Exceptional Children

Vol. 80, No. 2, pp. 169-186.

Unidentified Language Deficits in Children With Emotional and Behavioral Disorders: A Meta-Analysis

ALEXANDRA HOLLO

JOSEPH H. WEHBY
Peabody College of Vanderbilt University

REGINA M. OLIVER
University of Nebraska-Lincoln

ABSTRACT: *Low language proficiency and problem behavior often co-occur, yet language deficits are likely to be overlooked in children with emotional and behavioral disorders (EBD). Random effects meta-analyses were conducted to determine prevalence and severity of the problem. Across 22 studies, participants included 1,171 children ages 5-13 with formally identified EBD and no history of developmental, neurological, or language disorders. Results indicated prevalence of below-average language performance was 81%, 95% CI [76, 84]. The mean comprehensive language score was 76.33 [71, 82], which was significantly below average. Implications include the need to (a) require language screening for all students with EBD, (b) clarify the relationship between language and behavior, and (c) develop interventions to ameliorate the effects of these dual deficits.*

Children with emotional and behavioral disorders (EBD) exhibit maladaptive social and behavioral responses characterized as severe, chronic, and pervasive (Gresham, 2005). Whether identified and served through educational or mental health channels, outcomes for children with EBD are likely to include school failure, dropping out, unemployment, substance abuse, and contact with mental health or criminal justice systems (Bradley, Doolittle, & Bartolotta, 2008). Difficulty in academic, social, emotional, and behavioral functioning contribute uniquely to the negative outcomes experienced by children with EBD; however, problems in each of these areas interact in ways that are not yet well understood (Tomblin, Zhang, Buckwalter, & Catts, 2000). One variable that is strongly related to performance in each of these areas is children's language proficiency (Beitchman, Cohen, Konstantareas, & Tannock, 1996). Language development is the

foundation of, and inexorably intertwined with, adaptive academic, social, and behavioral performance (Im-Bolter & Cohen, 2007; Toppelberg & Shapiro, 2000).

An extensive body of literature has described interrelations among language, learning, and behavioral problems in school-age children. Although causal or directional mechanisms of these relations have yet to be established, descriptive evidence supports a strong association between linguistic and behavioral competence (Hooper, Roberts, Zeisel, & Poe, 2003; Zadeh, Im-Bolter, & Cohen, 2007). That is, children who exhibit problem behavior tend to have low language proficiency, and children with low language proficiency tend to exhibit problem behavior (Benner, Nelson, & Epstein, 2002). As many scholars have noted, although children with a range of maladaptive behavioral profiles are at risk for communication disorders, low language proficiency is often overlooked in children whose challenging behavior is highly salient to adults (e.g., Cohen, Davine, Horodezky, Lipsett & Isaacson, 1993; Donahue, Cole, & Hartas, 1994).

Many researchers have documented the association between language and behavioral competence, but only one systematic review to date has focused on children formally identified as EBD. In 2002, Benner et al. conducted a comprehensive search and narrative summary of the literature. They concluded that 71% of students with EBD had concurrent language impairments. Moreover, 64% were deficient in expressive language, and 56% in receptive skills. The current investigation also synthesized studies reporting prevalence estimates of language impairment (LI) in children with formally recognized EBD, but adds to the literature in several ways. Most importantly, meta-analytic methods incorporated both systematic review procedures and quantitative synthesis of data reported in primary research studies. Meta-analysis was used to compute not only prevalence rates, but also a second dependent variable: mean standard scores on comprehensive language assessments. The methodology also permitted analysis of moderator variables for both prevalence and means. Additionally, it has been demonstrated that for many children with psychiatric disorders, concurrent language deficits often are undetected (e.g., Cohen et al., 1993). To determine the extent of this problem, the population of interest was children with EBD and unidentified language deficits. Finally, stringent inclusion criteria were employed to minimize alternative explanations for high prevalence rates.

For children with EBD, undetected LI can have serious consequences. Researchers have noted that children's language deficits often are misperceived as low intelligence; inattention; noncompliance; or deliberate dishonesty, disrespect, and defiance (Cohen et al., 1993; Donahue et al., 1994). Such characterizations may add stress, frustration, and blame to interactions that already are likely to be challenging, and may contribute to negative or coercive interactions (Sutherland & Morgan, 2003). Finally, problem behavior may be exacerbated if adults' verbal input is too complex for students to comprehend (Harrison, Gunter, Reed, & Lee, 1996). Instruction and interventions requiring intact language skills therefore may be counterproductive.

Identifying prevalence and severity of language deficits, as well as areas of linguistic strengths and weaknesses, may provide a foundation on which to build supports for children with EBD. This information is vitally important to ensure that children receive appropriate assessments, resources, and treatment. The purpose of the current study was to provide a quantitative synthesis of research examining unidentified language impairment in school-age children with EBD. An overview of issues surrounding identification of language and behavioral disorders is provided, followed by discussion of sources of heterogeneity among primary studies highlighted in prior research. Methods and procedures employed in the meta-analyses designed to either minimize or analyze differences among studies also are presented.

DETERMINING CO-OCCURRENCE OF LANGUAGE AND BEHAVIOR DISORDERS

Children may be identified as having emotional and behavioral disorders through either educational or mental health channels; however, these two pathways are independent and not mutually exclusive. According to criteria outlined in the *Diagnostic and Statistical Manual of Mental Disor-*

ders (DSM-IV; American Psychiatric Association, 2000), children may receive diagnoses for affective (e.g., mood or anxiety), disruptive (e.g., attention deficit/hyperactivity), or behavioral disorders (e.g., oppositional defiant or conduct disorders; ODD or CD respectively). Regardless of the presence or absence of a psychiatric diagnosis (Della Toffalo & Pedersen, 2005), children may receive educational services under the disability label emotional disturbance (ED), as defined by the Individuals With Disabilities Education Improvement Act (IDEA, 2006). Because the current definition of ED remains controversial and may underrepresent the population of interest (Gresham, 2005), the term EBD is used here to include all children with either ED labels or DSM diagnoses, regardless of the source of identification or setting of services received.

Identifying prevalence and severity of language deficits, as well as areas of linguistic strengths and weaknesses, may provide a foundation on which to build supports for children with EBD.

Defining and diagnosing language disorders also can be controversial, as can the terms used to describe children with language disorders, deficits, delays, or impairments. For example, specific language impairment (SLI) has been defined as expressive (production) or receptive (comprehension) delays in the absence of explanatory factors such as neurological, sensory, motor, or environmental deficits (Tomblin et al., 1997). Inclusion criteria for specifying a delay, however, is less straightforward. Some researchers and practitioners have used discrepancy criteria (e.g., between verbal and nonverbal intelligence or achievement and chronological age expectations), and others have used various cutoff scores on diagnostic measures (Law, Boyle, Harris, Harkness, & Nye, 2000; Tomblin et al., 1997). Note that disorders affecting the mechanics of speech (e.g., articulation, fluency) are excluded from these definitions.

For the present discussion, language *deficit* or *impairment* is abbreviated LI to avoid confusion with terminology from other disciplines (e.g., LD, or learning disability). LI is used here to denote low language proficiency identified through various assessment methods and defined by different diagnostic criteria. The term may or may not include children with a diagnosed *disorder* such as SLI. The purpose of the current study was not to diagnose children, but to synthesize descriptive data as reported in primary studies. The data most commonly reported related to expressive or receptive deficits occurring in one or more areas of language, including semantics (meaning) and syntax (grammar). Deficits in each of these areas commonly overlap, and all have been linked with children's problem behaviors (Beitchman et al., 1996; Harrison et al., 1996).

Prior Research

In the primary research studies summarized by Benner et al. (2002), prevalence estimates ranged from 35% to 97%. Although unable to assess potential moderator variables empirically, Benner and colleagues suggested that between-studies heterogeneity may be due to differences in measurement variables, diagnostic standards, and participant characteristics. Limitations and recommendations noted by those and other authors are outlined, as are the strategies used to address those concerns in the current study.

Measurement Variables. There are many ways to determine the presence, type, and severity of LI, and method of assessment is known to affect those outcomes (Friberg, 2010; Law et al., 2000). For example, the number and type of assessments used to determine case status may vary according to the purpose for the assessment (Law et al., 2000). In research studies, a single diagnostic instrument often is used to determine case status or presence/absence of LI. In practice, however, speech-language pathologists (SLPs) are encouraged to use multiple assessments that include standardized and naturalistic measures (e.g., spontaneous language samples) and use of clinical judgment when determining eligibility for services and developing treatment plans (Friberg, 2010). Potential moderators of prevalence outcomes are, therefore, whether estimates were obtained (a) via a single measure or an assessment battery, and (b) for the purpose of informing research (e.g., describing language performance) or

practice (e.g., providing services). Benner et al. (2002) also noted that test instruments may produce different estimates as a function of technical adequacy. Although it was not possible to analyze all differences (e.g., compare versions of the same test), moderator analyses included whether different measures produced significantly different results.

Diagnostic Standards. Determining prevalence of LI is complicated further by use of varying cutoff scores to determine case status (Law et al., 2000; Tomblin et al., 1997). Benner et al. (2002) noted that prevalence varied across studies according to cutoff criteria established by individual researchers: Studies with more stringent criteria reported lower rates of LI and vice versa. In studies using a single test to determine LI in children with EBD, cut scores typically were 1, 1.5, or 2 standard deviations below standardized norms. These roughly corresponded to mild, moderate, and severe deficits; however, a child with a score of 77 could be considered LI in one study and not another. In the current study, differences in diagnostic standards were minimized by assigning a priori definitions and cutoffs. A 1–2 standard deviation cutoff on a single measure was considered less stringent, and scores reported in this range were assigned to the mild category. Scores 2 or more standard deviations below the mean were assigned to the moderate/severe category.

Reporting standards also varied across studies. Some authors reported test results as sample means or as a prevalence estimate defined as proportion of the sample with scores 1, 1.5, or 2 standard deviations below the mean. Others reported only the number of children achieving case status, or prevalence in the sample, without reference to mean scores. Additionally, some studies reported case status for the entire sample, including proportions of students with mild, moderate, and severe LI as well as those without deficits. Others reported only presence/absence of language impairment, without indication of severity. For studies with stringent cutoff criteria, results for all children often were not reported (thus omitting children with mild deficits and underestimating the numerator in overall prevalence rate). For studies that did not specify severity, findings were included only in the moderate to severe category. Although it is possible some children may have received standard scores above 70 (2 *SD*), the authors' use of stringent criteria to assign LI diagnoses supported this decision. In the current investigation, the below-average category included only those studies reporting the full range of language proficiency.

Participant Characteristics. Benner et al. (2002) noted that studies conducted in schools reported higher prevalence than studies conducted in "more restrictive clinical settings" (p. 51), defined as treatment centers, speech, or psychiatric clinics. Level of restriction (day or residential program) and setting (school or clinic) were reported in several studies, and were assessed as moderator variables. Researchers (Cohen et al., 1993; Nelson, Benner, & Cheney, 2005) also have questioned whether form or severity of problem behavior is related to linguistic proficiency. That is, children may exhibit topographies of behavior characterized as internalizing (e.g., withdrawn, depressed, anxious), externalizing (e.g., destructive, defiant, aggressive), or combined, and behavioral subtypes may be related to differences in language profiles. Although assessing these associations may have important implications for supporting children with EBD, insufficient data were available to conduct these analyses.

Methodology

Meta-analysis is a quantitative method specifically designed for synthesizing research studies (Borenstein, Hedges, Higgins, & Rothstein, 2009; Lipsey & Wilson, 2001). An advantage of using meta-analysis is that data used to calculate outcomes are weighted to account for precision of primary data. Larger sample sizes provide more precise estimates and are assigned higher weights, which is considered more accurate than simply averaging across studies (Lipsey & Wilson, 2001). It is also possible to analyze moderator variables to explain between-group heterogeneity. Given the data available in reports included in the current study, however, it was not possible to follow Benner et al.'s (2002) recommendation to analyze moderator variables related to underlying cognitive processes of language development such as attention or intelligence. Instead, an attempt was

made to control for those factors by employing very conservative inclusion criteria.

Summary and Research Questions

Although presence of unidentified language deficits in school-age children with EBD is a well-documented phenomenon, and although unidentified language deficits may contribute to poor proximal and distal outcomes, questions remain regarding the extent, severity, and types of deficits these children are likely to experience. This information is critically important to developing interventions that have the potential to improve children's functioning across multiple areas. The current investigation builds upon an earlier review (Benner et al., 2002) by including additional studies (conducted after 2001; published and unpublished) and specifying predetermined measurement and diagnostic criteria. This study extends the literature by including analyses of mean standardized test scores in addition to prevalence estimates, using conservative inclusion criteria to minimize differences in participant characteristics, and employing meta-analytic methodology to answer the research questions. This study was designed to answer research questions related to two primary outcomes for students with EBD:

1. What is the prevalence and severity of unidentified deficits in comprehensive, receptive, and expressive language proficiency? Is prevalence moderated by differences in behavioral topography (internalizing or externalizing), program type (day or residential), setting (school or clinic), number of measures (single or multiple), or purpose of assessment (research or practice)?
2. Is mean performance on standardized measures of comprehensive language proficiency significantly below average in specific language components (comprehensive, expressive, receptive, semantic, and syntactic)? Is mean performance moderated by differences in measures (specific comprehensive tests) or participant characteristics (severity or topography of problem behavior)?

METHOD

Eligibility Criteria

Studies were selected according to eligibility criteria for participant characteristics and for outcome variables as outlined in this section. To avoid possible publication bias (Lipsey & Wilson, 2001), no restrictions were in place regarding where studies were conducted, language, date, or type of publication (e.g., journal, dissertation, technical report). All study designs were eligible for inclusion (e.g., descriptive, group experimental or quasiexperimental, single subject), providing preintervention descriptive data were reported.

Participant Characteristics. As in the earlier review (Benner et al., 2002), samples were children with formal designations of EBD. To increase the probability that samples were drawn from a relatively homogeneous population, the following criteria also were defined for participant samples. First, all participants were in Grades K-8 (ages 5-13). Preschool-age children were excluded because they are typically not identified as EBD until kindergarten or later. Additionally, LI is assessed differently in preschool-age children, and may result in inflated prevalence rates relative to those above kindergarten age (Law et al., 2000). High school students were excluded because it was unclear how delinquency, substance use, school dropout, and retention may affect EBD samples at that age (e.g., if students with LI were more likely to be arrested or drop out of school, prevalence estimates would be artificially low). Second, participants with EBD had to be free of conditions known to co-occur with LI. That is, studies were excluded if study reports indicated participants had below-average IQs (defined in most of the primary studies as full scale IQ $\leq$ 80 or 85; a cutoff of 80 was adopted here), or any neurological condition related to problems with language or learning, including any intellectual or developmental disability, schizophrenia, autism spectrum, or attention deficit/hyperactivity disorders. Finally, studies were excluded if participants had pre-existing LI; were receiving speech, hearing, or language-related services; or were sampled because of suspected LI.

Outcome Variables. Studies contributed data to prevalence estimates if they reported either (a) number or proportion of the sample with clinical

diagnoses of language delay, impairment, or deficit; or (b) number of children scoring below 1 standard deviation, within 1-2 standard deviations, or below 2 standard deviations on a comprehensive standardized language measure (i.e., composed of expressive and receptive language tasks or composite scales). Studies contributed data to mean outcomes if they reported sample means and standard deviations on the same types of standardized language measures. Because standard errors are not comparable across criterion-referenced, informal, or naturalistic measures, outcomes include only norm-referenced measures.

Literature Search, Screening, and Study Selection

To locate studies reporting language outcomes in children with EBD, keyword searches were conducted in electronic databases (i.e., ERIC, PsycINFO, PsycARTICLES, Linguistics & Language Behavior Abstracts, PubMed, ProQuest Dissertation Abstracts, OpenSigle, British Education Index, ISI Web of Knowledge, and Google Scholar). Searches were conducted in February and August 2011, and included permutations of the terms emotional disturbance, EBD, behavior*disorder*, language, communicat*, psych*, impair*, disorder, child*, and school. Forward and backward searches of relevant reviews and frequently cited studies also were conducted. Titles and abstracts of identified articles then were screened for broad inclusion criteria (i.e., presence of EBD and any empirical data). Next, potentially relevant studies were retrieved and full articles were reviewed by two master's-level research assistants. Retained articles were read by the first and third authors to determine whether data were reported according to inclusion criteria. Finally, included studies were coded by the first and third authors.

Analytic Strategies

Separate random effects meta-analyses were planned to describe comprehensive, receptive, and expressive language ability for two outcomes: prevalence and means. Samples within meta-analyses are comprised of research reports rather than individual participants, and each report provides one or more research findings (Lipsey & Wilson, 2001). In the current meta-analyses, findings were prevalence estimates and mean scores on standardized measures. These descriptive outcomes each represent a single variable, unlike more commonly reported effect sizes for between-group or bivariate relationships such as Cohen's *d* or Fisher's *z*. Means were point estimates of central tendency, and prevalence was the proportion of children in each sample with LI. For both outcomes, τ represented the standard deviation and I^2 represented the proportion of true heterogeneity to random error, or the signal-to-noise ratio. Additional analyses were conducted to assess publication bias, or "the file drawer problem" (Borenstein et al., 2009, p. 379), in which only significant results from large studies are published.

To avoid violating the assumption of independence, the following decision rules were implemented to ensure each participant sample within a primary research report was included only one time in each meta-analysis. First, several studies produced more than one report (e.g., a longitudinal study, or dissertation later published in a journal). If the same sample data were reported in multiple articles, the earlier report was cited. If the earlier report was a dissertation study, the published version was cited (e.g., Griffith, Rogers-Adkinson, & Cusick, 1997; Rogers-Adkinson, 1995). If authors reported different data in separate reports, the article providing data for the largest number of participants per outcome was selected to improve precision of the overall estimates (Lipsey & Wilson, 2001). Second, the category boundaries for prevalence of mild and moderate/severe LI were exhaustive and exclusive. In two studies, however, cutoff scores were reported as 1.5 standard deviations. In these cases, half of the sample was counted in the mild category, and half in moderate/severe.

Prevalence. To answer the first research question regarding prevalence of LI in children with EBD, individual meta-analyses were planned for comprehensive, expressive, and receptive language at two levels of severity: mild and moderate/severe. For studies reporting results for the total sample at all levels of severity, the two levels were combined to represent all children in the sample with below-average performance. Rather than using authors' varying definitions of LI, cat-

egories were determined a priori. Children with scores 1-2 standard deviations below the mean on a comprehensive language measure (71–85 where M = 100 and SD = 15) were classified as having mild deficits. The moderate/severe category included children with scores two or more standard deviations below the mean ($\leq$ 70). Children with clinical diagnoses of language disorder also were included in this category if severity was not otherwise specified. The category of below average ability included all children with LI diagnoses or scores below 85.

Prevalence was the proportion of participants in a sample of children with EBD identified as having low language ability. To calculate effect sizes for proportions, Lipsey and Wilson (2001) recommend transforming proportions and their standard errors to a logarithmic scale using these formulas:

$$ESp = log_e\left[\frac{p}{1-p}\right] \qquad SEp = \sqrt{\frac{1}{np} + \frac{1}{n(1-p)}}$$

Whereas values for proportions are constrained between zero and one, logits have an infinite range and, therefore, provide a more accurate estimate of the distribution of proportions around the mean (between-study variance). Logits were then converted back to a scale of 0 to 1 to facilitate interpretation of prevalence rates and confidence intervals. Statistical significance was determined by examining 95% confidence intervals for overlapping data points.

Mean Scores. To answer the second research question—whether mean standard scores of children with EBD are below average—means and standard deviations were coded for studies reporting comprehensive, expressive, receptive, semantic, or syntactic language scores from standardized assessments. Standard errors were computed by dividing the standard deviation by the square root of N. To determine whether effect sizes were statistically different from 85 (the cutoff for average language ability on standardized measures), z was calculated using the formula z = (mean ES – 85)/(*SE* of mean ES).

Moderator Analyses. Analyses of moderators were planned for meta-analyses with sufficient true heterogeneity as defined by $I^2 > 75\%$ (Lipsey & Wilson, 2001). For prevalence, variables included topography (internalizing or externalizing) and severity of problem behavior (defined as more or less restrictive programs; i.e., day or residential), setting (school or clinic), purpose for data collection (research or practice; e.g., data were collected to determine eligibility for services, often reported via chart review), and assessment (single or multiple measures). The variable assessed for mean scores was the measure used to obtain outcome data. The only tests meeting inclusion criteria were various editions of the Clinical Evaluation of Language Fundamentals (CELF-R, Semel, Wiig, & Secord, 1987; CELF-3, Semel, Wiig, & Secord, 1995) and the Test of Language Development-Intermediate (TOLD-I, Hamill & Newcomer, 1982; TOLD:I-2, Hamill & Newcomer, 1988; or TOLD:I-3, Hamill & Newcomer, 1997). Analyses therefore compared means obtained from the CELF and the TOLD.

RESULTS

Study Selection

Search procedures identified 1,631 articles from database searches and 98 from ancestral searches. After screening titles and abstracts for those 1,729 articles, 194 articles were retrieved and screened. In the first stage of screening, 103 articles were excluded: 74 did not include participants with EBD (e.g., samples were characterized as at-risk, delinquent, ASD, or ADHD), 25 did not include empirical data, and four were unavailable. Interrater agreement for this stage of screening was 97%; disagreements were included in the next round of screening. For the remaining 91 studies, sample age or IQ scores exceeded cutoff criteria in 31 cases, 25 had insufficient data, and 12 included participants with previously identified or suspected LI. These constraints resulted in rejection of several well-known studies (e.g., Baltaxe & Simmons, 1988; Gualtieri, Koriath, van Bourgondien, & Saleby, 1983; McDonough, 1989) but were considered necessary to avoid overestimating prevalence of LI in students with EBD. Similarly, some samples from the Benner et al. (2002) review were rejected entirely or only a subset of the sample was retained (see Table 1). Interrater agreement for screening was 89%; disagreements were resolved by discussion.

TABLE 1

Participant Characteristics Reported for Samples Included in Meta-Analyses

First Author (Year)	*Primary Ethnicity*	*SES*	N	*% Male*	*Age* M (SD) *or Range*	*FSIQ* M (SD), *Range, or Other*
Behar (1982)	—	Low-mid	58	70	9.1 (3.0)	90.7 (12.0)
Benner (2005)	EU	—	84	79	8.6 (1.7)	96.1 (14.8)
Camarata (1988)[b]	—	—	21	77	10.9 (1.2)	Normal limits
Cohen (1989)[a]	—	5.5	37	78	6.9 (1.4)	95.4 (11.0)
Cohen (1993)[b]	—	Low-mid	288	78	8.3 (2.2)	104 (15.2)
Cohen (1998)[b]	—	Low-mid	97	66	10.0 (2.1)	98.0 (11.8)
Curtwright (2007)	Mixed	—	63	86	8.6 (1.7)	< 70 excluded
Giddan (1996)[a]	—	Low-mid	55	72	9.5 (1.9)	—
Griffith (1997)[a, c]	Mixed	Mixed	21	81	8.95 (1.41)	Average
			20	85	10.7 (1.5)	Average
Heneker (2005)	—	—	11	—	5–11	—
Hyter (2001)	—	—	6	100	8.5–12.9	85–115
Keefe (1992)[a]	—	—	19	84	9.67	105 (8.6)
Kotsopoulos (1987)[a]	—	Lower	41	87	8.6	< 70 excluded
Lassman (2007)	—	—	27	—	9–12	—
Mack (1992)[a]	EU	—	20	100	11.8 (1.0)	96.9, 79–123
Miniutti (1991)[a]	Non-EU	Lower	27	85	7.8 (.79)	< 70 excluded
Nelson (2005)[c]	EU	65% FRL	57	82	7.75 (1.18)	93.79 (14.66)
			39	84	10.97 (.86)	100.50 (13.77)
Novak (1992)	Mixed	—	31	77	7–13	< 85 excluded
Rinaldi (2003)	HS	92% FRL	61	84	10.83	89.83 (12.07)
Rogers-Adkinson	—		19	74	11.1	85–128
(2003)[c]			19	79	11.0	
Ruhl (1992)[a]	—	—	30	73	11.9	Normal range
Trautman (1990)[b]	EU	Lower	37	85	6–12	Median 91.5

Note. Hyphen indicates information was not specified in the original research report. EU = greater than 65% European descent/White/Caucasian. Non-EU = sample described as "predominantly non-white." HS = greater than 65% Hispanic. Mixed = no single ethnicity or SES group constituted greater than 65% of the sample. SES = Socioeconomic status. FRL = Free/reduced-price lunch. EU = European descent/White/Caucasian. HS = Hispanic. FSIQ = Full Scale Intelligence Quotient.

[a]Study was included in Benner et al. (2002) review; full sample was included in this study. [b]Study was included in Benner et al. (2002); only part of the sample was included in this study. [c]Data reported for two independent samples.

The final number of articles meeting inclusion criteria was 25; however, three included samples were described in multiple reports. After applying the previously outlined decision rules, 22 studies were retained. No experimental studies were identified (group or single subject designs). In the two quasiexperimental studies (Heneker, 2005; Hyter, Rogers-Adkinson, Self, Simmons, & Jantz, 2001), only preintervention data were used. All other identified studies employed nonexperimental (i.e., descriptive or comparative) group designs. Of the studies identified in the earlier review by Benner et al. (2002), 12 were included in the current analyses (see Table 1). Of the 22 total studies, 12 studies contributed data to mean outcomes, and 18 contained data used to compute prevalence estimates (see Table 2). Data were reported for two independent samples in three

TABLE 2

Study Characteristics and Outcome Variables for Studies Included in Meta-Analyses

1st Author (Year)	*Program*	*Setting*	*Purpose*	*Pub-Loc*	*Measure*	*Language Domain*	*Mean* (SD)	*% Mild*	*% Mod/Severe*
Behar (1982)	Res	Clinical	Practice	J-US	Multiple	Expressive			25.90
						Receptive			12.10
Benner (2005)	Day	School	Research	J-US	CELF-3	Comprehensive		70.24	15.47
						Expressive		71.43	16.67
						Receptive		52.38	10.71
Camarata (1988)	Day	School	Research	J-US	TOLD-I	Comprehensive		23.81	71.43
Cohen (1989)	Day	Clinic	Practice	J-CA	Multiple	Comprehensive			28.57
Cohen (1993)	Day	Clinic	Practice	J-CA	Multiple	Comprehensive			34.38
Cohen (1998)	Day	Clinic	Practice	J-CA	Multiple	Comprehensive		38.10	41.27
Curtwright (2007)	Day	School	Practice	D-US	Multiple	Comprehensive			65.39
Giddan (1996)	Res	Clinic	Practice	J-US	Multiple	Comprehensive			35.00
Griffith (1997)	Res	School	Research	J-US	TOLD	Comprehensive	77.30 (10.43)		
						Expressive	73.80 (11.71)		
						Receptive	84.35 (9.02)		
	Day					Comprehensive	74.45 (11.43)		
						Expressive	70.05 (12.94)		
						Receptive	81.75 (12.87)		
	Mix					Comprehensive		41.46	39.02
						Expressive		39.02	46.34
						Receptive		56.10	7.32
Heneker (2005)	Day	School	Research	J-UK	Multiple	Expressive		54.55	18.18
						Receptive		36.36	27.27
Hyter (2001)	Day	School	Research	J-US	TOLD-I:2	Comprehensive	72.00 (6.20)	66.67	0.00
Keefe (1992)	Day	School	Research	J-US	TOLD-I:2	Comprehensive	92.53 (13.49)		
						Expressive	90.84 (18.03)		
						Receptive	91.05 (12.87)		
						Syntax	87.32 (18.13)		
						Semantics	94.42 (12.96)		
Kotsopoulos (1987)	Day	Clinic	Practice	J-CA	Multiple	Comprehensive		19.51	53.66

continues

TABLE 2. *Continued*

1st Author (Year)	*Program*	*Setting*	*Purpose*	*Pub-Loc*	*Measure*	*Language Domain*	*Mean* (SD)	*% Mild*	*% Mod/Severe*
Lassman (2007)	Day	School	Research	D-US	CELF-3	Receptive	71.48 (16.89)	25.93	48.15
Mack (1992)	Res	Clinic	Research	J-CA	Multiple	Comprehensive		15.00	60.00
Miniutti (1991)	Day	School	Research	J-US	CELF-R	Comprehensive	62.33 (13.20)	0.00	81.00
						Expressive	63.26 (14.10)		
						Receptive	66.96 (12.10)		
Nelson (2005)	Day	School	Research	J-US	CELF-3	Comprehensive	84.95 (15.96)		
						Expressive	84.26 (16.47)		
						Receptive	87.67 (16.02)		
						Comprehensive	86.08 (16.24)		
						Expressive	81.59 (15.63)		
						Receptive	92.54 (17.91)		
Novak (1992)	Day	School	Research	D-US	CELF-R	Comprehensive	65.50 (16.12)	13.33	73.33
Rinaldi (2003)	Day	School	Research	J-US	TOLD-I:3	Syntax	83.51 (14.37)		
						Semantics	80.59 (13.98)		
Rogers-Adkinson (2003)	Day	School	Research	J-US	TOLD-I:3	Comprehensive	67.26 (4.34)		
						Expressive	62.52 (6.66)		
						Receptive	77.26 (7.46)		
						Syntax	65.42 (5.61)		
						Semantics	75.89 (11.19)		
						Comprehensive	85.26 (7.18)		
						Expressive	83.21 (9.89)		
						Receptive	90.21 (9.60)		
						Syntax	83.31 (11.84)		
						Semantics	89.57 (6.29)		
Ruhl (1992)	Day	School	Research	J-US	TOLD-I	Comprehensive	72.70 (12.9)		
						Expressive	74.80 (16.70)		
						Receptive	80.50 (7.70)		
						Syntax	74.7 (14.4)		
						Semantics	79.90 (11.10)		
Trautman (1990)	Day	Clinic	Practice	J-US	Multiple	Comprehensive			27.03

Note. Pub = Publication type. J = Journal. D = Dissertation. Loc = location. US = United States. UK = United Kingdom. CA = Canada. Res = Residential. Mild = Standard scores 71–85. Moderate/severe = ≤ 70 or clinical LI diagnosis.

TABLE 3

Prevalence (Proportion) of Students With EBD and Unidentified Language Deficits

Component	*Severity*	*Prevalence [95% CI]*	N_s	N_p	Q_{total}	% I^2	τ
Comprehensive	Below (≥ 1 *SD*)	80.6% [76, 84]	9	367	7.07	0	0
	Mild (1–2 *SD*)	33.8% [20, 50]	9	367	50.3***	84.1	.89
	Mod/sev (≥ 2 *SD*)	46.5% [36, 57]	14	838	81.78***	84.1	.68
Expressive	Below (≥ 1 *SD*)	85.7% [79, 91]	3	136	1.83	0	0
	Mild (1–2 *SD*)	55.9% [32, 77]	3	136	11.69**	82.9	.77
	Mod/sev (≥ 2 *SD*)	26.5% [15, 43]	4	194	12.20**	75.4	.62
Receptive	Below (≥ 1 *SD*)	64.8% [57, 72]	4	163	1.15	0	0
	Mild (1–2 *SD*)	45.0% [32, 59]	4	163	7.16	58.1	.42
	Mod/sev (≥ 2 *SD*)	17.8% [8, 36]	5	221	21.82***	81.7	.95

Note. Random effects models. Prevalence = Proportion of children with EBD and language deficits. N_s = Number of samples providing prevalence estimates from primary studies in each meta-analysis. N_p = Total number of participants in each meta-analysis. Degrees of freedom = $N_s - 1$. Q_t = Statistic of total heterogeneity across all studies. I^2 = Ratio of true to random heterogeneity. τ = standard deviation. EBD = emotional and behavioral disorders.

** $p < .01$. *** $p < .001$.

studies (e.g., data for different age groups were reported separately), so a total of 25 samples contributed to outcomes. Reliability for coding was 90%; disagreements then were consensus coded.

Participant and Study Characteristics

The total number of participants with EBD included in the sample of 22 primary research studies was 1,171 (range = 6–288; median = 27). The mean age of the sample was 9.49 years (*SD* 1.6 years); however, several authors reported only grade level or age range of the sample or means without standard deviations (see Table 1). Because this and other variables (e.g., IQ, race/ethnicity, socioeconomic status, specific DSM diagnoses or behavioral topographies) were not reported consistently, their roles as moderators could not be examined. The only moderator variable of interest for subgroup analysis reported in every study was measurement instrument (see Table 2). Nearly all of the studies represented samples of convenience; only one (reported in both Benner, 2005 and Nelson et al., 2005) employed random selection from a larger population of students with EBD. Study settings and program types generally were not well-described; thus, designations for program type, setting, and purpose may reflect authors' professional affiliations as much as study or participant characteristics.

Prevalence

Main Effects. Prevalence of below average, mild, and moderate/severe language deficits was computed for comprehensive, expressive, and receptive language ability (see Table 3). Results of a meta-analysis including all children with LI regardless of severity showed that the prevalence of below-average comprehensive language ability was distributed around a mean of 81%, with 34% and 47% of deficits characterized as mild and moderate/severe, respectively (see relevant tables for confidence intervals). Note that summing the point estimates of mild and moderate/severe deficits does not equal the percentage of below average deficits, due to differences in the number of samples and participants contributing data to each outcome.

The next meta-analyses examined expressive and receptive components of language ability (see Table 3); however, data were not available for analysis of semantics or syntax. The prevalence

TABLE 4

Moderator Analyses: Differences in Prevalence of Moderate/Severe Deficits by Study Characteristics

Variable	*Group*	*Prevalence*	*[95% CI]*	N_s	N_p	Q_w	% I^2	τ
Setting	School	56.9%	[33, 78]	7	272	58***	89.6	1.15
	Clinic	38.7%	[32, 46]	7	566	13*	53.8	0.28
Measures	Single	55.2%	[27, 80]	6	157	52***	90.0	1.32
	Multiple	42.6%	[34, 52]	8	629	30***	90.0	0.47
Purpose	Research	55.9%	[31, 78]	7	229	52***	88.5	1.20
	Practice	40.9%	[32, 51]	7	609	27***	77.7	0.46

Note. Random effects models. N_s = Number of samples providing prevalence estimates from primary studies in each meta-analysis. N_p = Total number of participants within samples. Degrees of freedom = $N_s - 1$. Q_w = The statistic of within-group heterogeneity. I^2 = The ratio of true to random heterogeneity. τ = The standard deviation.
*$p < .05$. *** $p < .001$.

estimate for below-average expressive deficits was 86%. The proportions of children experiencing mild and moderate/severe deficits in expressive language were centered around means of 56% and 27%, respectively. Prevalence of receptive deficits was lower than expressive in the below average, mild, and moderate/severe categories (65%, 45%, and 18%, respectively).

Moderator Analyses. Before investigating moderators, it was necessary to determine whether there was sufficient heterogeneity in outcome variables. Examination of heterogeneity statistics in Table 3 reveals very little between-study variance in outcomes in the category for below-average comprehensive language (Q_t = 7.07, $p > .05$). Although the Q statistic may be unreliable when applied to small samples (Borenstein et al., 2009), this conclusion was supported by I^2 = 0%, indicating that the ratio of true to random heterogeneity was insufficient for examining moderator variables within that category. However, heterogeneity was sufficient to examine moderator variables in both the mild and moderate/severe categories. Because (a) fewer studies reported mild language deficits, and precision increases with sample size and (b) conducting tests in both categories would increase the likelihood of Type I error, the decision was made to examine moderators within the moderate/severe category only.

Separate meta-analyses were conducted for three of the five proposed moderator variables: setting, measures, and purpose for data collection (see Table 4). The analyses for program type and topography of behavior could not be conducted as too few studies provided relevant data. In the remaining moderator analyses, noticeable patterns emerged. First, the number of samples in each subgroup was equal or nearly equal; however, the number of participants was two to three times higher in clinical, multiple measures, and practice subgroups (relative to school, single measure, and research subgroups, respectively). Prevalence estimates followed a reverse pattern: Moderate/severe language deficits were 18% higher in schools than clinical settings, 13% higher when obtained by single measures, and 15% higher in studies conducted for research purposes other than those informing practice. Consequently, prevalence estimates for the school, single measure, and research subgroups ranged from 55% to 57%, and estimates for the clinical, multiple measure, and practice subgroups ranged from 39% to 43%. None of these differences were statistically significant.

Mean Scores

Main Effects. Eleven samples contributed data to the meta-analysis of mean comprehensive language ability (see Table 5). The overall mean language score was 76.33. Mean expressive and receptive scores were 75.92 and 82.23, respectively. Although each point estimate was below

TABLE 5

Effect Sizes: Mean Performance on Standardized Language Assessments for Students With EBD

	Mean [95% CI]	N_s	N_p	z	p	Q_{total}	% I^2	τ
Comprehensive	76.33 [71, 82]	11	279	–2.97	0.002**	208.9	95	9.38
Receptive	82.23 [77, 87]	10	270	–1.65	0.13	125.8	92	7.42
Expressive	75.92 [69, 83]	9	243	–1.48	0.004*	137.9	94	9.96
Semantic	84.03 [78, 90]	5	140	–0.31	0.377	43.24	91	6.54
Syntactic	78.63[70, 88]	5	140	–1.30	0.097	93.16	96	10.6

Note. Random effects models. Mean = Outcome point estimate, or weighted mean standard scores. N_s = Number of samples providing mean scores from primary studies in each meta-analysis. N_p = Total number of participants. I^2 = The ratio of true to random heterogeneity. *Z* tests indicate whether means are statistically different from 85, or 1 *SD* below norm means. EBD = emotional and behavioral disorders.
*$p < .05$ **$p < .01$.

the cutoff score of 85, or 1 standard deviation below average language ability, only mean comprehensive ($z = –2.97$; $p = 0.002$) scores reached statistical significance. The analysis to determine whether children had relative strengths or weaknesses in semantic or syntactic skills was a synthesis of five studies, all of which reported results obtained from TOLD composites. Participants scored somewhat higher on semantic (84.03) than syntactic composites (78.63) on this measure, but neither were significantly below average.

Moderator Analyses. The only moderator analysis for mean standard scores was to determine whether differences in comprehensive, expressive, and receptive outcomes varied by test instruments (CELF or TOLD). All of these scores represent children with EBD in school settings only, as mean scores for comprehensive standardized tests were not reported in any of the clinical studies. Generally, scores reported for each measure were very similar, although scores on the CELF were slightly lower and more variable (wider confidence intervals; larger standard deviations), than those obtained by the TOLD. In the moderator analysis for comprehensive language, the CELF was administered to 153 participants in four different samples. The weighted mean score for this subgroup was 74.77. The TOLD was used more often (7 samples), with fewer participants ($n = 126$), and the mean was 77.14. For expressive language, scores from the CELF were 72.91, and scores from the TOLD were 79.17. Scores for receptive language on the CELF were 79.68 and 83.82 on the TOLD. None of the differences in scores obtained by different tests were statistically significant.

Additional Analyses. Analyses were conducted to identify "missing" studies due to publication bias. That is, studies with significant positive results are more likely to be submitted and accepted for publication, published in English, and cited in other publications (Lipsey & Wilson, 2001). Results of Eggers statistical tests, the trim-and-fill method, and an analysis of funnel plots all indicated that small studies were not missing from the analyses and presence of publication bias was unlikely.

DISCUSSION

The purpose of the current study was to synthesize decades of research describing prevalence, severity, and types of unidentified language deficits in the population of students with EBD. Research questions were answered using quantitative methodology designed to synthesize outcomes across primary research reports, and stringent inclusion criteria to control for alternative explanations due to participant characteristics. This study adds to the literature an analysis of mean standardized test scores in several domains of language performance.

The prevalence estimate of previously unidentified language deficits in children with EBD was distributed around a mean of 81%; thus, in answer to the first research question, it is likely that four out of five children with EBD had at least mild LI that escaped the attention of rele-

vant adults. Surprisingly, this estimate was even higher than that reported in Benner et al. (2002), which included a more broadly defined sample of children with EBD. Although results are not directly comparable due to differences in methodology and sampling procedures, the pattern of results was repeated for prevalence of expressive and receptive deficits: Both were higher than the earlier estimate by 21% and 9%, respectively.

Regarding severity of those deficits, of the 838 participants in 14 studies, 47% had deficits categorized as moderate to severe. That is, nearly half the children across studies had either a diagnosis of LI or standard scores below the 3rd percentile in comprehensive language proficiency. The majority of children evaluated in this study had at least a mild language deficit or impairment. This estimate is far higher than in the general population of school-age children, in which prevalence has been estimated at 3% to 14% depending on criteria used to determine case status (Law et al., 2000; Tomblin et al., 1997).

The results of the meta-analyses to determine mean standardized test scores supported the above findings and confirmed the second research question: Language proficiency for students with EBD was well below that of typical peers, even among students without documented deficits. The sample mean comprehensive standard score was 76.33, or 1.5 standard deviation below normative means, which is often cited as the cutoff criteria for determining LI (Law et al., 2000).

Analyses regarding moderators did not reveal variables predicting systematic differences in either prevalence or mean outcomes, indicating that these results may be consistent regardless of how children are identified, where they are served, or how they are assessed. Interestingly, though, there were few differences in the samples included in each subgroup analysis for prevalence. That is, examination of Table 2 shows that three of the subgroups overlapped considerably: studies conducted in school settings also were conducted primarily for research purposes and employed a single measure. On average, prevalence was 15% higher than in studies using multiple assessments in clinical settings in which the purpose of assessment was to inform practice. Clinical studies also were more likely to use a retrospective design, in which data were collected via chart review. Due to lack of available data, the moderator analyses regarding form and severity of problem behavior could not be performed.

Limitations

The current study shares a limitation common to all meta-analytic research: Results are dependent on the quality and quantity of primary studies. The greatest limitation regarding quality was that participants and study characteristics seldom were clearly described. For example, relevant studies may have been excluded if not enough information was provided to determine eligibility. Conversely, if samples were not well-defined, children with overall cognitive, neurological, or attention deficits may have been included inadvertently. Imprecise descriptions may have led to incorrect specification of children's educational placements, which would affect accuracy of the moderator analyses. Inconsistent reporting standards also may have contributed to improper specification of severity categories. Furthermore, limited descriptions of participants also limits external validity or the ability to generalize results.

Regarding quantity of primary studies used to compute results, it was anticipated that there would be more than 18 studies contributing to the prevalence effect sizes, as Benner et al. (2002) identified 18 studies a decade ago. The low number of identified studies was likely due to strict inclusion criteria: Six studies in the Benner et al. review did not meet inclusion criteria for the current analysis. Many more potentially relevant articles also were excluded on the basis of participant age, IQ, or secondary conditions known to cooccur with LI.

Sample size also may have played a role in failure to reject the null hypotheses regarding differences in mean expressive and receptive language proficiency: Larger sample sizes may have increased power to detect an effect. It is also possible, however, that conducting multiple meta-analyses introduced family-wise error and increased the probability of Type I error in analyses of children's comprehensive language abilities. Finally, it is possible that outcomes were affected by variations among test instruments. Although every attempt was made to avoid comparing apples and oranges (Lipsey & Wilson, 2001) by re-

stricting the types of tests that were included in analyses, even direct comparisons of instruments included different versions of tests (and therefore test items, subtests, normative groups, and sensitivity/specificity).

Conclusions and Future Directions

Results of all meta-analyses supported and extended conclusions drawn by Benner, Nelson, and Epstein over a decade ago. Descriptive research has been instrumental in revealing that co-occurrence of mild, moderate, and severe language deficits is clearly widespread among children with EBD. It is important to remember that the children assessed in the primary research studies had no previous history of receiving language-related services, or indeed of ever being previously evaluated for LI. Additionally, it is unlikely that results were attributable to developmental, neurological, or physical (e.g., hearing or speech) factors known to co-occur with both language and behavioral problems. Although it is possible that these results underestimate the problem due to such conservative inclusion criteria, it must also be noted that results may reflect children's behavioral performance during testing (e.g., lack of attention, effort, or cooperation), thus inflating estimates of deficits in language proficiency. Still, the finding that so many children with EBD also had previously unidentified LI echoes the refrain that has been noted throughout the literature: Children's problem behaviors are so salient that they effectively eclipse other intervention needs.

This study also confirms that all children with EBD should be screened for language problems as early as possible. Even without resource-intensive intervention, it is possible that simply educating adults about the link between language and behavior problems could affect some change. Anecdotally, researchers have noted that simply recognizing that problem behaviors such as noncompliance could be in part due to deficits in comprehension helps adults become "less likely to fault the children for their misbehavior" (Cohen et al., 1993, p. 600) and more likely to perceive the child "in a more positive light" (Gallagher 1999, p. 7). Whether this phenomenon can be replicated has yet to be demonstrated empirically.

Another promising area of research is clarifying relationships among linguistic and behavioral variables. For example, results of this and other studies have yet to indicate directionality; that is, whether LI contributes to development of EBD or vice versa. In addition, fine-grained analyses of patterns of language performance among children with different behavioral profiles may be an important precursor to developing targeted interventions. In the current study, some researchers did report behavioral information; however, it was not possible to calculate language performance by behavioral topography given the available data. There are conflicting results in the literature regarding type of problem behavior and components of language (e.g., whether receptive or expressive deficits are more prevalent or severe in children with internalizing or externalizing behavior; see Benner et al., 2002; Cohen et al., 1993; Nelson et al., 2005). To answer these questions, descriptive studies are warranted to investigate how specific types of language and behavior are associated and how those interactions affect children's academic, social, and behavioral development.

Once specific relations among relevant variables are better understood, future research must include interventions to address combined difficulties in language and behavior. Language deficits limit children's ability to benefit from instruction, talk-based therapies, and complex behavior management plans. Interventions by teachers, parents, and therapists must include consideration of children's linguistic needs. In addition, researchers must determine whether interventions to increase children's communication skills will decrease problem behavior.

Language deficits limit children's ability to benefit from instruction, talk-based therapies, and complex behavior management plans.

Another fruitful area for future research may be examining adults' use of language in interactions with children with EBD. If adults use language that is beyond students' comprehension, they may inadvertently increase the occurrence of

problem behavior. Harrison et al. (1996) suggested that verbal instruction may be aversive to students due to a mismatch in the form, content, or function of teacher talk and students' ability to reproduce or comprehend it. Ample evidence supports the presence of coercive interactions and decreased instruction in classrooms for students with EBD, resulting in negative outcomes for teachers and students alike (Sutherland & Morgan, 2003). Determining the effects of teacher talk on children's problem behavior is an important area of study, as is the ability of adults to monitor and modify language use and the effect of those adaptations on children's behavior.

This synthesis of decades of descriptive research confirms that children with EBD are highly likely to have co-occurring language impairment. It is now incumbent upon researchers to identify important targets for change and develop empirically grounded interventions to ameliorate the harmful effects of co-occurring LI and EBD. Supporting language development and effective communication for these children may be a critical step in interrupting the maladaptive academic, social, and behavioral outcomes so often experienced by this population.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

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

Baltaxe, C. A. M., & Simmons, J. Q. (1988). Communication deficits in preschool children with psychiatric disorders. *Seminars in Speech and Language, 9*(1), 81–91. http://dx.doi.org/10.1055/s-2008-1064453

*Behar, D., & Stewart, M. A. (1982). Aggressive conduct disorder of children: The clinical history and direct observations. *Acta Psychiatrica Scaninavica, 65*(3), 210–220. http://dx.doi.org/10.1111/j.1600-0447.1982.tb00841.x

Beitchman, J. H., Cohen, N. J., Konstantareas, M. M., & Tannock, R. (Eds.). (1996). *Language, learning, and behavior disorders: Developmental, biological, and clinical perspectives.* New York, NY: Cambridge University Press.

*Benner, G. J. (2005). Language skills of elementary-aged children with emotional and behavioral disorders. *Great Plains Research, 15*(2), 251–265.

Benner, G. J., Nelson, J. R., & Epstein, M. H. (2002). Language skills of children with EBD: A literature review. *Journal of Emotional and Behavioral Disorders, 10*(1), 43–59. http://dx.doi.org/10.1177/10634266020100010 5

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, England: Wiley. http://dx.doi.org/10.1002/9780470743386

Bradley, R., Doolittle, J., & Bartolotta, R. (2008). Building on the data and adding to the discussion: The experiences and outcomes of students with emotional disturbance. *Journal of Behavioral Education,17*, 4–23. http://dx.doi.org/10.1007/s10864-007-9058-6

*Camarata, S. M., Hughes, C. A., & Ruhl, K. L. (1988). Mild/moderate behaviorally disordered students: A population at risk for language disorders. *Language, Speech, and Hearing Services in Schools, 19*, 191–200.

*Cohen, N. J., Barwick, M. A., Horodezky, N. B., Vallance, D. D., & Im, N. (1998). Language, achievement, and cognitive processing in psychiatrically disturbed children with previously identified and unsuspected language impairments. *Journal of Child Psychology and Psychiatry, 39,* 865–877. http://dx.doi.org/10.1111/1469-7610.00387

*Cohen, N. J., Davine, M., Horodezky, M. A., Lipsett, L., & Isaacson, L. (1993). Unsuspected language impairment in psychiatrically disturbed children: Prevalence and language and behavioral characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry, 32*(3), 595–603. http://dx.doi.org/10.1097/00004583-199305000-00016

*Cohen, N. J., Davine, M., & Meloche-Kelly, M. (1989). Prevalence of unsuspected language disorders in a child psychiatric population. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*, 107–111. http://dx.doi.org/10.1097/00004583-198901000-00020

*Curtwright, B. J. (2007). Prevalence of language disorders among children with severe behavioral problems referred for a psychiatric evaluation by a large urban school district. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 68*(4-A), 1405–1405.

Della Toffalo, D. A., & Pedersen, J. A. (2005). The effect of a psychiatric diagnosis on school psychologists' special education eligibility decisions regarding emotional disturbance. *Journal of Emotional and Behavioral*

Disorders, 13(1), 53–60. http://dx.doi.org/10.1177/10634266050130010501

Donahue, M., Cole, D., & Hartas, D. (1994). Links between language and emotional/behavioral disorders. *Education and Treatment of Children, 17*(3), 244–254.

Friberg, J. C. (2010). Considerations for test selection: How do validity and reliability impact diagnostic decisions? *Child Language Teaching and Therapy, 26(1),* 77–92. http://dx.doi.org/10.1177/0265659009349972

Gallagher, T. M. (1999). Interrelationships among children's language, behavior, and emotional problems. *Topics in Language Disorders, 29*, 1–15. http://dx.doi.org/10.1097/00011363-199902000-00003

*Giddan, J. J., Milling, L., & Campbell, N. B. (1996). Unrecognized language and speech deficits in preadolescent psychiatric services. *American Journal of Orthopsychiatry*, *66*(1), 85–92. http://dx.doi.org/10.1037/h0080158

Gresham, F. M. (2005). Response to intervention: An alternative means of identifying students as emotionally disturbed. *Education and Treatment of Children, 28*(4), 328–344.

*Griffith, P. L., Rogers-Adkinson, D. L., & Cusick, G. M. (1997). Comparing language disorders in two groups of students with behavioral disorders. *Behavioral Disorders, 22,* 160–166.

Gualtieri, C. T., Koriath, U., van Bourgondien, M., & Saleby, N. (1983). Language disorders in children referred for psychiatric services. *Journal of the American Academy of Child Psychiatry, 22*(2), 165–171. http://dx.doi.org/10.1016/S0002-7138(09)62330-X

Hamill, D. D., & Newcomer, P. L. (1982). *Test of Language Development-Intermediate*. Austin, TX: PRO-ED.

Hamill, D. D., & Newcomer, P. L. (1988). *Test of Language Development-Intermediate-2*. Austin, TX: PRO-ED.

Hamill, D. D., & Newcomer, P. L. (1997). *Test of Language Development-Intermediate, Third Edition*. Austin, TX: PRO-ED.

Harrison, J. S., Gunter, P. L., Reed, T. M., & Lee, J. M. (1996). Teacher instructional language and negative reinforcement: A conceptual framework for working with students with emotional and behavioral disorders. *Education and Treatment of Children, 19*(2), 183–96.

*Heneker, S. (2005). Speech and language therapy support for pupils with behavioural, emotional and social difficulties (BESD)-A pilot project. *British Journal of Special Education, 32*, 86–91. http://dx.doi.org/10.1111/j.0952-3383.2005.00376.x

Hooper, S. R., Roberts, J. E., Zeisel, S. A., & Poe, M. (2003). Core language predictors of behavioral functioning in early elementary school children: Concurrent and longitudinal findings. *Behavioral Disorders, 29,* 10–24.

*Hyter, Y. D., Rogers-Adkinson, D. L., Self, T. L., Simmons, B. F., & Jantz, J. (2001). Pragmatic language intervention for children with language and emotional/behavioral disorders. *Communication Disorders Quarterly, 23(1),* 4–16. http://dx.doi.org/10.1177/152574010102300103

Im-Bolter, N., & Cohen, N. J. (2007). Language impairments and psychiatric comorbidities. *Pediatric Clinics of North America, 54*, 525–542. http://dx.doi.org/10.1016/j.pcl.2007.02.008

Individuals With Disabilities Education Act, 20 U.S.C. §§1400 *et seq*. (2006 & Supp. V. 2011)

*Keefe, C. H., Hoge, D. R., Shea, T. M., & Hoenig, G. (1992). Language deficits: A comparison of students with behavior disorders and students with learning disabilities. *ICEC Quarterly, 41*(3), 5–18.

*Kotsopoulos, A., & Boodoosingh, L. (1987). Language and speech disorders in children attending a day psychiatric programme. *British Journal of Disorders of Communication, 22*, 227–236. http://dx.doi.org/10.3109/13682828709019865

*Lassman, K. A. (2007). *Teacher language, student behavior, and receptive language skills in students with emotional behavior disorders* (Unpublished doctoral dissertation). Vanderbilt University, Nashville, TN.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language and Communication Disorders, 35*(2), 165–188. http://dx.doi.org/10.1080/136828200247133

Lipsey, M. W., & D. B. Wilson. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

*Mack, A. E., & Warr-Leeper, G. A. (1992). Language ability in boys with chronic behavior disorders. *Language, Speech, and Hearing Services in Schools, 23*, 214–223.

McDonough, K. M. (1989). Analysis of the expressive language characteristics of emotionally handicapped students in social interactions. *Behavioral Disorders, 14*(2), 127–139.

*Miniutti, A. M. (1991). Language deficiencies in inner-city children with learning and behavioral problems. *Language, Speech, and Hearing Services in Schools, 22,* 31–38.

*Nelson, J. R., Benner, G. J., & Cheney, D. (2005). An investigation of the language skills of students with emotional disturbance served in public school settings.

The Journal of Special Education, 39, 97–105. http://dx.doi.org/10.1177/00224669050390020501

*Novak, J. M. (1992). A pragmatic analysis of communication skills in children with behavior disorders. *Dissertation Abstracts International, 52*(9-B), 4689.

*Rinaldi, C. (2003). Language competence and social behavior of students with emotional and behavioral disorders. *Behavioral Disorders, 29*(1), 34–42.

Rogers-Adkinson, D. L. (1995). *Language disorders in students identified as severely behaviorally handicapped.* Dissertation Abstracts International, A: The Humanities and Social Sciences, *56*(6), 2120-A.

*Rogers-Adkinson, D. L. (2003). Language processing in children with emotional disorders. *Behavioral Disorders, 29*(1), 43–47.

*Ruhl, K. L., Hughes, C. A., & Camarata, S. M. (1992). Analysis of the expressive and receptive language characteristics of emotionally handicapped students served in public school settings. *Journal of Childhood Communication Disorders, 14*, 165–176. http://dx.doi.org/10.1177/152574019201400205

Semel, E., Wiig, E. H., & Secord, W. A. (1987). *Clinical Evaluation of Language Fundamentals—Revised.* San Antonio, TX: Psychological Corporation.

Semel, E., Wiig, E. H., & Secord, W. A. (1995). *Clinical Evaluation of Language Fundamentals—Third Edition*. San Antonio, TX: Psychological Corporation.

Sutherland, K. S., & Morgan, P. L. (2003). Implications of transactional processes in classrooms for students with emotional/behavioral disorders. *Preventing School Failure, 48*, 32–37.

Tomblin, J. G., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*, 1245–1260.

Tomblin, J. G., Zhang, X., Buckwalter, P., & Catts, H. (2000). The association of reading disability, behavioral disorders, and language impairment among second-grade children. *Journal of Child Psychology and Psychiatry, 41*(4), 473–482. http://dx.doi.org/10.1111/1469-7610.00632

Toppelberg, C. O., & Shapiro, T. (2000). Language disorders: A 10-year research update review. *Journal of the American Academy of Adolescent Psychiatry, 39,* 143–152. http://dx.doi.org/10.1097/00004583-200002000-00011

*Trautman, R. C., Giddan, J. J., & Jurs, S. G. (1990). Language risk factor in emotionally disturbed children within a school and day treatment program. *Journal of Childhood Communication Disorders, 13*(2), 123–133. http://dx.doi.org/10.1177/152574019001300201

Zadeh, Y. Z., Im-Bolter, N., & Cohen, N. J. (2007). Social cognition and externalizing psychopathology: An investigation of the mediating role of language. *Journal of Abnormal Child Psychology, 35*, 141–152. http://dx.doi.org/10.1007/s10802-006-9052-9

ABOUT THE AUTHORS

ALEXANDRA HOLLO (Tennessee CEC), Doctoral Candidate; and **JOSEPH H. WEHBY** (Tennessee CEC), Associate Professor, Department of Special Education, Peabody College of Vanderbilt University, Nashville, Tennessee. **REGINA M. OLIVER,** Assistant Research Professor, Center for Child and Family Well-Being, Department of Special Education & Communication Disorders, University of Nebraska-Lincoln.

Alexandra Hollo is now a Postdoctoral Fellow at the University of Louisville.

Preparation of this article was funded in part by the Office of Special Education Programs Leadership Training Grant H325D020022. The opinions expressed in this article are those of the authors and do not necessarily reflect those of the funding agency.

Address correspondence concerning this article to Alexandra Hollo, College of Education and Human Development, Department of Special Education, University of Louisville, Louisville, KY 40292 (e-mail: alex.hollo@louisville.edu).

Manuscript received February 2012; accepted October 2012.