension of adolescents with Attention Deficit Hyperactivity Disorder. *Exceptionality, 5*, 1-17.
- Evans, S. W., Pelham, W. E., Smith, B. H., Bukstein, O., Gnagy, E. M., Greiner, A. R., et al. (2001). Dose-response effects of methylphenidate on ecologically-valid measures of academic performance and classroom behavior in adolescents with ADHD. *Experimental and Clinical Psychopharmacology, 9*, 163-175.
- Fabiano, G. A., & Pelham, W. E. (2002). Measuring impairment in children with attention-deficit hyperactivity disorder. *The ADHD Report, 10*, 6-10.
- Faraone, S., Biederman, J., & Monuteaux, M. C. (2002). Further evidence for the diagnostic continuity between child and adolescent ADHD. *Journal of Attention Disorders, 6*, 5-13.
- Findling, R. L., Short, E. J., & Manos, M. J. (2001). Developmental aspects of psychostimulant treatment in children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1441-1447.
- Horner, R. H., Sugai, G., Todd, A. W., & Lewis-Palmer, T. (2005). School-wide positive behavior support: An alternative approach to discipline in schools. In L. Bambara & L. Kern (Eds.), *Individualized supports for students with problem behaviors: Designing positive behavior plans* (pp. 359-390). New York: Guilford.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.
- Mitsis, E. M., McKay, K. E., Schulz, K. P., Newcorn, J. H., & Halperin, J. M. (2000). Parent-teacher concordance for DSM-IV attention-deficit/hyperactivity disorder in a clinic-referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 308-313.
- Molina, B. S. G., Smith, B. H., & Pelham, W. E. (2005). Development of a school-wide behavior program in a public middle school: An illustration of deployment-focused intervention development, stage 1. *Journal of Attention Disorders, 9*, 333-342.
- Molina, B. S. G., Pelham, W. E., Blumenthal, J., & Galiszewski, E. (1998). Agreement among teachers' behavior ratings of adolescents with a childhood history of attention deficit hyperactivity disorder. *Journal of Clinical Child Psychology, 27*, 330-339.
- The Psychological Corporation. (1992). *Wechsler Individual Achievement Test: Manual*. San Antonio, TX: Harcourt Brace.
- Raggi, V., Evans, S. W., Hackethorne, B., & Thompson, J. (2003, June). *Diagnostic efficiency of the Impairment Rating Scale*. Poster presented at ISRCAP: 11th Scientific Meeting, Sydney, Australia.
- Richard, M. M. (1995). Students with ADD in postsecondary education. In K. Nadeau (Ed.), *A comprehensive guide to attention deficit disorder in adults* (pp. 284-307). New York: Brunner/Mazel.
- Robin, A. L. (1990). Training families with ADHD adolescents. In R. Barkley (Ed.), *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment* (pp. 462-497). New York: Guilford.
- Robin, A. L. (1998). *ADHD in adolescents*. New York: Guilford.
- Rounsaville, B. J., Carroll, K. M., & Onken, L. S. (2001). A stage model of behavioral therapies research: Getting started and moving on from Stage I. *Clinical Psychology: Science and Practice, 8*, 133-142.
- Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M., & Schwab-Stone, M. (2000). NIMH diagnostic interview schedule for children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 28-38.
- Simpson, R. G. (1991). Agreement among teachers of secondary students in using the Revised Behavior Problem Checklist to identify deviant behavior. *Behavioral Disorders, 17*, 66-71.
- Smith, B. H., Pelham, W. E., Gnagy, E., & Yudell, R. S. (1998). Equivalent effects of stimulant treatment for Attention-Deficit Hyperactivity Disorder during childhood and adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*, 314-321.
- Smith, B. H., Waschbusch, D. A., Willoughby, M., & Evans, S. W. (2000). The efficacy, safety, and practicality of treatments for adolescents with attention-deficit/hyperactivity disorder (ADHD). *Clinical Child and Family Psychology Review, 3*, 243-267.
- Swanson, J. M., McBurnett, K., Christian, D. L., & Wigal, T. (1995). Stimulant medications and the treatment of children with ADHD. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 17, pp. 265-322). New York: Plenum.
- Wechsler, D. (1991). *WISC-III manual*. New York: The Psychological Corporation/Harcourt Brace Jovanovich.
- Weisz, J. R., Southam-Gerow, M. A., Gordis, E. B., & Connor-Smith, J. (2003). Primary and secondary control enhancement training for youth depression: Applying the deployment-focused model of

treatment development and testing. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 165-182). New York: Guilford.

**Steven W. Evans, Ph.D.**, is a professor of psychology at James Madison University and director of the Alvin V. Baird Attention & Learning Disabilities Center. His research interests include treatment development for adolescents with ADHD and school-based models of care.

**Joshua Langberg** received his M.A. degree in Psychological Sciences from James Madison University in May of 2002. He completed his doctoral coursework in the University of South Carolina's Clinical/Community Ph.D. program and is currently an intern at Duke University Medical Center. His research interests focus on the assessment and treatment of adolescents with ADHD.

**Veronica Raggi** completed her M.A. degree in the Psychological Sciences program at James Madison University and is currently completing a Ph.D. at the University of Maryland, College Park, Clinical Psychology Program. Her interests are in school-based interventions and parent training for children and adolescents with ADHD.

**Jessica Allen** completed her M.A. degree in the Psychological Sciences program at James Madison University. She is currently employed at the Northwest Community Services Board in Luray, Virginia.

**Elizabeth Buvinger, B.A.**, completed this work while a research associate in the Alvin V. Baird Attention and Learning Disabilities Center at James Madison University. She is currently in the Psychological Sciences Program at JMU working toward her Master's Degree. Her research interests are primarily in brain-behavior relationship in ADHD and Autism.