

Three-Year Course of Learning Disorders in Special Education Students Classified as Behavioral Disorder

RICHARD E. MATTISON, M.D., STEPHEN R. HOOPER, PH.D., AND LESLIE A. GLASSBERG, ED.D.

ABSTRACT

Objective: To investigate the 3-year course of learning disorders (LDs) and academic achievement in a sample of students with psychiatric disorders who were newly classified by the special education category of behavioral disorder (BD). **Method:** The occurrence of four definitions for LD (both discrepancy and low achievement) based on the WISC-R and the Woodcock-Johnson Psychoeducational Battery was followed in 81 students with BD from the time of their enrollment in BD classes to their first reevaluation after 3 years. Odds ratios (ORs) were used to measure stability of LDs in these students. **Results:** The prevalence of any LD was 64.2% at baseline and 61.7% at follow-up. Most of the 10 possible LD categories showed significant ORs, and the average OR was 21.9. At follow-up after 3 years, students both with and without LD at baseline had approximately the same achievement standard scores in reading and mathematics, but a significantly lower score for written language. Standard scores for the students without LD consistently were significantly higher than the scores for students with comorbid LD. **Conclusion:** LDs in this unique sample of students with psychiatric disorders remained common and generally stable over the first 3 years. *J. Am. Acad. Child Adolesc. Psychiatry*, 2002, 41(12):1454–1461. **Key Words:** learning disorders, longitudinal stability, special education.

When child and adolescent psychiatrists consult to schools, they often face the issue of whether a student's academic dysfunction is secondary to a psychiatric disorder, a learning disorder (LD), or a combination of both. School personnel may too quickly ascribe a child's poor academic performance solely to a psychiatric disorder if a student shows noteworthy behavioral or mood symptoms and may thus overlook the possibility of comorbid LD. This question is of particular importance in a common group for consultation—special education students categorized as having emotional/behavioral disorders (EBD). LD is common in this group of students (53.2%) and often goes unclassified (Glassberg et al., 1999), which quite likely contributes in many of these students to their well-established poor academic, graduation, and employment

outcome (Wagner, 1995). Consequently, consultants must be prepared to help their school colleagues better appreciate comorbidity of LD in students with psychiatric disorders, including the long-term persistence of LD and the effects of special education programming on academic dysfunction in such students.

Although comorbidity between psychiatric disorders and LD is well recognized, our knowledge of the detailed course of LDs in children who also have psychiatric disorders is not substantial (American Academy of Child and Adolescent Psychiatry, 1998). Longitudinal studies of the course of psychiatric disorders (with or without treatment) have seldom included the initial percentages of specific LDs or subsequent interaction effects. Reviews of the natural course of LDs in general child populations or in samples of LD children (Kavale and Forness, 1995; Satz et al., 1998) suggest that LDs may be persistent and also negatively affected by the presence of psychiatric disorders, but reviewers caution that this longitudinal research is compromised by methodological shortcomings, particularly the definition of LD. Consequently, the prospective course of LDs in children with psychiatric disorders remains quite unclear. Furthermore, the outcome of LDs in students with EBD has never been studied.

Accepted June 25, 2002.

Dr. Mattison is with the State University of New York at Stony Brook, Dr. Hooper is with the University of North Carolina at Chapel Hill, and Dr. Glassberg is with the Special School District in St. Louis.

Correspondence to Dr. Mattison, Putnam Hall-South Campus, State University of New York, Stony Brook, NY 11794-8790; e-mail: richard.mattison@stonybrook.edu.

0890-8567/02/4112-1454©2002 by the American Academy of Child and Adolescent Psychiatry.

DOI: 10.1097/01.CHL.0000024865.60748.D4

Therefore, the primary goals of this study were to follow the course of specific definitions of LD as well as changes in academic achievement over the first 3 years of programming in a large epidemiological sample of students newly classified by the special education category of behavioral disorder (BD, a regional equivalent of EBD; Glassberg et al., 1999). We anticipated that the LD definitions would remain common and stable in this unique group of students. Furthermore, we wished to investigate the impact of the services on longitudinal achievement in BD students both with and without LD, and in BD/LD students who did or did not receive supplemental LD programming.

METHOD

Subjects

Glassberg et al. (1999) have determined the prevalence rates for four general definitions of LD in a suburban St. Louis epidemiological sample of 233 students aged 6 to 16 years. These pupils were classified for the first time between 1985 and 1989 by the special education category BD (the Missouri category that is similar to the federal special education classification of EBD). The parents of these original participants were contacted by letter in the spring of 1995 for permission to examine their children's files that summer. These records were kept by the special education agency that provided services to the students and approved of the project. The intent was to investigate the test data that existed from the routine 3- and 6-year reevaluations conducted by the special education agency for as many of the children as possible.

The original number of 233 participants was first reduced to 189 because 4 parents refused permission and 40 letters were returned because of unknown address. Before the first 3-year reevaluation, another 57 students had discontinued the agency's services (primarily moved away). After examining the files of these remaining 132 students (who also continued with the BD classification), the decision was made to analyze only the 3-year data for those 81 students who had been retested with the *same* IQ and achievement instruments used in the original study. This step allowed the preservation of the *identical* four LD definitions at both time points. (Although the original intent of the authors was also to examine the 6-year reevaluations, the 6-year test data proved insufficient to investigate for the 64 students who remained at that time.)

Baseline characteristics were compared between the final 81 students (the focus of the current report) and the original 152 students who were not investigated. No statistically significant differences were found between these two groups for the following baseline variables: sex, race, IQ (Verbal, Performance, and Full Scale), achievement standard scores (reading, mathematics, or written language), or prevalence for any of the four general LD definitions. Only two significant baseline differences were determined. The final group was significantly younger than the excluded group (8.6 ± 2.2 years versus 10.5 ± 2.7 years, respectively; $t_{233} = -5.82, p = .0001$), and the presence of any LD definition was significantly greater in the final group than in the excluded group (64.2% versus 47.4%, respectively; $\chi^2 [1, N = 233] = 6.01, p = .01$). Thus, in general, the 81 participants appeared representative of the original group of participants, and their greater per-

centage of any LD was advantageous for follow-up. Their demographic characteristics (Table 1) show that they are primarily male and the small majority are white.

The special education programming received by the subjects during the 3 years reflected the standard services that students with BD received nationally at that time. Most of the children were taught in small, structured classrooms by both special education teachers and aides, using behavioral modification and individualized curricula and instruction. Services to address specific LDs were provided to those students who were also classified as LD. The settings ranged from regular public schools (inclusion, resource rooms, or self-contained classrooms) to separate specialized public schools. The staff was supported by social workers, school psychologists, and other special education personnel.

Psychiatric diagnoses were not consistently available for the subjects because neither mental health evaluations nor psychiatric diagnoses were required by the state at that time as part of a student's evaluation for classification as BD. Behavioral checklists were not routinely available because an objective measure of psychopathology was also not required for the initial assessment. However, previous research has shown that the most frequent psychiatric diagnoses in students newly classified as EBD are attention-deficit/hyperactivity disorder (ADHD), followed by conduct/oppositional disorders and depressive disorders (Mattison and Forness, 1995).

Instruments

The 81 subjects with BD were tested at both baseline and the 3-year reevaluation (mean length of follow-up = 35.6 ± 4.9 months) by examiners of the special education agency with the WISC-R (Wechsler, 1974) and with at least one achievement test (age-based standard score for broad reading, mathematics, and/or written language) of the Woodcock-Johnson Psychoeducational Battery Test of Achievement (WJPEB) (Woodcock and Johnson, 1978). All 81 participants were tested with the WISC-R at both times. The following numbers of students had WJPEB testing in reading, mathematics, and written language, respectively: baseline (80, 77, and 66) and 3-year (72, 72, and 62). More specifically, 72 students had testing in reading for both times, 68 in mathematics, and 51 in written language.

TABLE 1
Baseline and Three-Year Reevaluation
Characteristics of the 81 Subjects

	Baseline	3-Year	<i>r</i>
Demographics			
Mean \pm SD age (yr)	8.6 ± 2.2	11.5 ± 2.2	—
Male (%)	86.4	—	—
White (%)	55.6	—	—
WISC-R IQ (mean \pm SD)			
Verbal	98.6 ± 15.1	96.4 ± 15.7	0.77
Performance	101.8 ± 14.3	101.2 ± 16.0	0.76
Full Scale	100.1 ± 14.4	98.6 ± 15.7	0.81
WJPEB standard scores (mean \pm SD)			
Reading ($n = 72$)	91.9 ± 12.9	92.0 ± 13.6	0.72
Mathematics ($n = 68$)	90.4 ± 13.3	89.7 ± 13.5	0.69
Written language ($n = 51$)	93.9 ± 13.3	90.4 ± 13.5^a	0.67

Note: All Pearson correlations are highly significant ($p < .0001$). WJPEB = Woodcock-Johnson Psychoeducational Battery Test of Achievement.

^a Baseline > 3-year, $F = 5.40, p = .02$.

Definitions of Learning Disability

Four general LD definitions were originally selected to be consistent with conceptualizations of LD that were current at that time, and these are described in greater detail elsewhere (Glassberg et al., 1999). The first definition was the Missouri *state definition* of LD, an ability-achievement discrepancy that was based on a regression model; LDs for specific achievement domains were not available. The second definition was *low achievement*—one or more achievement test with a standard score less than 78 and, therefore, at least 1.5 SD below the mean.

The final two definitions also used ability-achievement discrepancy. One definition (termed *simple discrepancy* in this report) required more than 22 points difference (i.e., at least 1.5 SD) between the WISC-R Full Scale IQ and the standard score of at least one WJPEB achievement domain. The other (termed *regression*) used the Reynolds regression formula (Reynolds and Stanton, 1990) to adjust for correlations between the WISC-R and the WJPEB while calculating the significance of any discrepancy. This last definition required a discrepancy of 1.5 standard errors or more of prediction to exist between actual achievement on any of the three tests and the expected level of achievement predicted by the Full Scale IQ.

No low IQ cutoff was used because the authors judged that several intrinsic factors could have influenced IQ scores in this unique population of students with BD, e.g., poor cooperation during testing whether from test avoidance, frustration, oppositionality, or limited attention. However, no subject had originally received the additional classification of mental retardation.

Data Analysis

First, the continuity of the LD definitions was investigated with odds ratios (ORs), which have rarely been used in LD longitudinal research. The ORs compared the odds of a specific LD definition occurring at the 3-year time point in students diagnosed with that same definition at baseline, relative to the odds of that definition occurring at the 3-year time point in students who did not meet that definition at baseline.

Then, the continuity of IQ and achievement scores over time was investigated with repeated-measures analysis of variance (ANOVA) (group \times time) for the students who met no LD definition at baseline and for the students who met at least one LD definition at baseline. Next, changes in IQ and achievement scores were compared with repeated-measures ANOVA (group \times time) between the students who were additionally classified at baseline as LD by state criteria and consequently received LD services, and the students who were not classified as LD by state criteria (and thus received no LD intervention) but did meet another study LD definition at baseline. Continuous variables were analyzed between groups with the Student *t* test, while dichotomous variables were analyzed with the χ^2 (or Fisher exact test if a cell contained five or fewer subjects). The accepted level of confidence was $p < .05$.

RESULTS

Stability of LD Definitions Over Three Years

First, Table 1 shows that the mean IQ scores for all 81 students were in the average range at both time points. The mean standard scores of the achievement tests were similar at both time points and in the low average–average range, but written language was significantly lower at fol-

low-up. The achievement scores were approximately 6 to 10 points lower than the mean Full Scale IQ scores at both times, with little difference among the three achievement domains. The Pearson correlations over 3 years for both instruments were strong ($r > 0.50$; Cohen, 1988), highly significant ($p \leq .0001$), and quite similar within each instrument.

The prevalence rates of subjects who met at least one LD definition were 64.2% at baseline and 61.7% at 3 years. Overall, 65.4% of the subjects who met any LD definition at baseline continued to meet an LD definition at follow-up. Table 2 provides a more specific breakdown according to each achievement area. Mathematics was the subject area for which students most frequently met criteria for LD. Both at baseline and follow-up, about 40% of the students met criteria for LD in mathematics according to at least one definition.

The investigation of the stability of specific LD definitions generally showed significant ORs (with a lower confidence interval > 1.00) in Table 2. Overall, the average OR among the 10 definitions was 21.9. Each achievement area showed two of three definitions with an OR of at least 10 (generally significant at $p < .0005$ or better). Overall, the reading definitions had the strongest

TABLE 2
Stability of LD Definitions From Baseline to Three-Year Reevaluation

LD Definition	Occurrence (%)		OR	95% CI
	Baseline	3-Year		
Missouri Reading ($n = 72$)	22.2	29.6	12.3†	3.6–41.6
Low	15.3	13.9	78.7†	11.4–545.0
Discrepancy	9.7	9.7	11.4*	1.9–69.7
Regression	29.2	20.8	21.3†	5.0–91.1
Any reading	34.7	25.0	13.7†	3.8–49.9
Mathematics ($n = 68$)				
Low	19.1	14.7	10.9***	2.5–48.6
Discrepancy	14.7	11.8	18.3***	3.4–100.3
Regression	27.9	39.7	2.8	1.0–8.4
Any math	39.7	41.2	2.7	1.0–7.4
Written language ($n = 51$)				
Low	11.8	19.6	40.0***	3.9–415.1
Discrepancy	7.8	19.6	17.1*	1.6–189.2
Regression	27.4	33.3	6.5**	1.7–25.0
Any written	31.4	39.2	6.4**	1.7–23.3

Note: low = low achievement; discrepancy = simple discrepancy; LD = learning disorder; OR = odds ratio; CI = confidence interval. * $p < .05$; ** $p < .01$; *** $p < .0005$; † $p < .0001$.

continuity, including two definitions with highly significant ORs greater than 20. Indeed, the strongest specific LD was low achievement in reading.

The findings for the averages of the general LD definitions across achievement areas also showed good stability. The low achievement definition was the strongest, averaging an OR of 43.2. Simple discrepancy was the next strongest definition with an average OR of 15.6, followed by the regression definition with an average of 10.2. However, within the three subject areas, the general definition with the highest stability varied. For reading and written language, the low achievement criteria yielded the highest stability, and the regression criteria yielded the next highest. For mathematics, the simple discrepancy criteria yielded the highest stability.

Three-Year Changes in the Psychoeducational Characteristics of BD Students With and Without LD at Baseline

The IQ and achievement characteristics at both baseline and follow-up were compared between the 29 BD children without any LD at baseline and the 52 BD children with any LD at baseline (Table 3). The psychoeducational characteristics for both groups of students generally remained constant. The 29 BD students without LD at baseline remained essentially average over the two time points, and their mean achievement scores were all significantly greater than the mean achievement scores for the LD students at both times. Using repeated-measures

ANOVA (group \times time), the mean Verbal IQ for the 52 BD students with LD was significantly lower at follow-up, as consequently was their mean Full Scale IQ. Their mean achievement scores showed less decrease, but remained 9 or more points lower than their mean Full Scale IQ. Thus neither group was found to have a significant change in any achievement area.

Three-Year Psychoeducational Changes in Students With Both BD/LD Who Received LD Intervention

The impact of LD services was investigated in the 52 students with both BD and LD at baseline. Eighteen of these comorbid participants were originally classified as LD according to state criteria, and thus consequently received services for their LD according to their individual educational plans. Repeated-measures ANOVA (group \times time) was used to compare changes in psychoeducational characteristics over the 3 years between this group and the other 34 BD/LD students who at baseline only met non-state LD definitions and thus did not receive LD services (Table 4).

The longitudinal results for both groups of students with BD/LD were generally similar and stable, except for written language. The comorbid students who received no LD services did significantly worse over time in this achievement area than the BD/LD students who received LD services. Indeed, the effect size was moderate (-0.62 ; Cohen, 1988) and very noteworthy compared with all other effect sizes for achievement in both groups (range = -0.03 to 0.16).

TABLE 3
Three-Year Changes in Psychoeducational Characteristics of BD Students With and Without LD at Baseline

Instrument	Mean Scores			
	No LD at Baseline (<i>n</i> = 29)		LD at Baseline (<i>n</i> = 52)	
	Baseline	3-Year	Baseline	3-Year
WISC-R IQ				
Verbal	97.8 \pm 14.5	99.4 \pm 14.5	99.1 \pm 15.6	94.8 \pm 16.2 ^a
Performance	99.0 \pm 11.2	101.7 \pm 12.6	103.3 \pm 15.6	100.9 \pm 17.8
Full Scale	98.1 \pm 11.0	100.6 \pm 13.0	101.2 \pm 16.0	97.5 \pm 17.0 ^b
WJPEB standard scores				
Reading (28,44)	97.8 \pm 11.7 ^c	97.5 \pm 14.3 ^d	88.1 \pm 12.3	88.5 \pm 12.1
Mathematics (25,43)	97.6 \pm 12.1 ^c	94.2 \pm 13.7 ^e	86.1 \pm 12.3	87.0 \pm 12.8
Written language (18,33)	100.6 \pm 13.0 ^f	96.8 \pm 14.3 ^e	90.3 \pm 12.2	86.9 \pm 11.9

Note: () denotes *n*'s at both time points for No LD at Baseline and LD at Baseline, respectively. BD = behavioral disorder; LD = learning disorder; WJPEB = Woodcock-Johnson Psychoeducational Battery Test of Achievement.

Baseline LD > 3-Year LD: ^a $F = 6.60, p = .01$, ^b $F = 9.05, p < .005$.

Baseline No LD > Baseline LD: ^c $p < .005$, ^f $p < .05$.

3-Year No LD > 3-Year LD: ^d $p < .01$, ^e $p < .05$.

TABLE 4
Three-Year Changes in Psychoeducational Characteristics of Baseline BD/LD Students
With and Without State LD Services

Instrument	Mean Scores			
	Received State LD Services (<i>n</i> = 18)		No State LD Services (<i>n</i> = 34)	
	Baseline	3-Year	Baseline	3-Year
WISC-R IQ				
Verbal	95.9 ± 13.9	93.4 ± 15.3	100.7 ± 16.3	95.4 ± 16.8
Performance	99.6 ± 11.5	100.4 ± 15.4	105.3 ± 17.2	101.2 ± 19.1
Full Scale	97.2 ± 12.3	96.2 ± 14.4	103.2 ± 17.4	98.2 ± 18.5
WJPEB standard scores				
Reading (14,30)	86.8 ± 12.4	88.8 ± 12.7	88.7 ± 12.4	88.3 ± 12.0
Mathematics (14,29)	84.2 ± 12.5	84.5 ± 12.1	87.1 ± 12.3	88.3 ± 13.1
Written language (13,20)	85.1 ± 12.6	86.8 ± 12.8	93.7 ± 11.0 ^a	86.9 ± 11.7 ^b

Note: () denotes *n*'s at both time points for State LD Services and No State LD Services, respectively. BD = behavioral disorder; LD = learning disorder; WJPEB = Woodcock-Johnson Psychoeducational Battery Test of Achievement.

^a Baseline No State LD Services > Baseline State LD Services: $p < .05$.

^b Baseline No State LD Services > 3-year No State LD Services: $F = 6.81, p = .01$.

DISCUSSION

In this study of the stability of LDs in children with psychiatric disorders, LDs in special education students with BD remained common and generally stable over their first 3 years of BD services, regardless of definition or achievement area. The original two groups of BD students with and without LD remained true to their initial groupings; i.e., the BD students without LD continued to obtain average academic achievement, while the BD/LD students continued to show noteworthy academic lags. The comorbid BD/LD students who did not receive LD services did significantly worse in written language, which was also the most deficient domain for all subjects.

Literature Comparison

The stability of LDs in children with psychiatric disorders has not been previously investigated with ORs. Indeed, most earlier longitudinal studies of children with LD have investigated samples from general populations or LD clinics and have rarely used ORs. Rather, they have demonstrated the persistence of LD by showing the continuance of low achievement scores at follow-up (Satz et al., 1998) or by finding varying percentages (12%–66%) of children with an LD definition at baseline who persisted with that same definition at follow-up (Badian, 1999; Share and Silva, 1986; Shaywitz et al., 1992b; Silver et al., 1999; Wright et al., 1996). ORs were determined

for the stability of reading and spelling LD (i.e., low achievement) in a general population sample ($n = 37$) of Australian children from ages 7 and 8 years to ages 11 and 12 years (Prior et al., 1999). The longitudinal ORs were 15.4 for reading disorder and 17.4 for spelling disorder, which are much lower than the reading and written language ORs found in our study with its more stringent definition of low achievement. ORs have rarely been determined in longitudinal studies of childhood psychiatric disorders. The stability of *DSM-III-R* psychiatric disorders (ADHD, conduct disorder, overanxious disorder, and oppositional defiant disorder) in a large general population sample of children aged 9 to 18 years was investigated at baseline and 2½ years later (Cohen et al., 1993). The OR range was primarily from 3.1 to 17.2. Thus the ORs (2.8–78.7) found in the current study for LD definitions were similar or at times stronger than previous results for *DSM-III-R* definitions of psychiatric disorders.

Previous studies have shown variable results for the longitudinal changes of IQ and achievement standard scores in LD students (Anderson et al., 2001; Shaywitz et al., 1992a; Wigle et al., 1988). Recently, Bielinski and Ysseldyke (2002) used a large longitudinal database to track the reading and mathematics performance for unchanging cohorts of both regular education students and special education students. All students were tested yearly from fourth to eighth grade. Although the special education students consistently scored significantly lower

than the regular education students in both reading and mathematics, the differences between the two groups of students remained constant in both achievement areas over time. Also, the learning curves were generally parallel between the two types of students in both achievement domains. Similar results were found for the groups with and without LD in this study.

Although the BD students who were also LD at baseline did not show any significant changes in their mean achievement scores over the 3 years, their Verbal and Full Scale mean IQ scores were significantly lower. This finding may represent the "Matthew effect" that has been discussed by Stanovich (1986), i.e., an extension of the concept that over time rich persons get richer while poor persons get poorer. More specifically, if children are not progressively exposed to and do not sufficiently master the advanced vocabulary/concepts of more complicated reading material, their cognitive abilities integral to Verbal IQ are affected negatively over time and their Verbal IQ may appear to decrease. In this case, at both time points the BD/LD students read at significantly lower levels than the BD students with no LD.

Finally, although LDs in written language/expression have rarely been investigated in children with psychiatric disorders, LD in written language was an important finding in this study. The consequence of neglecting this domain was shown in a recent study of LD in outpatient clinic children with ADHD (Mayes et al., 2000). Their prevalence rate was approximately 30% each for LD in reading, mathematics, and spelling, compared with 65% for written expression. The overall LD rate almost doubled with the inclusion of LD in written expression, leading the authors to emphasize that exclusion of this achievement domain may cause underestimation of LD rates in children with ADHD and the neglect of an LD that is amenable to several intervention strategies.

Limitations

The generalization of our findings to other groups of children with psychiatric disorders should be done cautiously because of the unique nature of our sample. The psychopathology of these children was serious enough in school to require special education services. Thus the severity and/or chronicity of their psychiatric diagnoses could have contributed to the persistence of any LD, as could the lack of community mental health treatment (both therapy and medication) commonly experienced by students with serious emotional disturbance (Mattison,

1999). Such factors would be important to include in future related research, in addition to the investigation of specific psychiatric diagnoses which could not be accomplished in this study.

While this study presented a unique opportunity for naturalistic longitudinal research that may not be repeated, inherent limitations occurred. For example, administration of the same IQ and achievement battery to all students who received BD services at both time points would have been desirable. Unfortunately, school examiners have varying instrument preferences, and testing in all three achievement areas is no longer always required by states (as was reflected by the inconsistent administration of achievement testing in all three domains to the subjects in this study). Also, individual education plans were not investigated to obtain specific descriptions of the special education interventions (both BD and LD) that the subjects received. The adequacy of such services is one factor that could influence LD outcome. In addition, our understanding of the academic changes for the subjects over the 3 years would have been enhanced through comparison with achievement testing that they received in their educational careers prior to BD classification, or through use of a comparison group such as newly classified LD students. Finally, it should be recalled that the participants in this study, although representative, accounted for 34.8% of the original 233 students and that this sample was suburban, which might limit the generalization of the findings to other populations of students (e.g., urban/inner city).

The objective LD definitions used in this study were originally selected to conservatively represent the predominant choices at that time. The debate between low achievement and discrepancy definitions continues (Buka et al., 1998; Kavale and Forness, 2000; Siegel, 1999; Vellutino et al., 2000). Researchers in reading disorders have also recently been suggesting newer definitions, such as the use of a percentile cutoff (e.g., 15th%) without consideration of IQ, or measurement of underlying neuropsychological processes such as phonological processing (Fletcher et al., 1998; Grigorenko, 2001). The generally good stability results found across definitions and achievement areas in this 3-year study do not recommend the superiority of any specific definition. However, the specificity of the LD definitions in this study, the use of the same IQ-achievement battery at both times, and the analyses with ORs should prove useful to future LD research in children both with and without psychiatric disorders.

Clinical Implications

This study demonstrates that clinicians, in both their roles as school consultants and advocates for patients, must be insistent that schools do a comprehensive evaluation for LD in students with EBD. Regardless of definition and achievement area, LDs proved not only common in students just beginning their BD programming (indeed, LD rates might have been even higher if all subjects had been tested in all three achievement domains), but LDs were also very persistent during the first 3 years of services. This need for increased clinician attention cannot be overemphasized because consistent with Individuals With Disabilities Education Act (IDEA) legislation, states are no longer required to conduct full assessment batteries for the reevaluation of students with EBD. Or, as in this study, state LD definitions may exclude students with EBD who fulfill other legitimate LD definitions. Students with EBD who have unrecognized or untreated LD are undoubtedly at higher risk for worsening academic performance over time, which subsequently negatively affects graduation and employment.

Clinicians must also be careful that neither they nor special educators assume that the academic performance of students with EBD is primarily related to their psychopathology and thereby overlook the role of comorbid LDs. In this study academic dysfunction was related to comorbid LDs, as the BD students without LD continued to achieve in the average range. Indeed, recent work has begun to demonstrate that teachers do *not* rate BD/LD students as significantly different from BD students without LD on behavioral checklists (Handwerk and Marshall, 1998; Mattison, 2001). Thus school evaluation teams must be careful not to presuppose that the degree of psychopathology correlates with the severity of academic dysfunction and thereby rule out consideration of accompanying LD. This issue is rightfully a matter of growing concern among special education researchers (Handwerk and Marshall, 1998; Rock et al., 1997) who worry that ambiguous federal definition guidelines lead to the underdiagnosis and subsequent undertreatment of BD students with accompanying LD, which further worsens their educational prognosis.

The 3-year findings in this study of relatively stable standard scores across the reading and mathematics achievement domains may be interpreted in at least two ways. First, sufficient BD programming appears to have frequently stabilized the academic performance of students

newly classified as BD (both with and without LD) in these achievement areas despite their ongoing psychopathology. This sustained performance may well have been related to the individual or small-group attention and coaching of learning skills that are typically provided in BD classrooms, general features that are proving important to helping children with LD (Swanson et al., 1999; Vaughn et al., 2000). Clinicians should be mindful of this potential capability of good BD programming, especially for their patients who are showing academic deterioration without special education.

A second interpretation may have expected more improvement in standard scores and then might explain the lack of progress as secondary to inadequate delivery of LD services to comorbid BD/LD students. Indeed, special education researchers have recently questioned the sufficiency of the academic instruction practices by teachers of BD students (Coleman and Vaughn, 2000; Gunter and Denny, 1998), let alone their ability to provide more complex LD services. Furthermore, BD teachers may too often overemphasize behavioral management as the prerequisite step before improvement in academic function can occur, when, in fact, proper academic instruction can selectively lead to a decrease in behavior problems (Penno et al., 2000).

Finally, our findings for written language have two implications. First, this area of achievement testing must not be neglected. Although LD in written language occurred in more than 30% of the BD subjects, its rate may have been higher because this achievement area was the least tested. Such undertesting is probably common nationally. Second, in addition to the finding of significantly decreased performance in written language for the BD/LD students who received no LD services, this achievement area was also weak for BD students without LD. This finding raises the question of how well students with BD are instructed in written language/expression, especially as they enter middle school when such skills become even more important. For example, how frequently are students with BD given writing assignments, followed by correction and subsequent instruction to improve their skills? Clinicians must ensure that educators pay attention to the writing curriculum received by students with BD.

REFERENCES

American Academy of Child and Adolescent Psychiatry (1998), Practice parameters for the assessment and treatment of children and adolescents with

- language and learning disorders. *J Am Acad Child Adolesc Psychiatry* 37(suppl):46S–62S
- Anderson JA, Kutash K, Duchnowski AJ (2001), A comparison of the academic progress of students with EBD and students with LD. *J Emotional Behav Disord* 9:106–115
- Badian NA (1999), Reading disability defined as a discrepancy between listening and reading comprehension: a longitudinal study of stability, gender differences, and prevalence. *J Learn Disabil* 32:138–148
- Bielinski J, Ysseldyke JE (2002), Effect of different methods of reporting and reclassification on trends in test scores for students with disabilities. *Except Child* 68:189–200
- Buka SL, Satz P, Seidman L, Lipsitt L (1998), Defining learning disabilities: the role of longitudinal studies. *Thalamus* 16:14–29
- Cohen J (1988), *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ: Erlbaum
- Cohen P, Cohen J, Brook J (1993), An epidemiological study of disorders in late childhood and adolescence: persistence of disorders. *J Child Psychol Psychiatry* 34:869–877
- Coleman M, Vaughn S (2000), Reading interventions for students with emotional/behavioral disorders. *Behav Disord* 25:93–104
- Fletcher JM, Francis DJ, Shaywitz SE et al. (1998), Intelligent testing and the discrepancy model for children with learning disabilities. *Learn Disabil Res Pract* 13:186–203
- Glassberg LA, Hooper SR, Mattison RE (1999), Prevalence of learning disabilities at enrollment in special education students with behavioral disorders. *Behav Disord* 25:9–21
- Grigorenko EL (2001), Developmental dyslexia: an update on genes, brains, and environments. *J Child Psychol Psychiatry* 42:91–125
- Gunter PL, Denny RK (1998), Trends and issues in research regarding academic instruction of students with emotional and behavioral disorders. *Behav Disord* 24:44–50
- Handwerk ML, Marshall RM (1998), Behavioral and emotional problems of students with learning disabilities, serious emotional disturbance, or both conditions. *J Learn Disabil* 31:327–338
- Kavale KA, Forness SR (1995), *The Nature of Learning Disabilities: Critical Elements of Diagnosis and Classification*. Mahwah, NJ: Erlbaum
- Kavale KA, Forness SR (2000), What definitions of learning disabilities say and don't say: a critical analysis. *J Learn Disabil* 33:239–256
- Mattison RE (1999), Use of psychotropic medications in special education students with serious emotional disturbance. *J Child Adolesc Psychopharmacol* 9:149–155
- Mattison RE (2001), Learning disabilities in special education students referred for psychiatric consultation. *Thalamus* 19:11–19
- Mattison RE, Forness SR (1995), Mental health system involvement in SED placement decisions. In: *Issues in Educational Placement: Students With Emotional and Behavioral Disorders*, Kauffman JM, Lloyd JW, Hallahan DP, Astuto TA, eds. Hillsdale, NJ: Erlbaum, pp 197–211
- Mayes SD, Calhoun SL, Crowell EW (2000), Learning disabilities and ADHD: overlapping spectrum disorders. *J Learn Disabil* 33:417–424
- Penno DA, Frank AR, Wacker DP (2000), Instructional accommodations for adolescent students with severe emotional or behavioral disorders. *Behav Disord* 25:325–343
- Prior M, Smart D, Sanson A, Oberklaid F (1999), Relationships between learning difficulties and psychological problems in preadolescent children from a longitudinal sample. *J Am Acad Child Adolesc Psychiatry* 38:429–436
- Reynolds CR, Stanton HC (1990), *Discrepancy Determinator-Revised: Technical and Interpretive Manual*. Bensalem, PA: Trainware
- Rock EE, Fessler MA, Church RP (1997), The concomitance of learning disabilities and emotional/behavioral disorders: a conceptual model. *J Learn Disabil* 30:245–263
- Satz P, Buka S, Lipsitt L, Seidman L (1998), The long-term prognosis of learning disabled children: a review of studies 1954–1993. In: *Specific Reading Disability: A View of the Spectrum*, Shapiro BK, Accardo PJ, Capute AJ, eds. Timonium, MD: York Press, pp 225–250
- Share DL, Silva PA (1986), The stability and classification of specific reading retardation: a longitudinal study from age 7 to 11. *Br J Educ Psychol* 56:32–39
- Shaywitz BA, Fletcher JM, Holahan JM, Shaywitz SE (1992a), Discrepancy compared to low achievement definitions of reading disability: results from the Connecticut Longitudinal Study. *J Learn Disabil* 25:639–648
- Shaywitz SE, Escobar MD, Shaywitz BA, Fletcher JM, Makuch R (1992b), Evidence that dyslexia may represent the lower tail of a normal distribution of reading disability. *N Engl J Med* 326:145–150
- Siegel LS (1999), Issues in the definition and diagnosis of learning disabilities: a perspective on Guckenberger v. Boston University. *J Learn Disabil* 32:304–319
- Silver CH, Pennett HDL, Black JL, Fair GW, Balise RR (1999), Stability of arithmetic disability subtypes. *J Learn Disabil* 32:108–119
- Stanovich KE (1986), Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read Res Q* 21:360–407
- Swanson HL, Hoskyn M, Lee CM (1999), *Interventions for Students With Learning Disabilities: A Meta-Analysis of Treatment Outcomes*. New York: Guilford
- Vaughn S, Gersten, R, Chard DJ (2000), The underlying message in LD intervention research: findings from research syntheses. *Except Child* 67:99–114
- Vellutino FR, Scanlon DM, Lyon GR (2000), Differentiating between difficult-to-remediate and readily remediated poor readers: more evidence against the IQ-achievement discrepancy definition of reading disability. *J Learn Disabil* 33:223–238
- Wagner MM (1995), Outcomes for youths with serious emotional disturbance in secondary school and early adulthood. *Future Child* 5:90–112
- Wechsler D (1974), *The Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation
- Wigle SE, White WJ, Parish TS (1988), A longitudinal comparison of high IQ and low IQ LD students. *Read Improvement* 25:282–285
- Woodcock RW, Johnson MB (1978), *Woodcock-Johnson Psychoeducational Battery*. Hingham, MA: Teaching Resources
- Wright SF, Fields H, Newman SP (1996), Dyslexia: stability of definition over a five year period. *J Res Read* 19:46–60