School Psychology Review,
2007, Volume 36, No. 2, pp. 256–273

Cumulative Benefits of Secondary School–Based Treatment of Students With Attention Deficit Hyperactivity Disorder

Steven W. Evans, Zewelanji N. Serpell, Brandon K. Schultz, and Dena A. Pastor
James Madison University

Abstract. School-based services are well suited for youth with chronic conditions who manifest much of their impairment in the school setting, such as youth with attention deficit hyperactivity disorder (ADHD). A significant literature on such programs at the elementary level exists, but little has been developed and tested at the secondary level. The current study is a 3-year treatment outcome study of a school-based training and consultation program for middle school youth with ADHD. Social and academic outcomes for program recipients were compared to that of participants in a community care control group. Findings revealed cumulative long-term benefits for the treatment group as measured by parent ratings of ADHD symptoms and social functioning. Although teacher and parent reports indicated no cumulative academic benefits, within-year analyses suggested a trend toward benefits in student grade point average. Implications for the treatment of ADHD in secondary schools are discussed.

School-based mental health professionals are increasingly called upon to embrace a public health perspective that includes meaningful linkages with mental health providers in the community and provides evidence-based treatments (Hoagwood & Johnson, 2003). Unfortunately, there are very few school-based evidence-based treatments and there are numerous implementation obstacles for those that do exist (Evans & Weist, 2004). School-based services are well suited for children with chronic conditions who manifest much of their impairment in school settings such as children with attention deficit hyperactivity disorder (ADHD). Although many studies have evaluated school-based treatments for elementary school children with ADHD (e.g., DuPaul & Stoner, 2003), very little information exists that can be used to guide the school-based treatment of secondary school students with the disorder (Wolraich et al., 2005). This article reports the findings of a 3-year treatment

This project was supported by a research grant from the Virginia Tobacco Settlement Foundation and funding from the Alvin V. Baird Attention & Learning Disabilities Center at James Madison University. The authors express appreciation to Liz Buvinger and Nillie Fallah for their help on this project as well as to Frank Gearing, Krista Swensson, Fred Kennon, the teachers and administrators in the Rockingham County Schools and Augusta County Schools, and the students and parents who participated in this project.

Correspondence regarding this article should be addressed to Steven W. Evans, Alvin V. Baird Attention & Learning Disabilities Center, MSC 9013, James Madison University, Harrisonburg, VA 22807; E-mail: evanssw@jmu.edu

outcome study of a school-based training and consultation program for middle school youth with ADHD.

ADHD Treatment in Secondary School

Presenting problems for children with ADHD typically include academic difficulties, discipline problems at school and at home, and conflict with peers. Adolescents with ADHD have many of these same problems, but often with more serious consequences, such as school dropout and legal problems. In addition, because of physical and social maturation, adolescents encounter new sets of problems such as automobile accidents, traffic tickets, difficulty in romantic relationships, vocational problems, and substance use or abuse. Furthermore, the school setting and related procedures and expectations change as students move from elementary to secondary schools. For example, less monitoring by a single teacher occurs in secondary schools. Factors like these complicate the provision of school-based services and the assessment of their effect (Evans, Allen, Moore, & Strauss, 2005). However, reports from the Multimodal Treatment of Children with ADHD (MTA) study and others suggest that it is a combination of psychosocial and medication treatment that is most likely to improve functioning of youth with ADHD (Conners et al., 2001). Furthermore, although medication has been found to provide short-term benefits for adolescents with ADHD (Evans et al., 2001), use drops precipitously throughout adolescence (Pelham, Molina, Meichenbaum, Gnagy, & Greenhouse, 2003). These reports underscore the need for continued development and evaluation of psychosocial treatments for secondary school youth with this disorder.

Only a few preliminary studies have evaluated school-based interventions for secondary school students. These studies have examined the benefits of academic interventions (Raggi & Chronis, 2006), self-monitoring (Gureasko-Moore, DuPaul, & White, 2006), and behavioral techniques (Ervin, DuPaul, Kern, & Friman, 1998). However, most sample sizes are small and none of the studies address a system for the provision of multiple interventions over time.

One approach that could facilitate treatment over time is school-based consultation. In fact, a study that sought to increase middle school teachers' knowledge about adolescents with ADHD and equip them with helpful interventions revealed that the most frequently requested services across 57 school districts were in-services to faculty and consultation about individual students (Shapiro, DuPaul, Bradley, & Bailey, 1995). In addition, positive outcomes were noted for the teachers in this study, including reports of an increase in knowledge about ADHD and adolescence as well as use of the information and interventions they had learned. The investigators did not measure individual student outcomes and noted that this would be an important next step in this line of research and treatment development.

We embarked on a process of developing and evaluating psychosocial interventions for middle school youth with ADHD in the form of an after school program called the Challenging Horizons Program (CHP). Preliminary findings suggested that the CHP effectively improves the social and academic functioning of many of students (Evans, Axelrod, & Langberg, 2004; Evans, Langberg, Raggi, Allen, & Buvinger, 2005). Yet, the feasibility of disseminating the program was questionable as schools would have to provide staff after hours and this would limit the ability of some districts to provide the service. Given the simplicity and potential transportability of many of the CHP interventions, we opted to create an alternate model of the CHP that could be implemented during regular school hours and staffed by educators, school-employed mental health professionals, and paraprofessionals. To this end, a community development team comprised of parents, teachers, school administrators, physicians, and the investigators collaboratively modified the CHP interventions and procedures (see Evans, Green, & Serpell, 2005) to create a training and *consultation* model of the CHP (CHP-C). This model reduces the program cost and was rated by school personnel as more acceptable and feasible than the after

school model of the CHP. The current study sought to assess whether the modified program (CHP-C) could effectively improve functioning and lead to symptom reduction.

Current Study

This study sought to compare social and academic outcomes for participants in a treatment group receiving CHP-C and a community care control group. Given that a chronic care model of treatment is recommended for individuals with ADHD, treatment participants in our study received the CHP-C psychosocial interventions for up to 3 school years (excluding the summer months). We explored whether participants would benefit from treatment in a cumulative manner, resulting in increasing differences between groups over time. That is, if treatment benefit is time limited, then our data would reveal a pattern of increasing differences between the groups within a school year followed by a return to baseline and equivalence between treatment and control groups at the beginning of the subsequent academic year. However, there is some evidence to suggest that psychosocial treatments may extend beyond the period of time that they are provided (MTA Cooperative Group, 2004), and we therefore anticipated increasing divergence of the treatment group from the control group over time.

Method

Participants

Participants were 79 young adolescents with ADHD, between the ages of 10 and 14 years ($M = 11.93$; $SD = 0.72$) at the start of participation, recruited when they were in the sixth grade from five middle schools in rural Virginia. Forty-two of these students attended schools randomly preassigned to treatment and 37 attended schools preassigned to the control condition. Ninety-four percent of the sample was Caucasian, which is consistent with the rural population living in these school districts. The final sample consisted primarily of boys (77%) and 50.6% of the sample lived in homes with a household income less than $40,000. Less than 10% of the sample lived in homes with household incomes greater than $80,000. Inclusion criteria required that participants: (a) meet *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; DSM-IV-TR; American Psychiatric Association, 2000) diagnostic criteria for one of the subtypes of ADHD; (b) have an IQ equal to or greater than 80 using the Kaufman Brief Intelligence Test (Kaufman & Kaufman, 1990); and (c) not meet diagnostic criteria for current or previous psychotic episodes (e.g., childhood onset schizophrenia) or pervasive developmental disorder. Diagnostic data as well as cognitive ability and achievement data at baseline are summarized in Table 1.

Procedures

Four public middle schools were initially recruited to participate in this study; two were randomly preassigned to the treatment group and two to the community care control group. As a result of difficulties reaching target recruitment numbers at control schools, an additional school from the same school district was recruited and assigned to control. Participants were recruited from the sixth-grade classes of these five schools in 2 consecutive school years, producing two cohorts; the first cohort was recruited in the fall of 2003, and the second in the fall of 2004.

Recruitment flyers requesting study participation of sixth-grade students with problems related to impulsivity, hyperactivity, or inattention were sent to all parents at participating schools. Caregivers who responded to the flyers were prescreened for eligibility using telephone administration of the ADHD items of the Disruptive Behavior Disorder Scale (DBD; Pelham, Evans, Gnagy, & Greenslade, 1992). If the phone screen indicated that the child met the prescreening criteria (parent endorsed six or more DSM-IV-TR criteria for either ADHD factor or indicated that child had a previous ADHD diagnosis), then an evaluation was scheduled for the parent and child. Before the evaluation, parents were asked to distribute behavior rating scales to each of their child's four core subject teachers.

Table 1
Diagnostic Results and Cognitive and Achievement Test Scores of Participants at Baseline

Diagnosis	Group		
	Control ($n = 37$)	Treatment ($n = 42$)	Total ($n = 79$)
ADHD			
Predominately inattentive	32.43%	40.48%	36.71%
Combined	67.57%	59.52%	63.29%
Comorbid conditions	70.27%	54.76%	62.03%
Cognitive ability: K-BIT FSIQ	$M = 105.62$	$M = 102.59$	$M = 104.01$
	$SD = 11.03$	$SD = 12.37$	$SD = 11.79$
Academic achievement: WIAT-II			
Word reading	$M = 100.68$	$M = 97.90$	$M = 99.20$
	$SD = 10.14$	$SD = 15.67$	$SD = 13.36$
Numerical operations	$M = 95.73$	$M = 92.86$	$M = 94.20$
	$SD = 15.71$	$SD = 14.19$	$SD = 14.90$
Spelling	$M = 97.49$	$M = 96.36$	$M = 96.89$
	$SD = 11.50$	$SD = 16.63$	$SD = 14.38$
Written expression	$M = 99.62$	$M = 94.40$	$M = 96.85$
	$SD = 13.27$	$SD = 15.35$	$SD = 14.56$

Note. ADHD = attention deficit hyperactivity disorder; K-BIT FSIQ = Kaufman Brief Intelligence Test Full-Scale IQ; WIAT-II = Wechsler Individual Achievement Test—II.

During the initial evaluation, parents and children signed informed consent and assent forms and trained graduate students administered the Diagnostic Interview Schedule for Children—IV (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000), which is a structured interview conducted with the parent to evaluate children for ADHD and other psychiatric disorders. A licensed clinical psychologist (SWE) incorporated data from this interview and from parent and teacher ratings to determine diagnosis based on the criteria listed in DSM-IV-TR. The Diagnostic Interview Schedule for Children—IV diagnoses were established using the standard scoring procedures for the instrument, and children met diagnostic criteria for ADHD if they met criteria according to the Diagnostic Interview Schedule for Children—IV and one or more teachers endorsed impairment consistent with symptoms related to ADHD (i.e., social, academic, or overall impairment on the Impairment Rating Scale [Fabiano et al., 2006] was a three or greater). Diagnoses of other disorders were based entirely on the results of the parent version of the Diagnostic Interview Schedule for Children—IV. The Kaufman Brief Intelligence Test was administered to obtain an estimate of verbal and nonverbal intelligence and the Wechsler Individual Achievement Test—II (The Psychological Corporation, 2001) was used to assess achievement. Once eligibility was confirmed, control participants were provided a list of local providers and summaries of their intake evaluation and told they were free to pursue treatments of their choice in their communities and schools. Treatment group participants received CHP psychosocial interventions at school provided by school staff, were monitored on a monthly basis, and were permitted to opt for a medication assessment.

All participants returned to the lab for follow-up evaluations every 6 months. Attri-

tion rates for the treatment group were 4.76% in Year 1, 7.5% in Year 2, and 12.20% in Year 3. For the control group, attrition rates were 12.5% in Year 1, 21.21% in Year 2 and 5.71% in Year 3. There were no significant group differences on data from any of the parent ratings between participants who dropped out and those who did not. This resulted in a sample size of 29 in the treatment group and 24 in the control group who completed the final evaluation.

Participants in both groups were asked about their involvement in treatment services (aside from CHP-C interventions) at each of the 6-month evaluations using the Services Use in Children and Adolescents—Parent Interview (SCA-PI; Hoagwood et al., 2004). The SCA-PI gathers parent report data about the services received by the child participant from schools, pediatric health care providers, juvenile justice, mental health, and day treatment across various modalities including psychosocial, medication, and education. Test–retest analyses over a 2-week period of time have indicated excellent agreement with an overall kappa for all services equal to .97 (Hoagwood et al., 2004). The data on medication use was included in the analyses.

Data from the SCA-PI were also used to examine possible overlap in services between participants in the treatment group and those in the community care control group. During the follow-up visits, between 17 and 40% of the control group reported receiving individual counseling since their last lab visit, whereas 29–63% reported receiving services at school (mostly tutoring). CHP-C interventions provided to the treatment group did not include individual counseling or tutoring—thus suggesting little, if any, overlap in services received by participants in the two groups.

Treatment: CHP-C

CHP-C included 15 psychosocial intervention sessions described in an 80-page treatment manual and interactive CD-ROM. Interventions primarily targeted academic skills, such as assignment tracking, note taking skills, and organization. The remaining interventions targeted social skills, such as social problem solving and conversation skills. CHP-C treatment manuals were presented and described to educators during 1-day trainings at each of the two treatment schools, held before the beginning of each school year.

During the school years, each participant in a treatment school was assigned an individual educator (referred to as a "mentor") who coordinated the CHP-C interventions for the student, starting at the time of their entry into the study, continuing through middle school, and ending at the point of their withdrawal or the end of the study (whichever came first). Educators who chose to fill the role of mentor did so voluntarily, but there were specific efforts to make informed matches based on prior student–educator relationships (e.g., a participant on the school wrestling team was mentored by his coach). Although all interventions were coordinated and, in most cases, implemented by mentors, some interventions were implemented by other teachers, counselors, or paraprofessionals. Participating teachers were primarily women (74%) with a mean age of 43.24 ($SD = 10.98$). The average years of teaching experience were 14.79 ($SD = 9.99$) and teachers reported teaching an average of 4.09 classes per day ($SD = 1.16$) with an average class size of 21.38 ($SD = 4.97$). Five respondents (6.3%) identified themselves as special educators. Most respondents (68.8%) reported having a bachelor's degree, whereas the remainder indicated having a master's degree or higher. There were no significant differences on any of these variables between teachers who volunteered to be mentors and those who did not.

To support the efforts of the mentors and other teachers, a certified school psychologist (BKS) provided ongoing school-based consultation. The school consultant was available to mentors at each treatment school an average of 8 hr per week throughout each school year. Much of this time was spent in direct consultation with various mentors (approximately four to five times per week), meeting with teacher teams, meeting with school counselors, participating in child study and individu-

alized education program meetings, and maintaining contact with educators and families by telephone and e-mail. The consultant used a problem-solving approach to consultation that involved forming cooperative partnerships with mentors, identifying specific mentor concerns, exploring intervention options, supporting mentors in implementing interventions, and then helping mentors modify interventions as needed. The consultant presented CHP-C intervention options in tentative terms and mentors had the option to accept or reject his recommendations. Preliminary analysis of data collected from mentors indicated that most agreed that they had adequate access to the school consultant (86%), that meetings with the consultant were helpful (88.4%), and that they were well informed and clear about their role in CHP-C (86.1%).

The CHP-C interventions selected most frequently by teachers targeted organization and social skills. Establishing contingencies, a frequent monitoring system, and prompts for the student to independently and accurately record assignments in the assignment notebook were the most frequently used interventions. Interventions targeting organization of binders, book bags, and lockers using regular monitoring and behavioral contingencies were also frequently provided. Other academic interventions used included study skills instruction and practice and note taking. Social skills interventions were administered by school counselors in small group weekly sessions targeting social problem solving, the development and practice of prosocial behavior, and anger or frustration management. Finally, some mentors provided parents with a training video and consulted with them on homework management strategies.

The initial plans for measuring adherence to treatment procedures included observation and product measures. However, given the complexity of schedules in middle schools, teachers wanted the flexibility to provide the interventions at their convenience and indicated that establishing set times, as would be necessary for observation, would greatly reduce the likelihood that interventions would be provided. Teachers also reported that the increased paperwork associated with product measures would similarly reduce the likelihood of implementation. We therefore adjusted our assessment of adherence and used teachers' self-reports in conjunction with ratings by the school consultant. Mentor teachers reported that between 87% and 89% of the interventions were implemented as described, although they rarely (34% of the time) used tangible rewards, such as snacks or privileges. School consultant ratings of teachers' adherence suggested less adherence than did the self-report measures. These ratings indicated that slightly less than 50% of the mentors were likely to implement more than two interventions at a time whether the child needed them or not (47.1%), make adjustments to the interventions based on parent or consultant feedback (41.4%), and implement the interventions as described in the manual (42.9%).

As mentioned previously, CHP-C also included a monthly monitoring system and medication algorithm. The monthly monitoring system used parent and teacher ratings to determine the need for additional treatment and to prompt treatment modifications guided by parental choice. Monthly ratings of functioning and impairment were collected from participants' four core subject teachers and parents via secure web-based scales or via paper and pencil scales mailed home or to the school. Return rates across schools each month ranged from 25% to 100% for teachers (modes: Year 1 = 100%, Year 2 = 100%, Year 3 = 100%) and 25% to 100% (modes: Year 1 = 100%, Year 2 = 50%, Year 3 = 40%) for parents. When monthly ratings showed the child had met a prespecified threshold indicating he or she might benefit from additional treatment, parents were contacted and informed that their child had met a "trigger" and they could opt to begin a medication trial or make an adjustment to the psychosocial interventions. If parents opted to adjust psychosocial treatment, the school consultant encouraged and supported mentors in making changes to psychosocial interventions. If parents opted to adjust medication, an assessment was provided by a local pediatrician who served as the participant's medical pro-

vider. This provider used a medication protocol that specified guidelines for medication trials and the use of weekly feedback from parents and teachers to assist with decision making about the most efficacious type and dose of medication for the participant. Once the medication trial was complete, the monthly monitoring system went back into effect and if the child met the trigger the next month, parents were again contacted for a treatment adjustment decision.

Outcome Measures

Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1993). Completed by parents at the initial evaluation and at the follow-up evaluations, the BASC is a comprehensive behavior rating scale that measures internalizing and externalizing behaviors. The standardized scores on the hyperactivity and inattention subscales were used for analyses.

Disruptive Behavior Disorders Rating Scale (DBD; Pelham et al., 1992). Administered at both the initial visit for diagnostic purposes and on a monthly basis to track ADHD symptoms, the DBD is a symptom rating scale completed by parents and participants' core subject teachers. The ADHD portion of the scale was administered monthly and includes 18 DSM-IV-TR symptoms of ADHD that parents or teachers indicate the degree to which each was present (0 = *not at all*; 1= *just a little*; 2 = *pretty much*; 3 = *very much*). The raw parent and teacher scores for hyperactivity–impulsivity symptoms and inattention symptoms were used in the analyses.

Grades. Participants' report card grades were used to calculate an overall grade point average (GPA) for each grading period from core subjects (science, math, history, and reading). GPA was based on a 5-point scale with $F = 0$ and $A = 4$.

Impairment Rating Scale (IRS; Fabiano et al., 2006). The IRS was completed by parents and teachers at the initial visit for diagnostic purposes and on a monthly basis to track functioning. It is a brief rating scale that assesses general functioning in various domains including relationships with peers, siblings, parents and teachers, academic functioning, self-esteem, effect on family or classroom life, and overall impairment. Parents or teachers indicate the degree to which functioning in these domains is or is not a problem that requires additional treatment. A 6-point visual response scale is used for each domain and high scores indicate greater impairment and need for additional treatment. A series of studies by Fabiano et al. (2006) have established evidence of concurrent, convergent and discriminant validity, and cross-informant reliabilities that were equivalent to or better than frequently used symptom scales. Parent ratings of social and academic functioning were used in the analyses along with teacher ratings of academic functioning and effect of problem behaviors on the classroom.

Social Skills Rating System (SSRS; Gresham & Elliot, 1990). The SSRS is a narrow-band, norm-referenced instrument that focuses on social interactions and is designed to screen and classify children and adolescents suspected of having social behavior problems. Parents completed this scale at the initial visit and at each 6-month follow-up visit. This is a widely used scale with strong psychometric properties that are reported in the manual. The total social skills standard score as rated by parents was used in the analyses.

Results

Parent ratings of hyperactivity–impulsivity and inattention symptoms from the DBD were used as measures of ADHD symptoms along with the hyperactivity and inattention *t* scores from the parent version of the BASC. Measures of social functioning used in the analyses included parent ratings of social impairment on the IRS, SSRS, and BASC. School functioning was included in the analyses using ratings of academic performance (parent and teacher) and classroom disruption (teacher) from the IRS, as well as GPA.

Each outcome measure was modeled separately using hierarchical linear modeling

Table 2
Estimated Intercepts and Slopes When Modeling Change in Parent-Rated Outcome Measures Across Years

		Intercepts		Slopes	
Outcome Measure	Scale	Control	Treatment	Control	Treatment
ADHD symptoms	BASC—Hyperactivity	67.96	62.62	−.002	−.008
	DBD—Hyperactivity	15.05	13.21	−.003	−.006
	BASC—Inattention	71.02	72.06	.002	**−.008**
	DBD—Inattention	17.91	17.80	−.001	−.005
Social functioning	IRS	2.38	2.78	.000	**−.001**
	BASC	38.71	40.68	.000	.001
	SSRS	83.56	87.53	.004	.008
School functioning	IRS—Parent	4.13	4.41	−.001	−.001

Note. Values are boldfaced if the treatment group estimate is significantly different than that of the control group. The intercept represents the average outcome measure on the initial measurement occasion while the slope represents the average change in the outcome measure for each one day increase in time. Slopes for outcome measures on different scales or using different units of time are not comparable. ADHD = attention deficit hyperactivity disorder; BASC = Behavior Assessment System for Children; DBD = Disruptive Behavior Disorders Rating Scale; IRS = Impairment Rating Scale; SSRS = Social Skills Rating System.

(see the Appendix for details regarding the hierarchical linear models used for the analyses). The use of hierarchical linear modeling allowed the intercept and slope parameters describing students' change in the outcome measures over time to be modeled as a function of group membership (control and treatment). This model specification resulted in unique intercept and slope parameters for the treatment and control groups and tests of whether the treatment group parameters were significantly different than those of the control group. The estimated intercepts and slopes of the treatment and control groups are shown in Table 2. Table 3 provides between-group effect sizes (*d*) for the estimated outcome measures across time. School and medication effects were examined as possible moderators or mediators of treatment outcomes. Medication use was assessed at each 6-month evaluation using the Services Use in Children and Adolescents—Parent Interview (Hoagwood et al., 2004). All medications approved for treating ADHD at any dosage were treated equivalently and coded as a time-varying dichotomous variable that indicated whether the participant took medication for at least half of each semester. No school or medication effects were found in any outcome variables, except where specifically indicated in the sections that follow.

ADHD Symptoms

The intercepts of the treatment and control groups were not significantly different from one another on either the hyperactivity or inattention parent ratings using the BASC or the DBD. Although the parent ratings of hyperactivity (BASC and DBD) and inattention (DBD) indicated a sharper decrease over time for the treatment group, the treatment group's change over time was not significantly different than that of the control group. Change over time in inattention ratings as measured using the BASC was significantly different for the two groups, with the control group increasing over time and the treatment group decreasing (improving). Figure 1 shows the slopes by group

Table 3
Between-Group Effect Sizes of Estimated Treatment and Control Group Means for Parent-Rated Outcomes

Outcome Measure	Scale	Denominator for Effect Sizes	Days Since Initial Measurement Occasion					
			0	180	360	540	720	900
ADHD symptoms	BASC—Hyperactivity	10.00	0.53	0.64	0.74	0.84	0.95	1.05
	DBD—Hyperactivity	6.75	0.27	0.35	0.42	0.50	0.57	0.65
	BASC—Inattention	10.00	−0.10	0.07	0.24	0.41	0.59	0.76
	DBD—Inattention	6.05	0.02	0.14	0.26	0.38	0.50	0.62
Social functioning	IRS	2.15	−0.19	−0.07	0.05	0.17	0.28	0.40
	BASC	10.00	0.20	0.21	0.21	0.22	0.23	0.24
	SSRS	15.00	0.27	0.31	0.36	0.41	0.46	0.51
School functioning	IRS—Parent	1.92	−0.14	−0.15	−0.15	−0.16	−0.16	−0.17

Note. Positive effect sizes imply higher functioning by the treatment group. Effect sizes (d') were calculated by dividing the difference between the means of the two groups by the standard deviation. For the BASC and SSRS, the standard deviations of the normative standard scores were used for the denominator of the effect sizes. For all other outcome measures, the standard deviation of scores at the initial measurement occasion pooled across groups was used. ADHD = attention deficit hyperactivity disorder; BASC = Behavior Assessment System for Children; DBD = Disruptive Behavior Disorders Rating Scale; IRS = Impairment Rating Scale; SSRS = Social Skills Rating System.

over time for parent BASC ratings of inattention. Between-group effect sizes were quite small during the first year of treatment, but benefits from treatment were emerging after the first year and continued to increase until reaching a moderate effect size of 0.76 after 2.5 years. A similar pattern was apparent on the three other symptom measures, although these differences in slope were not statistically significant.

Social Functioning

The intercepts of the treatment and control groups' parent ratings of social functioning from the IRS, BASC, and SSRS were not significantly different from one another. Groups did not differ over time in BASC parent ratings of social functioning, and both groups appeared to become increasingly impaired. Although not statistically significant, the data from the SSRS indicated a sharper increase (improvement) over time for the treatment group. Change over time in social functioning as measured using parent ratings on the IRS was significantly different for the two groups, with the control group exhibiting increasing impairment over time and the treatment group improving over time. As with parent ratings of inattention, effect sizes for parent ratings of social functioning on the IRS and SSRS showed consistent trends towards advantages for the treatment group, with final effect sizes in the moderate range (0.40 and 0.51, respectively).

When medication was entered into the models, a significant medication effect appeared for the IRS ratings, but only for the control group. Students in the control group not taking medication were rated as less impaired at the initial measurement occasion

than students taking medication. A school effect also appeared for social functioning measured using the IRS, with significant differences in the intercepts of the control schools. These schools, however, did not differ from one another in how they changed over time.

School Functioning

Academic functioning. The intercepts of the parent ratings of academic functioning did not differ between the groups nor did the slopes, but the trend over 3 years for both groups indicated a slight improvement (see Table 2). In contrast, the teacher rating data revealed conflicting findings. Improvement and declines in ratings contrasted greatly between teachers and within time periods with no interpretable trends in the outcomes. Similarly, there was considerable variability in the medication effects. For history and math teacher ratings of academic functioning, no significant medication effects were found. An effect for medication was found, however, for science teacher ratings of academic functioning, but only for the treatment group. Students taking medication in the treatment group were rated high by science teachers on the initial measurement occasion and decreased over time, whereas the opposite was true of nonmedicated treatment group students.

Classroom disruption. The teacher rating data for classroom disruption yielded no clear patterns. Although there were some statistically significant differences between the groups in intercepts and slopes, they were always unique to one set of teachers (i.e., subject area) and contrasting patterns were sometimes found in the data reported by other teachers. There were no discernible trends or significant effects. Similarly, there were a few significant effects of medication and school, but as with the academic functioning findings, the effects were unique to one set of teachers

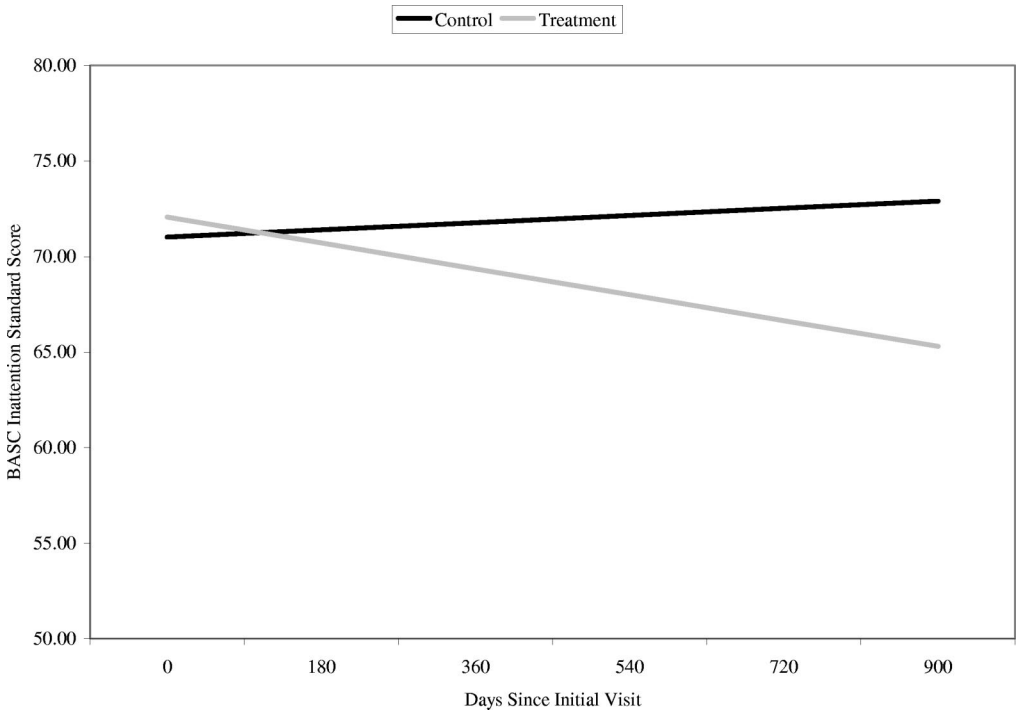

Figure 1. Estimated change over time in Behavior Assessment System for Children (BASC) parent ratings of inattention by group.

Table 4
Estimated Intercepts and Slopes When Modeling Change in Grades Within Year

	Sixth Grade		Seventh Grade		Eighth Grade	
Grade	Control	Treatment	Control	Treatment	Control	Treatment
Intercepts	2.25	2.63	2.22	2.52	1.99	2.61
First-semester slopes	−.18	−.21	−.16	−.12	.09	**−.27**
Second-semester slopes	−.06	.05	−.02	−.07	−.27	**.06**

Note. Values are boldfaced if the treatment group estimate is significantly different from that of the control group. The intercept represents the average grade point average at the first 6-week period while the slope represents the average change in grade point average for each 6-week increase in time.

and sometimes contradictory to reports from others.

Grades. The treatment and control groups did not differ significantly from one another in either their intercepts or slopes when change in GPA was modeled across years. As a result, within-year analyses were conducted. Analyses within year indicated that for grades during the sixth grade (representing average GPA after the first 6 weeks of school) the intercepts of the groups did not differ. Both groups decreased in GPA during the first semester and groups did not significantly differ in their first-semester slopes. In the second semester, the control group continued to decrease in GPA, whereas the treatment group increased in GPA (see Table 4). Although the groups did not differ significantly in their second-semester slopes, at the end of the year the mean grade for the treatment group was a C whereas that of the control group was a D. A similar pattern emerged for data from the eighth grade. That is, groups did not differ in their intercepts, but did differ significantly in their change in GPA during both the first and second semesters. In the first semester, the control group increased in GPA whereas the treatment group decreased. The opposite occurred in the second semester, with the control group decreasing in GPA over time and the treatment group increasing. The data from the seventh grade is different than data from the other two grades, indicating that the groups did not differ in their intercepts nor in their first-semester change in GPA and both groups demonstrated a decrease in GPA. Both groups continued to decrease in GPA during the second semester, with no significant difference between the groups in second semester slopes. Data from the GPA analyses are represented in Figure 2.

Analyses of the seventh-grade data also revealed significant school effects for the control schools. Specifically, the control schools differed significantly in their change in GPA during the first semester only, with one school increasing and the other school decreasing in GPA over time. Medication effects were also found for the control schools during the first semester, with medication associated with a lower intercept and lack of change in GPA.

Discussion

The primary finding of this study is that although some of the analyses did not reach the threshold of statistical significance, the pattern of effect sizes across the primary outcome variables presents persuasive evidence of a cumulative long-term benefit for young adolescents with ADHD receiving the school-based teacher consultation and training model of the CHP. Between-group effect sizes for parent ratings of inattention, hyperactivity, and social functioning were relatively small during the first year and increased over 3 years to the moderate range. Neither teacher nor

parent ratings of academic functioning revealed any cumulative academic benefits, despite the fact that CHP-C interventions focused primarily on school behavior and achievement. Similarly, no long-term benefit was found in student GPA; however, within-year gains in the second semester were found for treatment participants in the sixth and eighth grade. The finding of improved grades without improvements in teacher ratings is consistent with our past findings of poor inter-rater reliability between secondary school teachers on these rating scales (Evans, Allen et al., 2005), and contrasting and uninterpretable findings across teachers for the same students when measuring treatment response (Evans, Langberg et al., 2005).

Contrary to the benefits of combined treatment reported by the MTA investigators (e.g., Conners et al., 2001), psychosocial and medication treatment in the CHP-C did not interact to produce any advantages over non-medicated youth in the treatment condition. This finding is inconsistent with prevailing clinical wisdom that medication treatment can help individuals benefit from psychosocial treatments. Less than optimal medication dosing typical of community care (MTA Cooperative Group, 1999) may have contributed to the lack of a medication effect in this study. That is, when treatment group participants met the monthly threshold indicating a need for additional treatment, in 92% of the cases parents opted for adjustments in psychosocial interventions rather than medication trials. As a result, very few treatment participants received medication treatment that incorporated our medication treatment algorithm, and the vast majority of them received medication management resembling community care.

The cumulative benefits of the CHP-C on ADHD symptoms and social functioning is a potentially important finding, as it suggests that relatively small dosages of treatment provided over an extended period can lead to a cumulative benefit greater than what might be

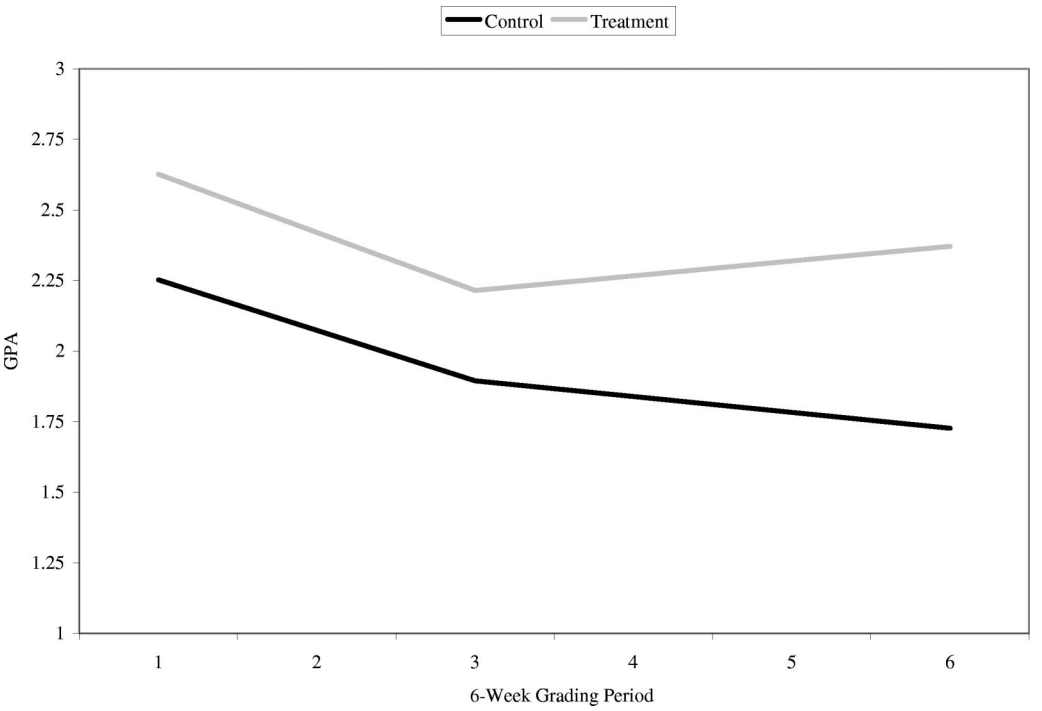

Figure 2. Estimated change over time in grade point average (GPA) by group within sixth-grade year.

obtained in the first few months of intervention. Persistence and patience over time may be as or more important than short-term intensive interventions. These results also have implications for intervention planning in schools. For example, individualized education programs and 504 plans (Section 504 under the Rehabilitation Act of 1973) that focus on accommodations rather than interventions are probably not in the best interest of students. Accommodations for adolescents with ADHD often take the form of reducing expectations. For example, Robin (1998) lists a series of commonly recommended accommodations for students with ADHD in secondary schools, a majority of which amount to a reduction in what is expected from the child, with little to no emphasis on improving behavior or performance (e.g., giving credit for late assignments, permitting extra time during tests). Our findings suggest that a rush to abandon ongoing interventions for accommodations might be premature.

Our findings that parent and teacher ratings of academic outcomes showed minimal response to treatment, despite the fact that the majority of the CHP-C interventions targeted school and academic functioning, were unexpected. However, one of the measurement problems confronted when assessing school functioning among adolescents is diminished parent and teacher knowledge of student performance, which results in highly inconsistent ratings, especially among teachers (Evans, Allen et al., 2005). Also, given the lack of teacher agreement, a consistent trend across school years is unlikely, because teachers change each year and cumulative benefits over time would require high rates of interrater reliability. The inconsistency in our teacher rating scale data provides further support for an emerging opinion that teacher rating data are of limited value in measuring change in school functioning among secondary school students. Many clinical researchers studying adolescents choose to not rely on these data (Barkley, Edwards, Laneri, Fletcher, & Metevia, 2001; Biederman, Wilens, Mick, Spencer, & Faraone, 1999; Conners et al., 1997). Investigators have reported that collecting rating scale data from secondary school teachers is also not helpful in the diagnostic process, despite the fact that impairment across settings is a criteria for diagnosis (Lahey et al., 1994), and practitioners tend to not use this information (Pliszka et al., 2000). An alternative method for using secondary school teacher ratings to assess outcomes is to focus assessment on behavior change within classroom and using teacher measures specific to the problems reported by the teacher in that classroom. This approach has been described by Evans and Youngstrom (2006) and may yield better information than we were able to obtain using the assessment procedures employed in this study.

An alternative measure of school functioning is grades. Our analyses indicated that although there was no cumulative benefit reflected in GPA, within-year analyses revealed an interesting pattern. In general, the first grading period of the year was the best for all participants. Group differences occurred in the second semester for both sixth- and eighth-grade participants, with treatment participants improving and their control counterparts declining. This pattern is similar to GPA outcome data reported for the after school model of the CHP (Evans et al., 2004; Evans, Langberg et al., 2005). To examine the potential meaning of these trends in GPA data, the percentage of total grades that were D or F were calculated by group. The treatment participants received a smaller percentage of D and F every semester over all three grades (19–34%) than the control participants (30–50%). Although randomization occurred within school district, it is certainly possible that there are school effects or a lack of equivalence between groups. Future studies using grades as outcome measures may consider gathering grade data from the year before participation in the project to help control for differences between schools. Although there are numerous factors that influence grades and compromise their value as a metric of school and academic functioning, grades have very high face validity. When parents assess how their children are doing in school, they are primarily concerned with grades. Furthermore, GPA

is widely used as criteria for admission to colleges and taking advanced courses. The grade data gathered in this study suggest that treatment may have had academic benefits, even if they were not apparent in parent and teacher ratings of academic functioning.

Limitations and Future Directions

There are several limitations to be considered when interpreting the findings of this study. First, attrition and sample size were limiting factors in this study and may have reduced the number of effects that reached statistical significance. Replication with larger samples is warranted. The sample used in this study was also not racially diverse, thus limiting generalizability; however, this work represents a population frequently ignored in the literature. Rural populations are rarely used in treatment studies despite representing a large portion of our population and experiencing rates of mental disorders comparable to those reported in urban areas while simultaneously facing reduced access to services (Kelleher, Taylor, & Rickert., 1992). It is important to point out that all of the interventions provided to participants in the treatment condition in this study were implemented by professionals employed by the school. The school psychologist who provided the training and consultation to the individuals providing the psychosocial interventions was employed by the research center. This school consultant did not directly provide services to the children or families. To the extent that consultation models are effective, they are valuable to school mental health systems because a large number of students may be served by an expert consultant. It was this potential cost–benefit advantage that led us to develop this consultation model of the CHP (Evans, Green, & Serpell, 2005) and future studies should measure costs and incorporate them into outcome analyses.

Future work must also consider the amount and type of services actually received by participants. In this study there was a great deal of variability in the services received by the participants in the treatment group and variations in dosage and types of treatment were attributable to many factors, including student and teacher variables. Untangling these factors and relating the resulting data to outcomes is an important consideration. Designs such as the one used in this study have been referred to as "randomized encouragement" (Frangakis & Rubin, 2002) for good reason. In the present study, the school consultant noted that some teachers went above and beyond what they were asked to do and offered many of the CHP interventions independent of the child meeting monthly triggers for needing additional treatment. Conversely, other teachers expressed both active and passive resistance to providing the interventions. Some teachers were very agreeable during meetings with the school psychologist, but subsequently did very little for students. Others refused and discouraged others from using the interventions with the students. For example, one teacher stated that because she had a master's degree in special education she knew that none of the CHP interventions were helpful. These types of active and passive resistance are sometimes not directly addressed, but they certainly pose a challenge to those working to improve mental health services in schools.

This variability in implementation points out the need for the consideration of fidelity and adherence data when examining outcomes. Although we gathered adherence data using teacher and consultant reports, improved methods for efficiently and accurately assessing adherence in the delivery of services in schools are sorely needed. Our data indicated that teachers believed they adhered to procedures at a much higher rate than the consultant perceived to be true. This could be owing to unclear communication of expectations by the consultant. Because the consultant had no authority over the teachers, his role was to encourage adherence and provide the teachers with options. Teachers may have assumed more freedom than we intended in how the services were provided. In addition, the differences between teacher and consultant reports may have been partly because of teachers' biases to present themselves favorably. This is an important challenge for future research on the development of this type of service and we

believe it may be a particularly challenging issue in secondary schools.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington DC: Author.

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.

Biederman, J., Wilens, T., Mick, E., Spencer, T., & Faraone, S. V. (1999). Pharmacotherapy of attention-deficit/hyperactivity disorder reduces risk for substance use disorder. *Pediatrics, 104*(2). Available at http://www.pediatrics.org/cgi/content/full/104/2/e20.

Carrig, M. M., Wirth, R. J., & Curran, P. J. (2004). A SAS macro for estimating and visualizing individual growth curves. *Structural Equation Modeling*, *11,* 132–149.

Conners, C. K., Epstein, J. N., March, J. S., Angold, A., Wells, K. C., Klaric, J., et al. (2001). Multimodal treatment of ADHD in the MTA: An alternative outcome analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40,* 159–167.

Conners, C. K., Wells, K. C., Parker, J. D. A., Sitarenios, G., Diamond, J. M., & Powell, J. W. (1997). A new self-report scale for assessment of adolescent psychopathology: Factor structure, reliability, validity, and diagnostic sensitivity. *Journal of Abnormal Child Psychology*, *25,* 487–497.

DuPaul, G. J., & Stoner, G. (2003). *ADHD in the schools: Assessment and intervention strategies* (2nd ed.) New York: Guilford Press.

Ervin, R. A., DuPaul, G. J., Kern, L., & Friman, P. C. (1998). Classroom-based functional assessment: A proactive approach to intervention selection for adolescents diagnosed with attention-deficit/hyperactivity disorder. *Journal of Applied Behavior Analysis, 31,* 65–78.

Evans, S. W., Allen, J., Moore, S., & Strauss, V. (2005). Measuring symptoms and functioning of youth with ADHD in middle schools. *Journal of Abnormal Child Psychology, 33,* 695–706.

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., Green, A., & Serpell, Z. N. (2005). Community participation in the treatment development process using assessment and community development teams. *Journal of Clinical Child and Adolescent Psychology, 34,* 765–771.

Evans, S. W., Langberg, J., Raggi, V., Allen, J., & Buvinger, E. C. (2005). Development of a school-based treatment program for middle school youth with ADHD. *Journal of Attention Disorders, 9,* 343–353.

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.

Evans, S. W., & Weist, M. D. (2004). Implementing empirically supported treatments in the schools: What are we asking? *Clinical Child and Family Psychology Review, 7,* 263–267.

Evans, S. W., & Youngstrom, E. (2006). Evidence-based assessment of attention-deficit/hyperactivity disorder: Measuring outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 45,* 1132–1137.

Fabiano, G. A., Pelham, W. E., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., et al. (2006). A practical measure of impairment: Psychometric properties of the Impairment Rating Scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child and Adolescent Psychology, 35,* 369–385.

Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics, 58,* 20–29.

Gresham, F. M., & Elliot, S. N. (1990). *Social Skills Rating System.* Circle Pines, MN: American Guidance Service.

Gureasko-Moore, S., DuPaul, G. J., & White, G. P. (2006). The effects of self-management in general education classrooms on the organizational skills of adolescents with ADHD. *Behavior Modification, 30,* 159–183.

Hoagwood, K. E., Jensen, P. S., Arnold, L. E., Roper, M., Severe, J., Odbert, C. et al. (2004). Reliability for the Services for Children and Adolescents-Parent Interview. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(11), 1345–1363.

Hoagwood, K., & Johnson, J. (2003). School psychology: A public health framework I. From evidence-based practices to evidence-based policies. *Journal of School Psychology*, *41,* 3–21.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test.* Circle Pines, MN: American Guidance Service.

Kelleher, K. J., Taylor, J., & Rickert, V. I. (1992). Mental health services for rural children and adolescents. *Clinical Psychology Review, 12,* 841–852.

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., et al. (1994). DSM-IV-TR field trials for attention deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry*, *151,* 1673–1685.

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). National Institute of Mental Health Multimodal Treatment Study of ADHD follow-up: 24-month outcomes of treatment strategies for attention-deficit/hyperactivity disorder. *Pediatrics, 113,* 754–761.

Pelham, W. E., Evans, S. W., Gnagy, E. M., & Greenslade, K. E. (1992). Teacher ratings of DSM-IIIR symptoms for the disruptive behavior disorders: Prevalence, factor analyses, and conditional probabilities in a special education sample. *School Psychology Review, 21,* 285–299.

Pelham, W. E., Jr., Molina, B. S. G., Meichenbaum, D., Gnagy, E., & Greenhouse, J. (2003, May). *Stimulant effects on long-term outcomes in ADHD individuals.*

Paper presented at New Clinical Drug Evaluation Unit meeting, Boca Raton, FL.

Pliszka, S. R., Greenhill, L. L., Crismon, M. L., Sedillo, A., Carlson, C., Conners, C. K., et al. (2000). The Texas Children's Medication Algorithm Project. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(7), 920–927.

Raggi, V. L., & Chronis, A. M. (2006). Interventions to address the academic impairment of children and adolescents with ADHD. *Clinical Child and Family Psychology Review, 9,* 85–111.

Raudenbush, S. W., & Bryk, A. G. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Reynolds, C. R., & Kamphaus, R. W. (1993). *Behavior Assessment System for Children.* Circle Pines, MN: American Guidance Service.

Robin, A. L. (1998). *ADHD in adolescents.* New York: Guilford Press.

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.

Shapiro, E. S., DuPaul, G. J., Bradley, K. L., & Bailey, L. T. (1995). *A school-based consultation program for service delivery to middle school students with ADD* (manual and videotape). Austin, TX: Pro-Ed.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal analysis: Modeling change and event occurrence.* New York: Oxford University Press.

The Psychological Corporation. (2001). *Wechsler Individual Achievement Test* (2nd ed.). San Antonio, TX: Harcourt Brace & Company.

Wolraich, M. L., Wibbelsman, C. J., Brown, T. E., Evans, S. W., Gotlieb, E. M., Knight, J. R., et al. (2005). Attention deficit hyperactivity disorder in adolescents: A review of the diagnosis, treatment and clinical implications. *Pediatrics, 115*(6), 1734–1746.

Date Received: August 7, 2006
Date Accepted: February 13, 2007
Action Editor: George DuPaul ■

Steven W. Evans, PhD, is the Alvin V. Baird Jr. Centennial Chair of Psychology and Director of the Alvin V. Baird Attention & Learning Disabilities Center at James Madison University. He is a faculty member in the Psychological Sciences program and works on developing and evaluating treatments for adolescents with ADHD.

Zewelanji N. Serpell, PhD, is a faculty member in the department of psychology and the associate director of the Alvin V. Baird Attention & Learning Disabilities Center at James Madison University. Her research focuses on the sociocultural factors in learning environments and interventions for children and adolescents with disruptive behavior disorders.

Brandon K. Schultz, EdS, is a school psychologist and Director of Clinical Services at the Alvin V. Baird Attention & Learning Disabilities Center. His research and practice focuses on school-based mental health services for adolescents with ADHD.

Dena A. Pastor, PhD, is a faculty member in the Center for Assessment and Research Studies at James Madison University. Her research focuses on applying latent variable models to examine student learning and development.

APPENDIX

HIERARCHICAL LINEAR MODELS USED TO EXAMINE GROUP DIFFERENCES OVER TIME

In equation form, the model examining group effects for all outcomes measures except GPA was specified as:

$$\begin{aligned} \text{Level 1.} \quad & Y_{jg} = \pi_{0g} + \pi_{1g} d_{jg} + e_{jg} \\ \text{Level 2.} \quad & \pi_{0g} = \beta_{00} + \beta_{01} group_g + r_{0g} \\ & \pi_{1g} = \beta_{10} + \beta_{11} group_g + r_{1g} \end{aligned} \tag{A1}$$

with Y_{jg} representing the outcome measure at time j for student g and with d_{jg} representing the number of days since the initial measurement occasion. One of the benefits of using hierarchical linear modeling with longitudinal data is that the number of time points can vary across individuals and the time points need not be equally spaced (Raudenbush & Bryk, 2002). In addition, all respondents can be used in the analysis, even those who supply data for only a single time point.

Before fitting hierarchical linear models to the data, the SAS macro developed by Carrig, Wirth, and Curran (2004) was used to determine if both linear and quadratic effects for time were needed to describe the overall trajectory as well as the individual student trajectories. Because the results of these analyses indicated that the overall and individual student trajectories were linear for all outcome measures, only linear time effects were included in the model.

Because the *group* variable used to predict the individual intercepts (π_{0g}) and slopes (π_{1g}) at Level 2 was coded 1 for the treatment group and 0 for the control group, β_{00} represents the overall intercept for the control group and β_{01} the difference of the overall treatment group intercept from the control group intercept. Similarly, β_{10} represents the overall slope (change in the outcome measure for each one-unit increase in time) for the control group and β_{11} the difference of the overall treatment group slope from the control group slope.

The variability of the Level 2 errors (r_{0g} and r_{1g}) indicates the extent to which the intercepts and slopes of the individual growth trajectories vary from person to person, once controlling for group membership. The distribution of the errors at Level 2 are typically assumed to be $N(0, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}$ typically equal to:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \text{var}\ (r_{0g}) & \\ \text{cov}\ (r_{0g}, r_{1g}) & \text{var}\ (r_{1g}) \end{bmatrix} \tag{A2}$$

Although the values of $\boldsymbol{\Sigma}$ were not essential to answering the research questions of the study, it was important to use a reasonable form for $\boldsymbol{\Sigma}$ because its specification does affect the precision of the parameters of interest (e.g., β_{01}, β_{11}; Singer & Willett, 2003). To this end, a series of nested models with different forms of $\boldsymbol{\Sigma}$ were fit to the data to find the most simplistic form of $\boldsymbol{\Sigma}$ that could be used without sacrificing model–data fit. All models had a Level 1 error variance–covariance structure equal to $\sigma_e^2\mathbf{I}$. Models were fit to the data specifying $\boldsymbol{\Sigma}$ to be different for each group and/or constraining certain elements in $\boldsymbol{\Sigma}$ to be equal to zero. The difference in the deviance statistics of nested models having different specifications of $\boldsymbol{\Sigma}$ were evaluated; restricted maximum likelihood estimation was employed as the differences between models that differed in only their random effects were being computed.

The estimated parameters of the most parsimonious models that emerged from the model comparison process were used to explain the effects of treatment. For the remainder of the appendix, these models are referred to as the *group* models. In the interest of space, the estimated random effects of the *group* models were not reported in this article, but are available from the authors upon request. The control group's estimated intercept and slope (β_{00} and β_{10}) and the treatment group's estimated intercept and slope ($\beta_{00} + \beta_{01}$ and $\beta_{10} + \beta_{11}$) are shown in Table 2. Of particular interest was the significance of β_{01}

and β_{11}, with the statistical significance of these parameters being evaluated using an $\alpha = .05$. Given the small sample size used in this study, it was particularly important to also assess the practical significance of the findings. To this end, the treatment and control group trajectories were plotted for each of the *group* models. In addition, an effect size (reported in Table 3) describing the standardized difference in group means was computed using predicted values from the *group* models.

To explore the extent to which medication had an effect on outcome measures, the time-varying predictor variable of medication and its interaction with time were added to Level 1 of the group models. At Level 2, the medication effects were specified as nonrandomly varying function of group membership:

$$\begin{aligned}
\text{Level 1.}\quad & Y_{jg} = \pi_{0g} + \pi_{1g} d_{jg} + \pi_{2g} meds_{jg} \\
& \quad + \pi_{3g}(d * meds)_{jg} + e_{jg} \\
\text{Level 2.}\quad & \pi_{0g} = \beta_{00} + \beta_{01} group_g + r_{0g} \\
& \pi_{1g} = \beta_{10} + \beta_{11} group_g + r_{1g} \\
& \pi_{2g} = \beta_{20} + \beta_{21} group_g \\
& \pi_{3g} = \beta_{30} + \beta_{31} group_g
\end{aligned} \tag{A3}$$

This model allowed us to explore if medication and its interaction with time was related to the outcome measures and if the nature of this relationship differed by group. Of particular interest was the significance of β_{20}, β_{21}, β_{30}, and β_{31}, with the statistical significance of these parameters being evaluated using $\alpha = .05$. Practical significance was assessed by examining the estimated model trajectories of the groups by medication status.

To ensure accurate estimation and interpretation of the school effects, three students in a single control school and two students who changed schools during the study were excluded when investigating the effect of school. Two treatment schools and two control schools remained, with the dummy-coded variables *trt.school* and *ctrl.school* representing affiliation with the second treatment and control school, respectively. These dummy-coded variables were included at Level 2 of the *group* model as predictors of individual intercepts and slopes:

$$\begin{aligned}
\text{Level 1.}\quad & Y_{jg} = \pi_{0g} + \pi_{1g} d_{jg} + e_{jg} \\
\text{Level 2.}\quad & \pi_{0g} = \beta_{00} + \beta_{01} group_g \\
& \quad + \beta_{02}(trt.school) \\
& \quad + \beta_{03}(ctrl.school) + r_{0g} \\
& \pi_{1g} = \beta_{10} + \beta_{11} group_g \\
& \quad + \beta_{12}(trt.school) \\
& \quad + \beta_{13}(ctrl.school) + r_{1g}
\end{aligned} \tag{A4}$$

These models allowed us to investigate if schools differed in their intercepts and slopes once controlling for group membership. Of particular interest was the significance of β_{02}, β_{03}, β_{12}, and β_{13}, with the statistical significance of these parameters being evaluated using $\alpha = .05$. Practical significance was assessed by examining the estimated model trajectories of the groups by school.

The only difference between the models investigating change in GPA from the model in Equation A1 is the use of the piecewise linear model. A piecewise linear growth model was used to model change in GPA during the first and second semesters separately. The model used to explore group difference in GPA over time was specified as:

$$\begin{aligned}
\text{Level 1.}\quad & Y_{jg} = \pi_{0g} + \pi_{1g} a_{jg} + \pi_{2g} a2_{jg} + e_{jg} \\
\text{Level 2.}\quad & \pi_{0g} = \beta_{00} + \beta_{01} group_g + r_{0g} \\
& \pi_{1g} = \beta_{10} + \beta_{11} group_g + r_{1g} \\
& \pi_{2g} = \beta_{20} + \beta_{21} group_g + r_{2g}
\end{aligned} \tag{A5}$$

with *a* coded as (0,1,2,2,2,2) and *a*2 coded as (0,0,0,1,2,3) to correspond to the six 6-week grading periods. Coding of time in this manner allowed π_{0g} to represent student *g*'s grade after the first 6 weeks of school, π_{1g} to represent their change in GPA each 6 weeks during the first semester, and π_{2g} to represent their change in GPA during each 6 weeks in the second semester. The methods previously described were used to find a proper form of $\boldsymbol{\Sigma}$ and to examine the group, medication, and school effects.