School Psychology Review,
2009, Volume 38, No. 1, pp. 14–27

SPECIAL TOPIC

Preventing Failure Among Middle School Students With Attention Deficit Hyperactivity Disorder: A Survival Analysis

Brandon K. Schultz and Steven W. Evans
James Madison University

Zewelanji N. Serpell
Virginia State University

Abstract. Middle school students with attention deficit hyperactivity disorder (ADHD) often experience academic decline within each school year, characterized by deteriorating grades from the first grading period to the last. This study reexamines whether the Challenging Horizons Program—Consultation Model (CHP-C) effectively reduces or delays academic failure. Report card data for two groups of students with ADHD—one that received the CHP-C and a control group that did not—were compared on whether and when grade point averages fell below passing. Results suggest that the CHP-C significantly reduced or delayed the onset of failure experiences in both sixth and seventh grades, even after the influence of student IQ was held constant. Practical implications of these findings are discussed.

Attention deficit hyperactivity disorder (ADHD) has been widely discussed in the professional literature, and it is estimated that more than 6,000 peer-reviewed articles have been devoted to the topic (Barkley, 2006, p. 76). Based on this extensive research base, ADHD is now recognized as a chronic disorder that often causes academic impairments

The authors thank Liz Buvinger and Nillie Fallah for their help on this project as well as 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. 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.

Correspondence regarding this article should be addressed to Brandon Schultz, Alvin V. Baird Attention and Learning Disabilities Center, MSC 9013, James Madison University, Harrisonburg, VA 22807; E-mail: schultbk@jmu.edu

that persist from elementary school into secondary school and beyond (American Psychiatric Association, 2000). However, little treatment research has focused exclusively on secondary school students with ADHD (Wolraich et al., 2005) and, as a result, little is known about the effectiveness of psychosocial interventions in addressing academic impairment in this population.

When examining academic impairment, researchers have relied on various measurement strategies. Recent reviews of the treatment literature suggest that the most commonly used indicators are student performance on standardized achievement tests, student performance on curriculum-based measurements (e.g., words read correctly per minute), and teacher rating scales (Raggi & Chronis, 2006; Trout, Lienemann, Reid, & Epstein, 2007). To a lesser extent, "universal measures" of school performance—including report cards, absenteeism, and suspensions or expulsions—have also been used for this purpose (Mattison, 2004, p. 359). Research examining universal measures has shown that, when compared to their peers, students with ADHD experience higher rates of suspensions, expulsions, and school dropout, as well as poorer report card performance and a reduced likelihood of postsecondary education (e.g., Barkley, Anastopoulos, Guevremont, & Fletcher, 1991; Klein & Mannuzza, 1991).

Among universal measures of academic outcomes, traditional A–F grading systems (hereafter referred to as "grades") are particularly appealing for research purposes because their widespread use and familiarity imparts face validity (Evans, Langberg, Raggi, Allen, & Buvinger, 2005). Moreover, evidence suggests that grades are a valid measure of academic impairment. For example, in a sample of middle school students in a special education setting, Mattison (2004) found that average grades (i.e., grade point average [GPA]) correlated more strongly with teacher ratings of student psychopathology than did other universal measures (e.g., disciplinary referrals, absences). In particular, a significant proportion (up to 25%) of the variance in grades was explained by teacher ratings of inattention and oppositional behaviors. Molina, Smith, and Pelham (2001) found similar relationships within a sample of middle school children in a general education setting and in another sample of children participating in a prospective study of ADHD. Specifically, findings from both samples suggested that teacher ratings of inattention and defiant behavior were significantly correlated with GPA. Molina and colleagues suggest that this relationship provides evidence for the ecological validity of GPA as a measure of ADHD-related academic impairment (p. 79).

However, grades are difficult to interpret because there are rarely objective measurement criteria across classrooms, teachers generally lack instruction in principles of measurement to consistently make valid judgments, and single-letter grades oversimplify student performances on myriad tasks (Allen, 2005). For such reasons, traditional grading practices are controversial within the field of education (cf. Kohn, 1999). From the perspective of school-based research, groupwise grade comparisons present several methodological challenges. First, it appears that grade outcomes are largely predicted by socioeconomic factors. For example, in a large study using a national sample of young adolescents, Mullis, Rathge, and Mullis (2003) examined the influence of family social networks (i.e., social capital) and socioeconomic status (i.e., resource capital) on students' grades. The results showed that indices of resource capital, such as income, parent education, and access to resources in the home (e.g., newspapers, study areas), significantly and positively predicted student grades. The authors concluded from these findings that children from families with greater resources are better equipped to benefit from educational opportunities than are children from less advantaged families.

Second, research suggests that student intelligence may also be an important predictor of academic performance. Specifically, student performance on standardized intelligence tests (IQ) has been shown to predict student achievement (e.g., Lassiter & Bardos, 1995) and student grades (Neisser et al., 1996). In terms of grades, Neisser and colleagues reviewed the relevant literature for a task force established by the American Psychological Association and estimated that the correlation between full-scale IQ and report card outcomes is .50. In other words, measured intelligence accounts for roughly 25% of the observed variance in grades.

Third, the interpretation of grade data is complicated by natural performance fluctuations over time. In our research we have found that secondary students with ADHD tend to exhibit a pattern of declining grades from the first to the last grading period in each school year. That is, without intervention, students' grades are typically strongest at the beginning of the school year, but then weaken to varying degrees by the end of the school year (Evans et al., 2005; Evans, Serpell, Schultz, & Pastor, 2007). Students who experience failure and do not receive adequate intervention can conceivably fall behind their peers in the curriculum, and this discrepancy can grow over time. Indeed, children with ADHD tend to exhibit increasing academic deficits when compared to their peers over several school years (e.g., Latimer, August, Newcomb, Realmuto, Hektner, & Mathy, 2003). Similar trends may occur within a school year, as it seems likely that once students fall behind it becomes increasingly difficult to recover as the classroom curriculum moves forward. Thus, the timing of failure events is consequential because the sooner the occurrence, the more likely intervention will be needed to ensure that the student benefits from instruction and ultimately earns credit for the class.

Given the difficulties in interpreting grade data, it is perhaps not surprising that researchers often report disappointing grade outcomes for otherwise effective interventions. In the case of school-based behavioral consultation, which appears to improve classroom strategies used by teachers (Sterling-Turner, Watson, & Moore, 2002) and is generally accepted by teacher consultees (Gilman & Gabriel, 2004), research suggests that there are little or no gains in the grades of students with ADHD (e.g., Evans et al., 2007; Jitendra et al., 2007; Kotkin, 1998). There are many potential explanations for this finding, but in our review of the literature it appears that grade data are inadequately assessed in school consultation studies. For example, researchers often select grade data from specific time points rather than the entire school year (e.g., Jitendra et al., 2007), which can result in equivocal findings depending on which time points are selected. Moreover, when parts of the grade trend are ignored, successful outcomes (e.g., prevention of periodic declines) can go unnoticed.

The Present Study

The aim of this study is to reexamine the effect of the Challenging Horizons Program, Consultation Model (CHP-C)—a school consultation program that targets the academic and social impairments associated with ADHD—on the grades of students with ADHD during their first 2 years in middle school. In a previous article (Evans et al., 2007), we reported that students who received the CHP-C did not significantly outperform students in a treatment-as-usual control condition in terms of grades, despite trends in the predicted direction. However, we did not account for the effects of student intelligence or family socioeconomic status on grade performance, nor did we specifically examine whether the program was effective at preventing failing grades. The threshold between passing and failing is critically important to stakeholders, and failure is likely to matter more to parents and youth than the difference between an A and a B, for example. Further, the timing of failure events is consequential to instructional deficits and the potential need for intensified intervention because of the cumulative nature of many classroom curricula.

Here, we present a survival analysis of the onset of failing GPA for the CHP-C participants, after controlling for student intelligence and family income. We hypothesized that, even after the covariates were held constant, students who received CHP-C consultation would experience fewer and delayed initial failure events when compared to the students who did not receive specialized school consultation.

Method

Participants

Seventy-nine students with ADHD between the ages of 10 and 14 years (*Mdn* = 11) were recruited from five rural middle schools in Virginia to participate in the CHP-C study. Boys represented 77% of the sample, which is consistent with the gender composition of ADHD children in clinic-referred samples (American Psychiatric Association, 2000). Most participants were Caucasian (93.7%), followed by Latino (3.8%); the remainder (2.5%) did not specify. Most families (73%) reported total yearly income of less than $60,000. In terms of educational background, 16 mothers (20.2%) reported a 4-year college degree or higher, 27 (34.2%) reported a 2-year college degree or equivalent, 28 (35.4%) reported a high school diploma, and the remaining 8 (10.1%) did not graduate from high school or complete an equivalent degree.

During intake, legal guardians provided informed consent, and then clinical evaluations verified a diagnosis of ADHD for each participant based on criteria set forth by the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition—Text Revision* (DSM-IV-TR; American Psychiatric Association, 2000). Our intake evaluation consisted of parent and teacher behavior and impairment rating scales, school records, and structured interviews with the primary caregiver. Final diagnostic decisions were based on parent responses to the Diagnostic Interview Schedule for Children, Fourth Edition (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000). Twenty-nine participants (36.7%) met the DSM-IV-TR criteria for ADHD Predominately Inattentive subtype and the rest of the sample (63.3%) met the criteria for ADHD Combined (inattention and hyperactivity/impulsivity) subtype. In addition, 36 (45.6%) of the students exhibited behaviors consistent with oppositional defiant disorder and another 11 (13.9%) exhibited behaviors consistent with conduct disorder. Per our exclusion criteria, there was no evidence of psychosis or post-traumatic stress disorder. According to parent reports, 29 participants (36.7%) were actively enrolled in special education services at their respective schools at the time of enrollment.

Our inclusion criteria required that all participants had an estimated full-scale IQ of 80 or higher (overall *M* = 104.0, *SD* = 11.8), as measured by the Kaufman Brief Intelligence Scale (Kaufman & Kaufman, 1990). We also examined academic achievement at intake using the abbreviated version of the Wechsler Individual Achievement Test, Second Edition (The Psychological Corporation, 2001), and these data (summarized in Table 1) were used for baseline comparisons of our study conditions (described in the next section).

Procedures

The CHP-C study used a two-wave cohort design, with participants entering the study in the 2003–2004 and 2004–2005 school years during their sixth-grade year. As a result, participants were enrolled at various times, and several students (*n* = 10) enrolled in the study in the seventh grade to replace participants who left the study during or after the sixth grade. Before the first wave of recruitment, two of the five participating schools were randomly assigned to be in the CHP-C treatment condition, where teacher consultees were encouraged to implement psychosocial interventions that have shown promise in an after-school setting (Evans, Axelrod, & Langberg, 2004). Fifteen specific interventions that target organization skills, study skills, behavioral self-monitoring, and social skills were outlined in a treatment manual. Teachers were taught these techniques in 1-day workshops before each school year, and a school psychologist (first author) provided ongoing consultation for approximately 8 hr per week at each treatment site throughout the school years. Consistent with best-practice recommendations for school consultation (e.g., Kratochwill, Elliott, & Callan-Stoiber, 2002), consultees in the CHP-C participated voluntarily. Most teacher consultees chose to focus their efforts on organizational interventions with

Table 1
Descriptive Variables for All Participants at Intake

	Treatment Group (n = 42)			Control Group (n = 37)		
	M	*SD*	Range	*M*	*SD*	Range
Demographics						
Participant age	11.8	0.6	10–14	12.0	0.8	10–14
Estimated family income	49,524	29,463	10–130[a]	39,329	21,622	10–70[a]
Number of siblings	1.2	0.9	0–3	1.2	1.0	0–3
IQ/Achievement						
K-BIT Full Scale IQ	102.6	12.4	81–132	105.6	11.0	81–130
WIAT-II Reading	97.9	15.7	46–133	100.7	10.1	85–123
WIAT-II Num Oper	92.9	14.2	55–124	95.7	15.7	46–135
WIAT-II Spelling	96.4	16.6	49–134	97.5	11.5	74–123
WIAT-II Writ Exp	94.4	15.4	48–126	99.6	13.3	72–126
DSM-IV-TR Symptom Counts[b]						
ADHD-IA Symptoms	8.1	1.1	6–9	8.2	1.0	6–9
ADHD-HI Symptoms	5.7	2.5	0–9	5.8	2.3	0–9
ODD Symptoms	3.6	2.4	0–8	4.3	2.3	0–8
CD Symptoms	1.0	1.5	0–6	1.0	1.7	0–7

Note. K-BIT = Kaufman Brief Intelligence Scale; WIAT-II Num Oper = Wechsler Individual Achievement Test, Second Edition Numerical Operations subtest; WIAT-II Writ Exp = WIAT-II Written Expression subtest; ODD = oppositional defiant disorder; CD = conduct disorder; DSM-IV-TR = *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text revision); ADHD-IA = attention deficit hyperactivity disorder inattention; ADHD-HI = ADHD hyperactivity/impulsivity
[a] The range of estimated family income is expressed in thousands.
[b] Symptom counts are based on structured interviews with the primary caregivers.

the support of the consultant. For example, all participating consultees in the sixth and seventh grades (n = 50) implemented organizational interventions targeting student assignment tracking, and some (28%) expanded these efforts to include the organization of bookbags and binders for students who needed the extra help. In addition, the consultant assisted school counselors during social skills groups targeting social problem-solving skills. Most (71%) of the CHP-C treatment condition students participated in the social skills groups at some point within their sixth- or seventh-grade years (see Evans et al., 2007, and Schultz & Cobb, 2005, for more details about the interventions and consultation strategies used within the CHP-C).

The three remaining schools were preassigned to a treatment-as-usual control condition. Initial comparisons of the treatment and control participants revealed no significant differences in age, estimated yearly family income, family size, IQ, academic achievement, or psychiatric comorbidities (based on symptom counts reported in structured caregiver interviews). These data are summarized in Table 1. As described in our previous article (Evans et al., 2007), there were no significant differences between the groups in gender composition, special education status, and medication use.

Analyses

Grades were collected for the four core courses (e.g., reading/English, math, science, and social studies) for all student participants, and GPAs were computed for each grading period, whereby A = 4.0, B = 3.0, C = 2.0, D = 1.0, and F = 0.0. Because our focus was prevention of failure, we coded failure events as grading periods during which GPAs dropped below 1.0. As such, failure was defined as grading periods for which most (if not all) core course teachers agreed that the student's academic performance was unsatisfactory.

To analyze differences in time-to-failure events between the treatment and control groups, we performed a *survival analysis* (cf. Kleinbaum & Klein, 2005; Singer & Willett, 1991). Survival analysis provides an estimate of the proportion of participants who experience a specific event over time. Initially, we summarized outcomes relative to the occurrence and timing of failure events using life table analysis; however, this approach does not adequately adjust for attrition (which is common in longitudinal studies), nor for the influence of covariates (Keiley & Martin, 2005). Thus, we also compared the two conditions using Cox regression analysis, which allows for a comparison of treatment conditions on the occurrence and timing of important events, while taking into account the role of covariates. Cox regression is a robust statistical procedure that can be used in many situations, as long as the *risk* of event associated with the levels of each covariate in the model is relatively consistent over time. This is referred to as the assumption of *proportional hazards* (Luke & Homan, 1998).

In the present study, we compared the proportions of students in the CHP-C treatment and control condition during which GPAs fell below 1.0 at each successive grading period while taking into account the influence of covariates (see next section). Grade data were available for six grading periods for all participants, based on the time that report cards were provided to families. Students who moved away from the area during either school year ($n = 3$) were removed from the analysis following the point of attrition (i.e., right-censored) because we were unable to gather complete report card data for these individuals.

Covariates

Previous research has shown that family income is the strongest predictor of adolescent grade performance among several indicators of family resources (Mullis et al., 2003). Based on this finding, we entered family yearly income as reported by participants' caregivers. In the CHP-C study, caregivers identified family income within $20,000 ranges (e.g., $40,000–$59,999) on a demographic questionnaire administered at intake, and the data were recorded as the median dollar amount for each category (e.g., $50,000). Three families (4%) did not respond to this item in the first year of the study and five families (8%) did not respond to this item in the second year of the study. We replaced missing family income data by imputing values based on the average family income at the mother's level of education because these variables were significantly correlated ($r = .45$, $p < .001$) within our sample. Family income data were updated before the start of the seventh grade year and original values were used when updates were not available. To account for family size, we then divided yearly income estimates by the total number of children living in the home.

In addition, we entered estimated student full-scale IQ as measured by the Kaufman Brief Intelligence Scale during intake. The grouping variable was entered into the model, followed by the covariates (student IQ and family income) in a hierarchical fashion. All statistical analyses were performed using the Statistical Package for the Social Sciences, version 15.0.

Results

To examine the relationships between failure experiences (i.e., whether student GPA dropped below 1.0 at any point in the school year) and the covariates, zero-order correlations were computed. Because failure events represent a continuous dichotomy, biserial correlation coefficients (r_b) were computed in these instances (Field, 2005, p. 132); otherwise, Pearson correlation coefficients were used. These results are summarized in Table 2. In general, it appeared that the correlations between failure events and the covariates were appreciably stronger for the control group than

Table 2
Zero-Order Correlations for Covariates and Failure Events by Study Conditions

	Treatment		Control		Total Sample	
	Failure Event	Student IQ	Failure Event	Student IQ	Failure Event	Student IQ
Sixth grade ($n = 69$)						
Income	−.11	.40*	−.28	.10	−.18	.29*
Student IQ	−.10		−.50*		−.18	
Seventh grade ($n = 59$)						
Income	−.25	.41*	−.30	−.08	−.25	.26
Student IQ	−.29		−.64*		−.29	

Note. Failure event was coded as follows: 0 = student grade point average (GPA) did not fall below 1.0; 1 = student GPA fell below 1.0. Estimated family income was adjusted for number of children in the home.
* $p < .05$.

the treatment group, with the strength of this relationship reaching statistical significance for student IQ in both the sixth and seventh grades. The correlation between student IQ and family income (adjusted for number of siblings) was nonsignificant for the control group, but significant for the treatment group in both sixth and seventh grades.

To assess the frequency of failure events across both conditions, life tables were computed. The results are provided in Tables 3 and 4 for the sixth- and seventh-grade data, respectively. Life tables provide a summary of the events of interest within a data set, which in the present study is GPA dropping below 1.0 (i.e., academic failure). The hazard function is the rate at which students experienced failure for the first time during each grading period. The survival function in the life tables indicate the proportion of students who did not experience initial failure in each grading period, which in this context can be considered *academic survival*. By the end of the sixth grade, 77% of students in the CHP-C condition had survived academically, whereas only 59% of students in the control condition had survived. By the end of seventh grade, 74% of students in the CHP-C had survived, whereas only 48% of students in the control condition had survived.

To assess whether the differences between our conditions were significant and meaningful in terms of relative failure risk, we compared the survival profiles of the treatment and control participants using Cox regression. To test the assumption of proportional hazards, we visually inspected the log-minus-log profiles, plotted the influence residuals for each of the predictors in the model (Luke & Homan, 1998), and assessed the goodness of fit by testing the relationship between the residuals and survival time (Kleinbaum & Klein, 2005, p. 151). Using these procedures, we identified a significant outlier ($z = 4.11$) in the sixth-grade data. This case was a student in the control condition who received special education services for the majority of the day, with three of the four core courses taught by the same teacher. As a result, we removed this case from the analysis because our event criterion (failing GPA) was presumed to be derived from student performance in multiple, independent classrooms. No other students in the sample received resource room assistance for more than two core courses. Following the removal of this case, the log-minus-log plots, plots of the residuals, and goodness of fit tests suggested that the assumption of proportional hazards was met for both sixth- and seventh-grade analyses.

Table 3
Life Table Describing the Grading Period at Which CHP-C Participants First Earned Grade-Point Averages Below 1.0 in Sixth Grade

	Grading Period	*n* Entering Period	*n* With Failing GPA	*n* Censored at End of Period	Hazard Function	Survival Function
Treatment	1	40	1	0	.03	.98
	2	39	0	0	.00	.98
	3	39	2	0	.05	.93
	4	37	3	1	.09	.85
	5	33	1	0	.03	.82
	6	32	1	31	.06	.77
Control	1	29	2	0	.07	.93
	2	27	4	0	.16	.79
	3	23	2	0	.09	.72
	4	21	1	0	.05	.69
	5	20	3	0	.16	.59
	6	17	0	17	.00	.59

Note. CHP-C = Challenging Horizons Program—Consultation Model; GPA = grade point average; censored = cases removed from analysis because of attrition; hazard function = rate of spontaneous event occurrence (failure) within each grading period; survival function = cumulative proportion of students maintaining GPAs above 1.0 at each grading period.

The results of the analyses (summarized in Table 5) suggest that for the sixth-grade data, the first model examining only the treatment effect (CHP-C vs. control) accounted for a significant proportion of the variance in time-to-failure events, $\chi^2(1, N = 68) = 4.38$, $p = .036$. The second model, which added student full-scale IQ, resulted in a significant improvement in prediction over the crude model, $\chi^2(1, N = 68) = 4.85$, $p = .028$, but the third model, which included the estimate of family income, did not, $\chi^2(1, N = 68) = 1.01$, $p = .314$. In the second model, both treatment condition ($B = 1.38$, $p = .007$, hazard ratio = 3.97) and IQ ($B = -0.05$, $p = .027$, hazard ratio = 0.95) were statistically significant, suggesting that both the treatment and student IQ significantly predicted academic survival outcomes in the anticipated directions. Survival profiles for the CHP-C and control groups in sixth grade are provided in Figure 1, adjusted for the influence of student IQ.

For seventh-grade data, it appeared that the first model including only the experimental conditions explained a significant amount of the variance in time-to-failure events, $\chi^2(1, N = 59) = 4.06$, $p = .044$. The second model offered a significant improvement over the first model, $\chi^2(1, N = 59) = 6.29$, $p = .012$, but the third model was not an improvement over the second model, $\chi^2(1, N = 59) = 0.84$, $p = .360$. Similar to the outcomes for sixth-grade data, both treatment condition ($B = 1.40$, $p = .007$, hazard ratio = 4.06) and IQ ($B = -0.06$, $p = .012$, hazard ratio = 0.95) were statistically significant in the second model, suggesting that the CHP-C treatment and student full-scale IQ significantly affected academic survival in the predicted directions. Survival profiles for the CHP-C and control groups in seventh-grade are provided in Figure 2, adjusted for the influence of the student IQ.

Discussion

In this study, we examined the extent to which a school consultation program (CHP-C; Evans et al., 2007) prevented or delayed the

Table 4
Life Table Describing the Grading Period at Which CHP-C Participants First Earned Grade Point Averages Below 1.0 in Seventh Grade

	Grading Period	*n* Entering Period	*n* With Failing GPA	*n* Censored at End of Period	Hazard Function	Survival Function
Treatment	1	36	2	0	.06	.94
	2	34	2	1	.06	.89
	3	31	2	0	.07	.83
	4	29	0	0	.00	.83
	5	29	1	1	.04	.80
	6	27	1	26	.07	.74
Control	1	23	4	0	.19	.83
	2	19	3	0	.17	.70
	3	16	1	0	.06	.65
	4	15	2	0	.14	.57
	5	13	0	0	.00	.57
	6	13	1	12	.15	.48

Note. CHP-C = Challenging Horizons Program—Consultation Model; GPA = grade point average; censored = cases removed from analysis because of attrition; hazard function = rate of spontaneous event occurrence (failure) within each grading period; survival function = cumulative proportion of students maintaining GPAs above 1.0 at each grading period.

onset of academic failure among middle school students with ADHD. We extend our previous analysis of the CHP-C by examining a practical index of passing or failing in core classes (defined as GPA dropping below 1.0). Our results suggest that students who received school consultation exhibited significantly fewer and delayed failure events in sixth and seventh grade, even after the influence of student IQ was statistically controlled. Specifically, students who did not receive the treatment experienced roughly four times the hazard of initial failure events than the students in the CHP-C across both grades when the effect of student IQ was held constant, although the wide confidence intervals around the hazard ratio suggests that this estimate may be unreliable. The results also suggest that IQ plays a unique role in predicting which students are likely to experience a failure event when the effects of the CHP-C treatment were held constant, but the estimate of family income did not (see following section, Limitations).

The present outcomes have practical implications for both school-based intervention and outcomes research. In terms of intervention, our results support the use of school consultation strategies to prevent student failure experiences in middle school, and conceivably, this reduction may help students avoid long-term problems in high school. For example, grades in middle school have been shown to strongly predict student dropout in high school (Garnier, Stein, & Jacobs, 1997), especially if poor grade performance leads to grade retention (Alexander, Entwisle, & Kabbani, 2001). Hence, by preventing failure experiences in the short-term, school consultation programs may prevent more serious risks in the long-term.

In terms of school-based outcomes research, the present study supports the potential for analyzing universal measures of school performance using survival analysis, as previously argued by Willet and Singer (1991). Treatment outcomes that rely on universal

Table 5
Cox Regression Results for Sixth- and Seventh-Grade Time-to-Failure Data

	B	*SE*	Wald	Significance	Hazard Ratio	95% Confidence Interval	
Sixth grade							
Model One[a]							
Group	0.92	0.46	4.09	.043	2.52	1.03	6.17
Model Two[b]							
Group	1.38	0.51	7.23	.007	3.97	1.45	10.82
FSIQ	−0.05	0.02	4.92	.027	0.95	0.91	0.99
Model Three[c]							
Group	1.41	0.52	7.38	.007	4.11	1.48	11.39
FSIQ	−0.04	0.02	3.89	.049	0.96	0.92	1.00
Income	0.00	0.00	0.93	.336	1.00	1.00	1.00
Seventh grade							
Model One[d]							
Group	0.91	0.47	3.80	.051	2.48	1.00	6.17
Model Two[c]							
Group	1.40	0.52	7.26	.007	4.06	1.47	11.25
FSIQ	−0.06	0.02	6.30	.012	0.95	0.90	0.99
Model Three[f]							
Group	1.42	0.53	7.21	.007	4.14	1.47	11.67
FSIQ	−0.05	0.02	5.25	.022	0.95	0.91	0.99
Income	0.00	0.00	0.79	.373	1.00	1.00	1.00

Note. FSIQ = estimated full-scale IQ; Income = family yearly income divided by number of siblings; Wald = test of significance.
[a] Overall $\chi^2 = 4.38$, $p = .036$.
[b] χ^2 change = 4.85, $p = .028$.
[c] χ^2 change = 1.01, $p = .314$.
[d] Overall $\chi^2 = 4.06$, $p = .044$.
[e] χ^2 change = 6.29, $p = .012$.
[f] χ^2 change = 0.84, $p = .360$.

measures (e.g., grades, suspensions, dropout rates) are likely to be compelling for stakeholders—more so than subjective behavior ratings or abstract psychometrics—and survival techniques can provide a relatively straightforward method of analyzing these data. In the present study, we defined academic survival as GPA above 1.0, which most stakeholders will recognize as the threshold for passing. Of course, alternative definitions of academic survival could be logically defended. For instance, some researchers may define academic survival using a higher GPA, or on the basis of a failing grade in any single class. But by any definition, time-to-event comparisons using universal measures of school performance are potentially more meaningful than alternative measurement strategies and statistical approaches.

Limitations

The present study has several important limitations. First, given the small sample size and the homogenous nature of our sample demographics, our results may not generalize to diverse populations or to school districts with alternative grading policies. For example, it seems likely that students receiving special and regular education services are graded in differing ways, but our small sample size did not allow for an analysis of these differences. It is noteworthy that a student who received resource room assistance for the majority of core courses represented a significant outlier in our data. Thus, future research examining academic survival would likely benefit by stratifying survival profiles based on special education status. In the present study, power analyses using the observed hazard rates from the life tables suggested that although our sample size was sufficient to detect overall hazard ratios above 2.94 and 3.13 for sixth and seventh grades, respectively (at the 0.80 level), analyses of subgroups were untenable.

Second, it is clear that a large proportion of the observed variance in failure events was unaccounted for, suggesting that unexamined factors played a critical role. For example, it seems likely that socioeconomic indicators may contribute to the unexplained variance, but these factors are notoriously difficult to calculate (cf. Duncan & Magnuson, 2003). Our attempt to examine the influences of family socioeconomic status did not appear to account for any of the variance in failure outcomes after the effects of the CHP-C and student IQ were held constant. One possible explanation of this finding is that the appreciable (albeit nonsignificant) difference in estimated yearly family income between our two conditions at intake (refer to Table 1) was attributable to socioeconomic differences between communities, and the effect of family resources on grades may be largely relative to local socioeconomic norms. Additional research is needed to evaluate this hypothesis, but our findings seem consistent with that explanation.

Third, we were unable to adequately assess treatment integrity within the CHP-C. Certainly, teachers' competence when delivering the interventions varied, but for practical reasons, strategies for measuring treatment integrity in the CHP-C were balanced against teacher acceptability of the procedures (see Evans, Schultz, & Serpell, 2008). Because of this necessary compromise, we cannot assess the degree to which the CHP-C interventions were implemented as intended or compare the academic survival profiles for students who received interventions as intended versus those who did not. Limitations relative to treatment integrity are endemic in consultation studies (Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000), and when adequate treatment integrity is achieved, outcomes do not necessarily exceed those of typical consultation practices (Jitendra et al., 2007). Thus, it is still unclear which mediating

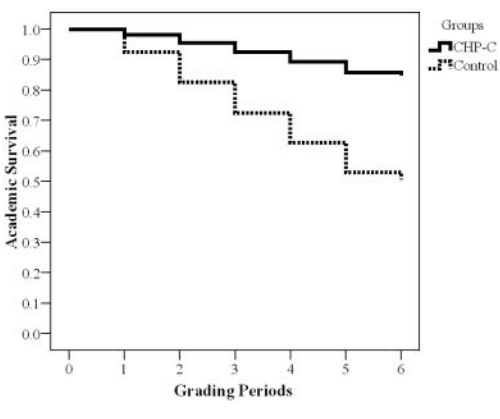

Figure 1. Academic survival profiles by study condition for sixth grade after controlling for the influence of student full-scale IQ.

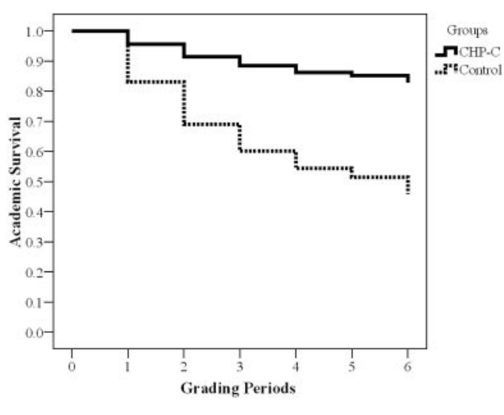

Figure 2. Academic survival profiles by study condition for seventh grade after controlling for the influence of student full-scale IQ.

processes are necessary for school consultation to be effective in improving the grades of students with ADHD. Unfortunately, the present study cannot shed light on whether the observed improvements in academic survival were attributable to specific components of the consultation-based interventions, or if the differences would be better explained by indirect effects, such as increased teacher awareness and understanding of student disabilities.

Conclusion

In our previous study of the CHP-C, our primary finding was that improvements in ADHD symptoms and impairment were relatively small in the first year of treatment, followed by increasing benefits over 2 or 3 school years when compared to students in a treatment-as-usual control condition. When grade data were analyzed, there were no significant long-term benefits associated with the CHP-C, but some positive short-term trends were noted (Evans et al., 2007). The findings from the present study suggested that when grades are analyzed from the standpoint of academic survival, the treatment effects were actually quite large. In both sixth and seventh grades, targeted school consultation appeared to significantly reduce the risk that a student with ADHD would earn a GPA below 1.0 (i.e., failing) when compared to treatment as usual. Thus, it appears that in addition to the gradual benefit of treatment on parent-rated measures of symptoms and impairment reported in the previous study, there are immediate benefits associated with *preventing* failure experiences. As discussed in the Discussion section, this difference is likely to be compelling to stakeholders, but in addition, the results suggest that grade outcomes in consultation studies are perhaps best analyzed by examining the relative academic survival profiles of participants, rather than comparing groups on continuous measures of GPA. In this manner, researchers can shift from assessing the unlikely possibility that psychosocial interventions will significantly improve grades over time, to a potentially more meaningful assessment of whether these interventions prevent failure.

References

Alexander, K. L., Entwisle, D. R., & Kabbani, N. S. (2001). The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record*, *103,* 760–822.

Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House, 78*(5), 218–223.

American Psychiatric Association. (2000). *Attention-deficit/hyperactivity disorder. Diagnostic and statistical manual of mental disorders* (4th ed., text revision, pp. 85–93). Washington, DC: Author.

Barkley, R. A. (2006). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment* (3rd ed.). New York: Guilford Press.

Barkley, R. A., Anastopoulos, A. D., Guevremont, D. C., & Fletcher, K. E. (1991). Adolescents with ADHD: Patterns of behavioral functioning, academic functioning, and treatment utilization. *Journal of the American Academic of Child and Adolescent Psychiatry*, *30,* 752–761.

Duncan, G. J., & Magnuson, K. A. (2003). Off with Hollingshead: Socioeconomic resources, parenting, and child development. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp. 83–106). Mahwah, NJ: Lawrence Erlbaum.

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., 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., Schultz, B. K., & Serpell, Z. N. (2008). Balancing feasibility and treatment adherence in school mental health research. *Advances in School Mental Health Promotion, 1,* 28–38.

Evans, S. W., Serpell, Z. N., Schultz, B., & Pastor, D. (2007). Cumulative benefits of secondary school-based treatment of students with ADHD. *School Psychology Review, 36,* 256–273.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage Publications.

Garnier, H. E., Stein, J. A., & Jacobs, J. K. (1997). The process of dropping out of high school: A 19-year perspective. *American Educational Research Journal, 34,* 395–419.

Gilman, R., & Gabriel, S. (2004). Perceptions of school psychological services by education professionals: Results from a multi-state survey pilot study. *School Psychology Review, 33,* 271–286.

Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research & Practice, 15,* 198–205.

Jitendra, A. K., DuPaul, G. J., Volpe, R. J., Tresco, K. E., Vile, R. E., Lutz, J. G., et al. (2007). Consultation-based academic intervention for children with attention

deficit hyperactivity disorder: School functioning outcomes. *School Psychology Review, 36,* 217–236.
Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test Manual.* Circle Pines, MN: American Guidance Service.
Keiley, M. K., & Martin, N. C. (2005). Survival analysis in family research. *Journal of Family Psychology, 19,* 142–156.
Klein, R. G., & Mannuzza, S. (1991). Long-term outcomes of hyperactive children: A review. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 383–387.
Kleinbaum, D. G., & Klein, M. (2005). *Survival analysis: A self-learning text* (2nd ed.). New York: Springer.
Kohn, A. (1999). Grading is degrading. *Education Digest, 65,* 59–64.
Kotkin, R. (1998). The Irvine paraprofessional program: Promising practice for serving students with ADHD. *Journal of Learning Disabilities, 31,* 556–564.
Kratochwill, T. R., Elliott, S. N., & Callan-Stoiber, K. (2002). Best practices in school-based problem-solving consultation. In A. Thomas & J. Grimes (Eds.). *Best practices in school psychology IV* (pp. 583–608). Bethesda, MD: National Association of School Psychologists.
Lassiter, K. S., & Bardos, A. N. (1995). The relationship between young children's academic achievement and measures of intelligence. *Psychology in the Schools, 32,* 170–177.
Latimer, W. W., August, G. J., Newcomb, M. D., Realmuto, G. M., Hektner, & J. M., Mathy, R. M. (2003). Child and familial pathways to academic achievement and behavioral adjustment: A prospective six-year study of children with and without ADHD. *Journal of Attention Disorders, 7,* 101–116.
Luke, D. A., & Homan, S. M. (1998). Time and change: Using survival analysis in clinical assessment and treatment evaluation. *Psychological Assessment, 10,* 360–378.
Mattison, R. E. (2004). Universal measures of school functioning in middle school special education students. *Behavioral Disorders, 29,* 359–371.
Molina, B. S., Smith, B. H., & Pelham, W. E. (2001). Factor structure and criterion validity of secondary school teacher ratings of ADHD and ODD. *Journal of Abnormal Child Psychology, 29,* 71–82.
Mullis, R. L., Rathge, R., & Mullis, A. K. (2003). Predictors of academic performance during early adolescence: A contextual view. *International Journal of Behavioral Development, 27,* 541–548.
Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77–101.
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*(2), 85–111.
Schultz, B. K., & Cobb, H. C. (2005). School consultation for middle school youth with ADHD: Lessons learned in the Challenging Horizons Program. *Report on Emotional and Behavioral Disorders in Youth, 5,* 91–99.
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.
Singer, J. D., & Willett, J. B. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin, 110,* 268–290.
Sterling-Turner, H. E., Watson, T. S., & Moore, J. W. (2002). The effects of direct training and treatment integrity on treatment outcomes in school consultation. *School Psychology Quarterly, 17,* 47–77.
The Psychological Corporation. (2001). *Wechsler Individual Achievement Test* (2nd ed.). San Antonio, TX: Harcourt Brace & Company.
Trout, A. L., Lienemann, T. O., Reid, R., & Epstein, M. H. (2007). A review of non-medication interventions to improve the academic performance of children and youth with ADHD. *Remedial and Special Education, 28*(4), 207–226.
Willett, J. B., & Singer, J. D. (1991). From whether to when: New methods for studying student dropout and teacher attrition. *Review of Educational Research, 61,* 407–450.
Wolraich, M., Wibbelsman, C. J., Brown, T. E., Evans, S. W., Gotlieb, E. M., Knight, J. R., et al. (2005). Attention-deficit hyperactivity disorder among adolescents: A review of the diagnosis, treatment, and clinical implications. *Pediatrics, 115,* 1734–1746.

Date Received: May 23, 2008
Date Accepted: October 27, 2008
Action Editor: Tanya Eckert ■

Brandon K. Schultz, EdD, NCSP, is Director of Clinical and Consultative Cervices at the Alvin V. Baird Attention and Learning Disabilities Center at James Madison University. He provided clinical assistance and school consultation services for the Challenging Horizons Program—Consultation Model. He is also Adjunct Professor in the undergraduate psychology department, and his research interests include school-based treatments for adolescents with disruptive behavior disorders.

Steven W. Evans, PhD, is the Alvin V. Baird Jr. Centennial Chair of Psychology and Director of the Alvin V. Baird Attention and 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 Assistant Professor of Psychology at Virginia State University, and former Associate Director of the Alvin V. Baird Attention and Learning Disabilities Center at James Madison University. Her research focuses on sociocultural factors in learning environments and school-based interventions for youth with disruptive behavior problems.