CrossMark

Associations Between Language and Problem Behavior: a Systematic Review and Correlational Meta-analysis

Jason C. Chow[1] · Joseph H. Wehby[2]

Abstract A growing body of evidence points to the common co-occurrence of language and behavioral difficulties in children. Primary studies often focus on this relation in children with identified deficits. However, it is unknown whether this relation holds across other children at risk or representative samples of children or over time. The purpose of this paper is to describe the results of a systematic review and two meta-analyses exploring the concurrent and predictive associations between language ability and problem behavior in school-age children. A systematic literature search yielded 1655 unduplicated abstracts, and a structured study selection process resulted in 19 eligible reports and 25 effect sizes for the concurrent analysis and 8 reports and 10 effect sizes for the predictive analysis. Eligible reports were then coded, and effect sizes were extracted and synthesized via random effects meta-analyses. Results estimate significant negative concurrent ($z = -0.17$ [-0.21, -0.13]) and predictive ($z = -0.17$ [-0.21, -0.13]) associations between language and problem behavior, and these relations hold across age, time, and risk status. Mean effect sizes for receptive and expressive language were significant. This study adds to the quantitative and descriptive literature by summarizing and corroborating the evidence that low language ability is associated with problem behavior. Further research is needed relative to differences in subconstructs of language and behavior, as well as a focus on intervention for students with these comorbid deficits.

✉ Jason C. Chow
  jcchow@vcu.edu

  Joseph H. Wehby
  joseph.wehby@vanderbilt.edu

[1] Department of Counseling and Special Education, Virginia Commonwealth University, 1015 W. Main Street, P.O. Box 842020, Richmond, VA 23284, USA

[2] Department of Special Education, Vanderbilt University, 230 Appleton Place, Peabody Box 228, Nashville, TN 37203, USA

Springer

Concomitant language and behavioral deficits in children and youth have been well documented in the research literature (Benner et al. 2002; Hollo et al. 2014; Yew and O'Kearney 2013). In spite of these known relations, the majority of children with emotional and behavioral disorders (EBD) are likely to have unidentified language deficits, as the immediate need for behavioral intervention may eclipse focus on diagnosis and intervention targeting language deficit; that is, problem behavior is often a more immediate concern relative to the impact it can have on classroom and school environments. Researchers have estimated that 68 to 97 % of students with emotional disturbance (ED) experience clinical language deficits (Camarata et al. 1988; Nelson et al. 2005), and a recent meta-analysis estimated that 81 % of students with EBD had language deficits that were unidentified (Hollo et al. 2014), highlighting that language deficits in these children went untreated, and these children likely only received services for their behavioral problems.

Correlational research points to a general negative association between language ability and problem behavior (Bornstein et al. 2013). In a meta-analytic review of published literature, children with specific language impairment (SLI) were twice as likely to develop a subsequent internalizing, externalizing, and attention-related behavior difficulty (Yew and O'Kearney 2013). While children with behavior disorders often experience language impairment that has gone undiagnosed, it is also true that many children with diagnosed language impairment will go on to develop behavior problems (St Clair et al. 2011; Yew and O'Kearney 2015). Other related populations have exhibited similar deficits. For example, delinquent youth and juvenile detainees have presented severe oral language deficits, particularly in receptive language (Lansing et al. 2014). This finding is relevant, as youth in correctional facilities share similar behavioral histories with students with EBD (Bradley et al. 2004). For the purposes of this paper, we will refer to this broad group of children and youth as having *behavior problems* or presenting *problem behavior*.

Because problem behavior and low language are known to contribute to long-term negative outcomes (Bradley et al. 2008; Nation et al. 2004), the relation between these constructs will be important in the development and implementation of intervention. While these deficits often co-occur, the direction of the relation is less clear. Hinshaw (1992) proposed four potential scenarios relating problem behavior to underachievement: (1) language deficits may lead to problem behavior; (2) problem behavior may hinder typical language development; (3) the relation is bi-directional; or (4) a mediating variable, or set of underlying variables, may explain this robust research finding. However, there is little understanding regarding the direction of these potential relations. As such, a clear quantification of the relation between linguistic and behavioral performance remains unknown. While the co-occurrence of these two factors is well documented, it is unclear how language disorders and behavioral difficulties may be related. For the purpose of this paper, we use this Hinshaw's framework to conceptualize the associations of interest.

Theoretical Framework

We posit that language is a fundamental underlying construct for learning and necessary for social development and that low language can subsequently negatively impact achievement and social behavior (see Fig. 1). Language influences academic skills and behavior, and academic skill and behavior influence each other. Success in both these constructs predict achievement, social success, and subsequently, later life outcomes.
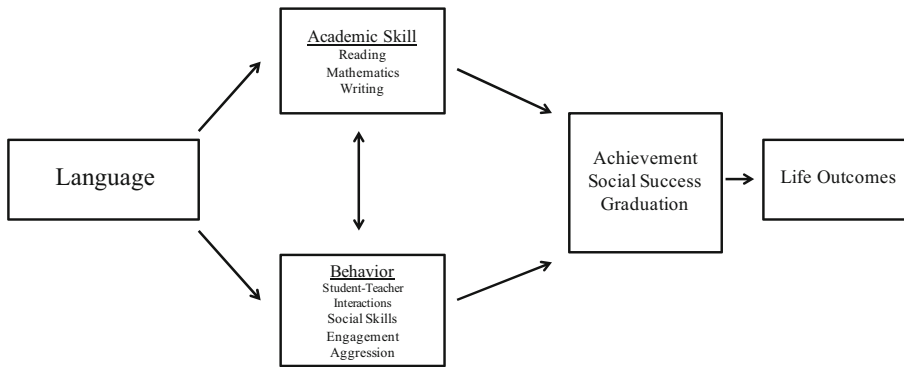
Fig. 1 Theoretical framework of the influence of language on achievement and behavior

Academic Skill

Oral language performance has been shown to predict achievement in both reading (Dickinson et al. 2010), writing (Dockrell et al. 2009; Kent et al. 2014), and mathematics (Chow and Jacobs 2016; Duncan et al. 2007; Fuchs et al. 2005). Children who struggle behaviorally have severe academic (Reid et al. 2004) and linguistic (Hollo et al. 2014) deficits, and children with LI are also at higher risk for academic failure (Cantwell and Baker 1980; Yoshimasu et al. 2011). Further, research continues to reveal low academic participation of students with behavior problems (Scott et al. 2011) as well as students with LI (Fujiki et al. 1999). Because of this consistent trend of low academic performance, understanding the role of classrooms and the academic context may provide an important conceptual starting point.

Understanding the Contribution of Classroom Context

When a child enters school, language forms the necessary foundation for academic and behavioral success (Tomblin et al. 2000). Language is required to understand academic instruction through listening to and reading content and to demonstrate understanding by speaking and writing. As comprehension demands of classroom discourse become less related to the instructional context as students progress through school (Nelson 1985), an even more complex environment with varying levels of linguistic demands unfolds. Effective communication skills are essential, as evaluation of understanding happens via speaking and writing (Cazden 2001).

In general, the research on these co-occurring deficits is student-centered. Descriptions and implications of the relation between language and behavior focus primarily on student-level characteristics. Because teachers use language as the primary mode of instruction, and students use language to interact with peers, proficiency in language is essential to navigate the academic and social context of classrooms and schools. Classroom interactions heavily and importantly involve teachers, and communicative clarity may be more valued in school than in any other context children encounter (Harmon and Watson 2012; Peets 2009); that is, the expectation of linguistic accuracy and communicative ability may be higher in the classroom than in other environments (e.g., playground, cafeteria, home). Further, it is likely that this consistency of the expectation of linguistic and communicative competency varies across different academic, social, and family contexts (Cazden 2001; Peets 2009). Thus, there is a need to understand the expectation of classrooms and how it may differ from other contexts.

Because language is used differently in different contexts, the concept of a "correct" English language and the potential consequences are important. Harmon and Watson (2012) argue for a "broadening of what we mean by 'grammatical' in order to recognize the fundamentally complex language structures that all speakers use" (p. 29). The *prescriptive* grammar movement began to describe many of the "wrongs" and "rights" of the English language (Parker and Riley 2005). A prescriptive approach to grammar contains a set of rules based on how to correctly use language and is the common standard in most classrooms in the USA. In contrast, *descriptive* grammar refers to how language is typically used. While students may communicate effectively outside of school with peers and family via descriptive grammar, one supposition is that the traditional prescriptive approach to language and teaching in the classroom may serve as a catalyst for the increasing mismatch between the linguistic expectations (and subsequent feedback) a child experiences. For example, if a child uses the phrase "she went" with the intention of communicating the phrase "she said," a prescriptive approach would consider "she went" to be incorrect, while a descriptive approach would consider it effective communication. The higher linguistic demands of the classroom and the discrepancy between prescriptive and descriptive grammar may be particularly relevant to teachers of students who engage in problem behavior, as an association between teacher instruction and student disruptive behavior has been demonstrated (Gunter et al. 1994).

Negative Student-Teacher Interaction

Students with behavior problems often are identified as having expressive language deficits that result in an inability to engage in and maintain fluent conversations (McDonough 1989), and receptive language deficits that impair their ability to comprehend abstract concepts and decontextualized language (Warr-Leeper et al. 1994). According to Patterson and Reid's (1984) theoretical categorizations based on the study of family member interaction, these unsuccessful communicative interactions between teachers and their students represent a coercive interaction model. Patterson (1982) describes this as a negative reinforcement trap—an interaction created when both parties are reinforced by behaviors that ultimately have a negative impact on desired outcomes. In the classroom, this cyclic interaction pattern may have devastating consequences for students' academic and social outcomes.

Harrison et al. (1996) proposed a conceptual framework for understanding teacher instruction and negative reinforcement. The authors hypothesized that teacher instructional language may be aversive stimuli to students with behavior problems and that these stimuli may maintain and reinforce problem behavior. For example, researchers have suggested that linguistically complex instruction may be misunderstood by or completely incomprehensible to students with language deficits (Kevan 2003). In a large meta-analysis, Titsworth et al. (2015) estimated that teacher clarity accounted for 13 % of the variance in student achievement. This finding suggests that instructional clarity is a meaningful predictor of student learning.

Because researchers have reported this negative cycle of interaction in classrooms with students at-risk for or with behavior problems (Sutherland and Morgan 2003) and many students with behavior problems have low language skills, it is possible that students engage in problem behavior to avoid instructional interactions they do not understand or comprehend and that the problem behavior is reinforced by the delivery of a desired consequence (e.g., escape/avoidance of academic demands). This is consistent with the hypothesis that negative social behavior may function as a communicative action.

Intuitively, the more negative interactions (and escape from academic tasks) occur, the less instruction students receive. One study estimated that teachers of students with behavior problems devote only 30 % of time to academic instruction (Wehby et al. 2003). Because these students consistently present low rates of academic engagement (Scott et al. 2011) compared to peers without disabilities (Briesch et al. 2014), the dismal academic performance of students with behavior problems is no surprise.

Given that (a) low language is associated with poor achievement, (b) children with problem behavior show significant academic deficits, and (c) low language is highly prevalent in children with behavior problems, it is likely that deficits in achievement, behavior, and language are inextricable. As we work toward the goal of improving the educational outcomes of students who struggle, considering the role of language and behavior on achievement may help our schools, teachers, and instruction increase the number of students for which our educational efforts are effective.

Purpose

To our knowledge, no systematic review has quantified and tested the significance of the correlation between language and problem behavior. Further, it is unclear for whom this association is relevant. While prevalence of LI in students with behavior problems and risk of later problem behavior for students with LI has been estimated, we aim to extend the literature by estimating the correlation between language and problem behavior in a more representative, heterogeneous sample of children. If measures of language and behavior are significantly related in representative and typically developing samples of children, attention to this relation may be important in general curriculum, and not just for students with identified deficits. Thus, we systematically review and statistically aggregate the magnitude of both the *concurrent* and *predictive* associations between standardized measures of language and problem behavior. The majority of research that examines this relation focuses on populations with a pre-existing deficit. Given the developmental continuum and the subjectivity that accompanies deficit, disability, and diagnosis identification, we aim to expand the focus and include studies with representative samples of children in an effort to determine the scope of the relation between language and behavior.

This systematic review reports findings from two separate meta-analyses. First, we estimate the magnitude of the *concurrent* correlation between language ability and problem behavior in school-age children. Then, we use qualifying studies to estimate the magnitude of the *predictive* correlation between language ability and future problem behavior. Further, we explore whether the overall correlation differs by subconstruct of language, subconstruct of behavior, age, and risk status. We explore potential subconstruct differences because primary research indicates that there may be differences in type of language skills across children with and without other disabilities, and externalizing and internalizing behaviors manifest themselves differently in children, and likely influence subsequent ratings of those behaviors. Thus, the following research questions drive our review:

1. What is the concurrent correlation between language ability and problem behavior in school-age children?
2. What is the predictive correlation between language ability and problem behavior in school-age children?

3. Do age, risk status, and measurement time predict individual study effect size (i.e., moderate)?

4. Are there substantively meaningful differences between subconstructs of language and behavior?

Methods

We conducted a systematic review of the literature to identify all potential eligible published and unpublished studies and used random-effects meta-analysis to synthesize data from primary reports (Lipsey and Wilson 2001). In meta-analysis, effect sizes are weighted based on individual study characteristics with the purpose of maximizing statistical reliability. Outcome and relevant variables were coded in each eligible study, and effect sizes were extracted for quantitative synthesis. Finally, meta-analyses and tests for moderators and publication bias were conducted within Stata/SE 12.0 (Stata 2011).

We used the same search procedure to identify reports for both meta-analytic reviews. In the final stage of eligibility identification, reports were coded as eligible for the current analysis, the predictive analysis, or both analyses.

Eligibility Criteria

Eligible studies were those that measured language and problem behavior at one or multiple time points. Eligible studies included participants, both typically developing and those with disabilities (e.g., high incidence) that did not have cognitive or other low-incidence deficits. Participants were defined as individuals between the ages of 3 and 21 (consistent with the age range of students covered under Part B of the Individuals with Disabilities Education Act (IDEA 2004)). Additionally, studies that focused primarily on English language learners or speakers of African-American English were not eligible, and language must have been assessed via measures in the primary language of participants. Further, studies that examined children with hearing loss and other medical conditions were excluded in an effort to rule out any behavioral or language performance deficits explained by cultural or other health-related differences. Intervention studies were only included if authors reported concurrent pre-test correlations. If multiple studies analyzed the same dataset (e.g., an extant data analysis or multiple papers from the same research group), we randomly selected one study to be included due to dependent samples.

Relative to language measures, eligible studies measured oral language and did not require the student to read words or respond to text prompts, tapping expressive or receptive language, only requiring the child to respond to verbal prompts or elicit a verbal response. This is an important distinction because, while interrelated, the cognitive processes of oral language, reading, and writing are substantively different. Because problem behavior may play a communicative, compensatory role due to low language skills (Hollo and Chow 2015), it was essential that the construct of interest included was that of *oral language*. Measures of behavior were those that indexed levels of *externalizing behavior* (e.g., aggression, conduct problems). This is because we are interested in synthesizing studies that measure the types of behaviors that are most problematic to teachers in classrooms. Studies that included measures of internalizing behavior only (e.g., only anxiety, depressive symptoms) were not included in the main analysis. In

addition, a defining characteristic for inclusion was the use of norm-referenced, standardized, psychometrically sound assessments.

Search Strategy

A comprehensive search strategy was used in an attempt to identify, retrieve, and code the entire population of eligible studies. No date restrictions were placed on any searches. We also aimed to guard against publication bias, the potential for the sample of studies published in peer-reviewed journals being substantively different from those that were not published. The following electronic databases were iteratively searched, current through December 15, 2015: ERIC, Linguistic and Language Behavior Abstracts, ProQuest, PsycARTICLES, PsycINFO, PubMed, Social Services Abstracts, and Sociological Abstracts. The search strategy included terms such as the following: (language OR "language ability" OR "verbal ability" OR "verbal language"); (behavio* OR "problem behavio*" OR "behavio* problem" OR maladaptive OR devian*); (child* OR student* OR teen* OR adolescen*); (predict* OR relation* OR associat*). In an attempt to identify grey literature, we searched the Campbell Collaboration, Cochrane Collaboration Central, Dissertations and Theses (Vanderbilt University and Global), Index to Theses in Great Britain and Ireland, and several conference websites. Reference lists of narrative and systematic reviews that were flagged during the abstract-screening phase were screened.

Screening and Coding

Our search strategy resulted in an initial pool of studies to be screened. First, titles and abstracts were reviewed, and any study that was clearly ineligible was eliminated. Studies that had any potential to be included were moved on to the next round of screening. For studies that were not excluded at the title and abstract stage, the first author conducted a full-text review. To ensure that eligible studies were not being excluded, a Master's level research assistant reviewed every study that was excluded by the first author at the full-text level. Any inconsistencies ($n = 1$) in the study exclusion process were resolved by discussion.

After abstract and full-text screening, a final set of eligible reports was identified for synthesis. The first author coded all studies on all relevant variables identified a priori: population [i.e., sample size, percent male, disability (if applicable), age (mean and standard deviation or range)], outcomes [i.e., language measure(s), behavior measure(s), correlation], study design; and report characteristics (i.e., year of publication, report type). If studies reported both, receptive and expressive language ability or internalizing and externalizing subscales of behavior, they were coded separately. For reliability, we randomly selected 30 % of the full-text articles, and a research assistant independently coded all variables and effect sizes in 30 %. Agreement was 100 %.

Summary Measures

Each primary study provided one or more effect sizes for analysis. If studies reported more than one eligible effect size, the following considerations were made. First, if a study reported the correlations between problem behavior and both receptive and expressive language, the correlations were averaged, resulting in an aggregate of language ability, which we equate with composite or comprehensive language scores. Secondary analyses estimate the relation between (a) total problem behavior and *receptive language*, (b) total problem behavior and

expressive language, (c) total language and *internalizing behavior*, and (d) total language and *externalizing behavior*. For studies that reported repeated measures of language and behavior, the time point from which to extract the effect size(s) was randomly selected.

The first author also contacted authors ($n = 11$) who used statistical techniques to examine the relation of interest and requested bivariate correlations. For example, the correlation between a latent language variable is substantively different than the simple aggregate of manifest variables of language due to the ability of confirmatory factor analysis to parse out measurement error from observable scores. This effort yielded a return of five reports. Finally, all correlations were transformed using Fisher's z transformation prior to synthesis. These reasons were twofold: First, this transformation has the effect of making the original correlations more normally distributed, which is advantageous when aggregating across studies as well as for assessing potential publication bias (Duval and Tweedie 2000). Second, this transformation guards against potential distortion of the standard errors of correlations in each individual study.

Analytic Strategies

Aggregating Effect Sizes Across Studies

In this meta-analysis, we used a random-effects model to account for expected between-study variability. Individual study effect sizes were combined to calculate an overall mean effect size. We conducted separate exploratory analyses to examine potential differences relative to different language and behavior constructs.

Heterogeneity

To examine heterogeneity, we examined the Q, τ^2, and I^2 statistics to evaluate the potential within-study heterogeneity in our sample. The Q indicates whether or not there is significant heterogeneity within the sample. The τ^2 describes the width of the distribution of effect sizes of studies included in this meta-analysis. The I^2 is an index of the amount of heterogeneity that is true between-study variability. In meta-analyses, it is essential to examine all three of these indices in tandem in order to best understand the amount and scope of heterogeneity within a particular collection of studies.

Moderator Variables

We conducted moderator analyses to examine whether the mean effect size differed based on within-study sample characteristics. We used meta-regression to identify whether the relation between language ability and problem behavior differed as a function of individual study sample age and risk status. We were interested in exploring if individual study effect size different based on participant mean age, or whether or not the sample had high-incidence disabilities or were at-risk for poor outcomes [e.g., low socioeconomic status (SES)].

Sensitivity and Publication Bias

Sensitivity analyses are important to perform in meta-analyses because they assess the extent to which results are robust to the assumptions and decisions made during the synthesis process (Borenstein et al. 2009) and to test for outliers. We also tested for bias

introduced to the sample by studies that utilize datasets for which imputation techniques were required. We used funnel plots, trim and fill analyses (Duval and Tweedie 2000), and Egger's tests (Egger et al. 1997) to explore whether potential publication bias existed in our sample of studies due to the lack of studies with small sample sizes.

Results

Study Selection

Figure 1 describes the process for which studies were identified for inclusion in the concurrent analysis. Electronic databases yielded 1625 reports, and 30 additional reports were identified via conference websites, literature reviews, and references lists. Titles and abstracts were screened, and 1362 studies were removed. A full-text review of the remaining 293 reports was conducted, resulting in 94 reports for detailed evaluation.

We performed the same search procedures for the second analysis. Inclusion criteria for this analysis were identical to those for the concurrent analysis until the final stage of eligibility screening. Here, studies that assessed language prior to the assessment of behavior were identified, and separate predictive correlations were extracted. At this final stage, 19 reports yielded 25 unique effects sizes for the concurrent analysis, and eight unique reports yielded 10 predictive effect sizes. Table 1 identifies all studies included in this systematic review and indicates the reports that contributed effect sizes to the concurrent, predictive, or both analyses.

Study Characteristics

Descriptive study characteristics are presented in Tables 1 and 2. All included studies used at least one standardized measure of language and behavior. Correlations were averaged for studies that measured either construct with more than one eligible measure. This sample of studies includes 15 peer-reviewed journal articles, four dissertations, and two conference presentations. Four studies used extant datasets that imputed data; the remaining 17 analyzed researcher-recruited samples. Relative to participant characteristics, seven studies included representative (i.e., sampled to represent the population) or typically developing samples, six included at-risk populations, and three studies recruited separate at-risk and comparison groups. In the concurrent analysis, the mean age of participants across included studies was 7.8 years, ranging from 4 to 16. Average effect size sample size across studies was 917 (range = 20 to 11,506; median = 139) in the concurrent analysis and 1976 (range = 75 to 11,506; median = 362) in the predictive analysis. In total, the concurrent and predictive analyses included data from 22,927 and 19,760 participants, respectively.

Analysis of Concurrent Effect Sizes

The primary results from a random-effects meta-analysis examining the concurrent association between language ability and problem behavior is presented in Fig. 2. Effect sizes and corresponding confidence intervals (95 %) are presented. The shaded box around each individual effect size visually represents the weight allocated to each effect size. The dashed line represents the significant overall mean effect size ($z = -0.17$; $p < 0.001$) and the overlaying diamond represents the confidence interval (95 % CI = [−0.21, −0.13]). Further, the prediction interval (95 % PI =

Table 1 Descriptive study characteristics

Study	Number	Age (years)	Percentage of male	Participants	Inclusion (study number)
Aro et al. (2012)	107	5	50.5	AR for dyslexia	1
Aro et al. (2012)	92	5	56.5	TD	1
Bornstein et al. (2013; study 1)	85	5.75	47.1	RS	1 + 2
Bornstein et al. (2013; study 2)	139	9.33	60.4	RS	1 + 2
Botting and Conti-Ramsden (2008)	139	16	69.8	SLI	1
Botting and Conti-Ramsden (2008)	124	16	62.1	TD	1
Bretherton et al. (2014)	1275	4	50.9	TD	1
Burchinal et al. (2006)	75		45	Head start (or similar)	2
Chow and Hollo (2014)	29	11.9 (1.9)	100	ED	1
•Chow and Wehby (2015)[c]	300	7.5	78	EBD or AR	1
•Clark (2011)[b]	126	4.88	59	TD	1
Ford et al. (2007)	111	12.1 (1.2)	41	Low SES	1 + 2
Haak et al. (2012)	1226	4.5	51.7	RS	1
•Khourdaji (2012)[b]	1364		51.7	RS	2
•Lassman (2006)[b]	26	5th–6th grade	–	EBD	1
Lindsay et al. (2007)	69	7.5–8.8	75.4	SLI	1 + 2
Nelson et al. (2005)	166	11.72 (3.67)	81.9	ED	1
•Pagani et al. (2010)[a]	1155	5.5	46.6	RS	1 + 2
•Peterson et al. (2013; Study 1)[a]	585	9	52.0	RS	1 + 2
•Peterson et al. (2013; Study 2)[a]	11,506	6.5	51.0	RS	1 + 2
Qi and Kaiser (2004)	60	4.5 (0.68)	53.3	SLI+TD	1
Redmond (2011)	20	7.8	60.0	SLI	1
Redmond (2011)	20	7.8	75.0	ADHD	1
Redmond (2011)	20	7.8	55.0	TD	1
•Spilt et al. (2015)[a]	4707	4–5	50.9	RS	1 + 2
•Spira (2006)[b]	139	5.5	48	Head start	1
Tomblin et al. (2000)	581	8	57	SLI+TD	1

En dash indicates that the data was not presented in report. Age not provided for samples only included in study 2. If participants were not all the same age, mean and *SD* or range is provided

AR at risk, *TD* typically developing, *RS* representative sample, *SLI* specific language impairment, *ED* emotional disturbance, *EBD* emotional and behavioral disorders

[a] Reports that used imputation

[b] Unpublished dissertations

[c] Conference presentations

[−0.30, −0.04]) is represented by the horizontal line extending from each side of the diamond. Because the prediction interval does not cross zero, this suggests that there is a 95 % chance that a new eligible study will report a negative association between language eligible and problem behavior. This prediction interval is different than the confidence interval, which describes the

**Table 2** Study measures and correlations

Study	Language measure(s)	Behavior measure(s)	r value
Aro et al. (2012; TD)	PPVT-R+BNT	BASC	0.0015
Aro et al. (2012; AR)	PPVT-R+BNT	BASC	0.051
Bornstein et al. (2013; Study 1)	TOLD-P+MSCA	PBQ	−0.3 (−0.27)
Bornstein et al. (2013; Study 2)	WJ-R+VABS	CBCL	−0.207 (−0.27)
Botting and Conti-Ramsden (2008; SLI)	CELF-R	SDQ	−0.11
Botting and Conti-Ramsden (2008; TD)	CELF-R	SDQ	−0.155
Bretherton et al. (2014)	CELF-P2	SDQ	−0.225
Burchinal et al. (2006)	CELF+PPVT	SSRS	(−0.25)
Chow and Hollo (2014)	CASL	TRF	−0.153
Chow and Wehby (2015)	WJ-3	TRF	−0.105
Clark (2011)	CELF-P2	PSBS	−0.245
Ford et al. (2007)	PPVT-3	TRF	−0.229 (−0.236)
Haak et al. (2012)	PLS-3	CBCL	−0.145
Khourdaji (2012)	WASI	CBCL	(−0.056)
Lassman (2006)	CELF-3	CBCL	−0.093
Lindsay et al. (2007)	CELF-R	SDQ	−0.327 (−0.30)
Nelson et al. (2005)	CELF-3+WISC-3	TRF	−0.079
Pagani et al. (2010)	PPVT	SBQ	−0.13 (−0.17)
Peterson et al. (2013; study 1)	NNSAT	TRF	−0.35 (−0.27)
Peterson et al. (2013; study 2)	PPVT-R	CBCL	−0.16 (−0.14)
Qi and Kaiser (2004)	PLS-3+PPVT-3	TRF+SSRS	−0.259
Redmond (2011)	TEGI+TNL	CBCL	−0.363
Redmond (2011)	TEGI+TNL	CBCL	0.272
Redmond (2011)	TEGI+TNL	CBCL	−0.063
Spilt et al. (2015)	PPVT-3	SDQ	−0.13 (−0.14)
Spira (2006)	PPVT-3	BASC	0.08
Tomblin et al. (2000)	PPVT-R+CELF+CREVT	CBCL	−0.30

Predictive correlations presented in parentheses

PPVT Peabody Picture Vocabulary Test (revised; 3rd edition), *BNT* Boston Naming Test, *TOLD–P* Test of Oral Language Development–Primary, *MSCA* McCarthy Scales of Children's Abilities, *WJ* Woodcock Johnson (revised; 3rd edition), *VABS* Vineland Adaptive Behavior Scales, *CELF* Clinical Evaluation of Language Fundamentals (revised; preschool–2; 3rd edition), *CASL* Comprehensive Assessment of Spoken Language, *PLS–3* Preschool Language Scale–3rd edition, *WISC–3* Wechsler Intelligence Scale for Children–3rd edition, *NNSAT* Nationally Normed Standardized Assessment Test, *TEGI* Test of Early Grammatical Impairment, *TNL* Test of Narrative Language, *BASC* Behavior Assessment System for Children, *PBQ* Preschool Behavior Questionnaire, *CBCL* Child Behavior Check List, *SDQ* Strengths and Difficulties Questionnaire, *SSRS* Social Skills Rating System, *TRF* Teacher Report Form, *PSBS* Preschool Social Behavior Scale, *SBQ* Social Behavior Questionnaire

accuracy of the overall mean effect size. Together, these estimates suggest that lower language is significantly related to higher problem behavior in children.

Results also indicate that there is significant heterogeneity across the sample of studies ($Q = 76.15$, $p < 0.001$). The I^2 (68.5) attributes 68.5 % of the variability to true between-study heterogeneity; that is, almost three quarters of the detected heterogeneity between studies is not
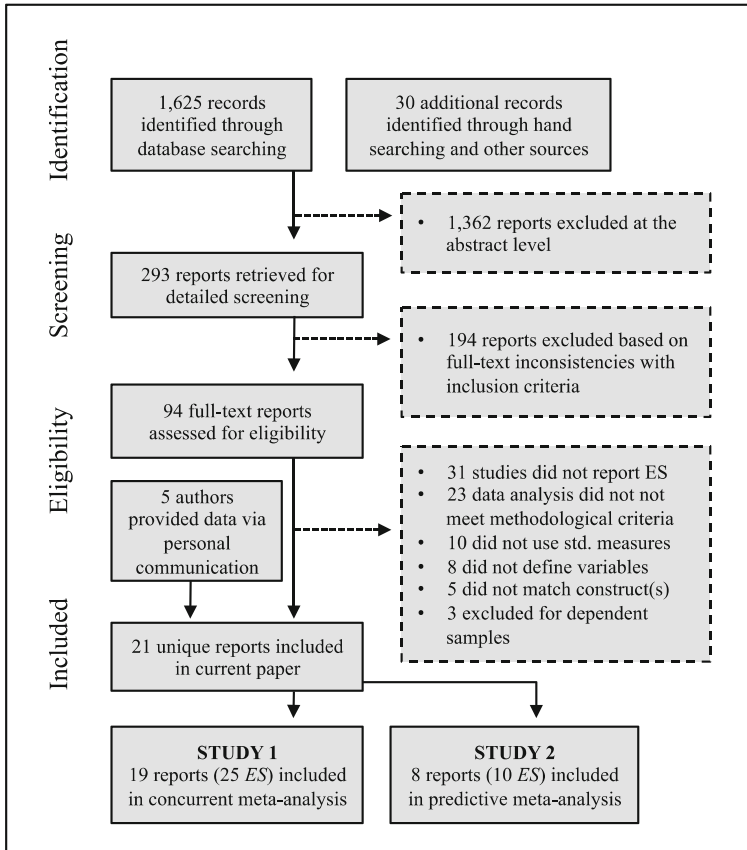
Fig. 2 Study identification flow diagram. *ES* effect size

due to chance. The $\tau^2 = 0.0034$ ($\tau = 0.058$) describes the distribution of potential effect sizes (θ). Although the τ^2 suggests some homogeneity, the Q, $\hat{I}^2$, and prediction interval point to a significant amount of heterogeneity and merits further exploration via moderator analyses.

Moderators

Meta-regression was used to explore age and risk status as moderators of effect size magnitude. Comparable to multiple regression, with meta-regression, study-level predictors are modeled to determine the relation between the moderator and mean weighted effect size. Because the final search for concurrent studies yielded 25 effect sizes, both predictors were entered in the same model. Results yielded neither age ($\beta = -0.006$, $p = 0.44$, 95 % CI = [−0.02, 0.01]) nor risk status ($\beta = 0.08$, $p = 0.19$, 95 % CI = [−0.04, 0.12]) as a moderator.

Additional Analyses

A final set of analyses focused on sensitivity and publication bias. Specifically, we tested whether our findings are robust to the assumptions and decisions made during the meta-analytic process. First, we tested whether the contribution of studies that used imputed data biased our results and

determined that these studies did not significantly predict the effect size ($p = 0.13$) after controlling for sample size. This suggests that there is a trivial relation between studies that imputed data and its corresponding effect size. Relative to publication bias, a funnel plot (see Fig. 3) suggests that publication bias is unlikely based on visual inspection of distribution of studies based on their standard errors and mean effect sizes. To further probe this question, we conducted a trim and fill analysis that yielded one filled effect size. However, the change between the random effects estimates of the original model ($z = -0.169$) and the filled model ($z = -0.171$) was negligible, suggesting an unbiased sample. Finally, the Egger's test ($p = 0.885$) provided additional evidence against publication bias. Overall, the synthesis of these tests suggests that this meta-analysis is free of outliers or publication bias.

Exploring Differences by Language and Behavior Constructs

A secondary aim of this study was to explore whether the mean effect size varied based on subconstructs of language and behavior. Thirteen studies allowed for comparison of receptive and expressive language, and nine studies allowed for comparison of internalizing and externalizing behavior. Mean effect sizes comparing receptive (-0.16 [-0.21, -0.10]) and expressive language (-0.13 [-0.21. -0.06]) and internalizing (-0.11 [-0.16, -0.05]) and externalizing behaviors
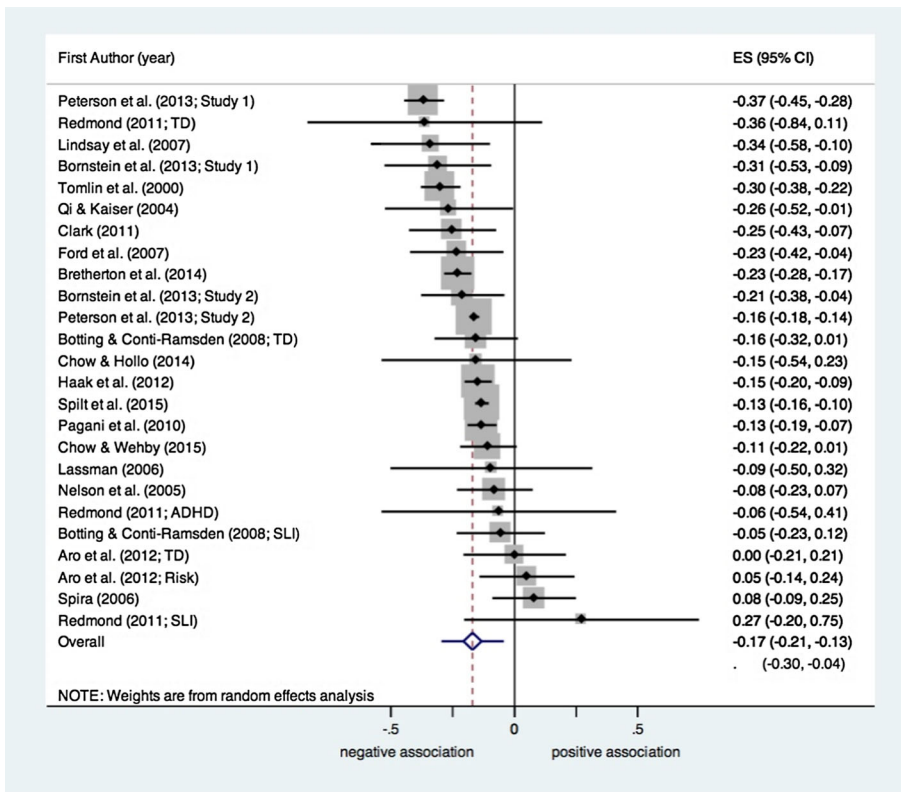
Fig. 3 Forest plot of random-effects meta-analysis examining the association (Fisher's *z*) between language ability and problem behavior. *Negative correlations* represent low language being associated with greater problem behavior(s)

(−0.15 [−0.23, −0.08]) were all significant at $p < 0.000$. These exploratory results suggest that these effect sizes are highly significant, but we cannot make statements about how different they are from each other. In both sub-analyses, the confidence intervals either almost or fully overlap.

Is Early Language Ability Associated with Later Problem Behavior?

In this analysis, we examined whether early language was associated with later problem behavior. Due to our sample size, we did not have the power to examine whether early language predicts later problem behavior [e.g., regress language (T1) on behavior (T2) controlling for initial behavior (T1)]. Thus, we conducted a random-effects meta-analysis to aggregate the bivariate correlations. To be eligible for this study, the measurement of oral language must have preceded measurement of behavior. The mean time between measurement points was 3 years (range = 1 to 5.5 years). The primary results from a random-effects meta-analysis are presented in Figs. 4 and 5. The overall mean effect size ($z = -0.17$; 95 % CI = [−0.21, −0.13]) was significant, and the prediction interval (95 % PI = [−0.27, −0.06]) also suggests that a new eligible study would likely yield a negative effect size.

Heterogeneity analyses indicate that there is significant heterogeneity across the sample of studies in this meta-analysis ($Q = 28.16$, $p = 0.001$). The I^2 estimated that 68 % of the variability was due to true between-study heterogeneity. The $\tau^2 = 0.0017$ ($\tau = 0.041$) describes the distribution of individual-study mean effect sizes (θ). Although the τ^2 estimate suggests some homogeneity, the Q and I^2 point to a significant amount of between-study heterogeneity.

Moderators

We conducted two exploratory single-predictor moderator analyses. We were interested in examining whether the length between measurement points and risk status predicted the effect
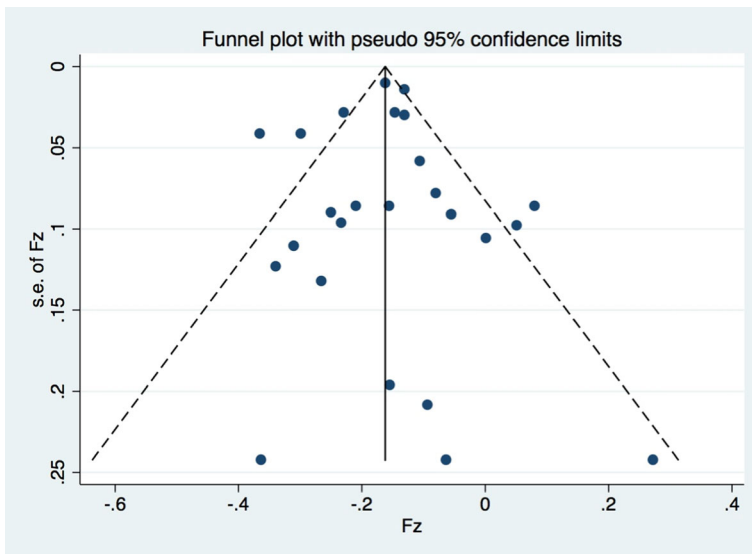
Fig. 4 Funnel plot of the individual study Fisher's z. The x-axis is the Fisher's z effect size and the y-axis is the standard error associated with each effect size
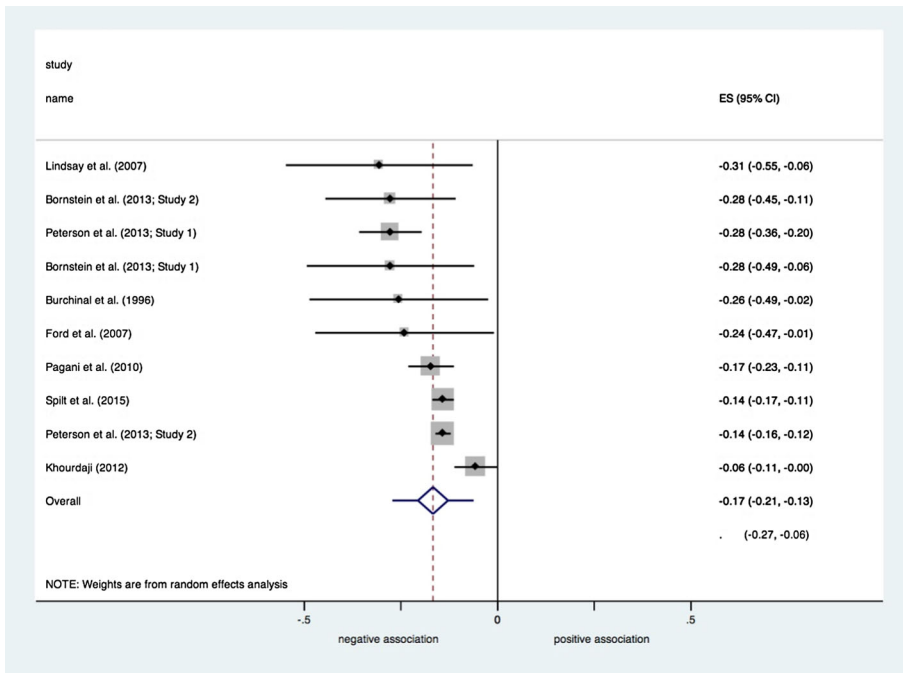
Fig. 5 Predictive associations between language and problem behavior. *Negative correlations* represent low language being associated with greater problem behavior(s)

size ($n = 10$). Results indicated that time between measurement ($\beta = -0.012$, $p = 0.51$, 95 % CI = [−0.05, 0.02]), and risk status ($\beta = 0.1$, $p = 0.148$, 95 % CI = [−0.04, 0.23]) did not significantly predict the effect size; that is, neither length of time between the measurement nor risk status significantly predicts the magnitude or the direction of the association between language ability and problem behavior.

Additional Analyses

To examine the potential for outliers or publication bias, we generated a funnel plot, conducted a trim and fill analysis (that yielded no trimmed or filled effect sizes), and employed the Egger's test ($p = 0.111$). Overall, the synthesis of these analyses provides convincing evidence that the reports included in this random-effects meta-analysis are free of publication bias.

Discussion

Summary of Evidence

This paper investigates the magnitude of the association between language ability and problem behavior in school-age children and youth. After a systematic literature review yielding 19 eligible reports for concurrent ($ES = 25$) and 8 for predictive ($ES = 10$) meta-analysis, we conducted random-effects meta-analyses to estimate the overall mean correlation between standardized measures of language and problem behavior. This

synthesis found significant negative effect sizes for both concurrent and predictive studies ($z = -0.17$). Illustratively, for a standard deviation lower on the Peabody Picture Vocabulary Test–III (Dunn and Dunn 1997), this effect size would represent an increase of 1.6 points on the Child Behavior Checklist (Achenbach 1991) problem behavior scale. These summative estimates corroborate the existing primary studies and literature reviews that report low language to be correlated with higher levels of problem behavior. These estimates are particularly convincing because the high end of both prediction intervals was still negative. This finding implies that any new study meeting inclusion criteria is very likely (95 %) to yield a negative correlation between language and problem behavior.

We also examined this relation across studies that include both at-risk and representative samples of children. Meta-regressions did not reveal significant differences between studies that included at-risk populations or for study sample mean age. This is potentially important because the association between language and problem behavior may not be unique to students with other noted deficits or risk factors and maybe stable over time. While the overall mean effect size is relatively small ($z = -0.17$), its significance coupled with both confidence and prediction intervals containing only negative values suggests a meaningful association.

In an exploratory fashion, we conducted additional analyses that estimated separate correlations for receptive and expressive language. These results suggest that each of these analyses is independently significant. However, the differences between the overall mean effect sizes for each construct are minimal. Further, the confidence interval around the mean effect size for expressive language is entirely subsumed by the receptive language mean effect size confidence interval. Thus, we are unable to make statements about the differences between each language and behavior construct.

Implications and Considerations

This paper extends the literature in several ways. We add to the existing systematic reviews by estimating a significant association between low language and problem behavior in children and youth with high-incidence disabilities, and representative and typically developing samples. These constructs were previously examined in samples of students with EBD (Hollo et al. 2014) and SLI (Yew and O'Kearney 2013). Our findings have implications for the classroom and for education more broadly.

The overall mean effect sizes for each analysis in this paper were statistically significant, and the confidence and prediction intervals in both meta-analyses further supported the findings. Considering the cumulative body of research that has described prevalence, co-occurrence, correlation, and prediction of language and behavior, researchers should consider allocating resources to focus on the identification of the mechanisms of change in this association. Then, probing these mechanisms in the development of intervention with the goal of improving social, behavioral, and educational outcomes for our children and youth.

The results of our moderator analyses indicated that both age and risk status were not significant predictors of the effect size. Interestingly, some of stronger negative effect sizes were from studies that examined representative samples of children. These findings suggest that the association of interest may not be unique to children and youth with co-occurring or identified deficits.

This study may also provide important implications for teacher preparation curriculum. While some countries set expectations for classrooms relative to language of instruction, approaching language variation in students, and comprehension (e.g., The National Curriculum in England (Department for Education 2014)), this is not universal. Behavior as a form of language and communication is widely known to be of importance in early childhood education as well as for students with severe disabilities. Our findings suggest that this recognition may be an important consideration for students with high-incidence disabilities as well for children without identified deficits. Because risk status did not predict effect size, risk characteristics may not preclude children from potentially important linguistic and behavioral associations. Further, because age did not predict effect size, teachers of children in upper elementary and secondary grades may be missing important pedagogical knowledge and strategies that involve considering language, language development, and communication.

Limitations

Although the exploratory findings of this meta-analysis support previous reviews and primary studies and extend these results to a wider range of school children, several factors limit the interpretation of the findings. First, although an effort was made to identify all eligible studies via a variety of search methods, it is still possible that some eligible studies were not captured by our search strategy. Second, only studies that used standardized, norm-referenced measures were included. Many studies were excluded due to measures not being norm-referenced, while meeting all other eligibility criteria. It is possible that the body of studies that used norm-referenced measures for both constructs is substantively different than those that only used one or none at all. For example, standardized comprehensive language assessments are expensive and time-consuming relative to other measures, and may measure language more globally than researcher-created measures. The small overall effect size may also be due to the sensitivity of the behavior measures (adult-rated). It is possible that error in these measures may deflate the effect size; that is, rating scales may not be sensitive to different topographies of classroom behavior that may be the most relevant. Further, only including studies that used normative measures of the constructs of interest may limit the picture of the association, and some studies included parent-rated behavior. It is likely that parent reports present differently than teacher reports (e.g., Hinshaw et al. 1992; Stone et al. 2010) due to the different contexts and relationships children have with these adults, and are a limitation of the current study. For example, parents report a more comprehensive perspective of their child's behavior, while teachers typically provide an academic-related rating and peer comparisons may influence overall ratings. Primary studies have attended to the differences between parent and teacher reports (e.g., Lindsay et al. 2007), and future meta-analytic reviews could aggregate these associations to advance the field. However, no peer- or self-report measures were included in this study.

Additionally, the measures used the primarily indexed externalizing problem behaviors. However, some individual studies included in this meta-analysis (e.g., SDQ total difficulties score) reported composite scores that included internalizing, emotional, and hyperactive symptoms. There is evidence that suggests several other types of problem behavior, such as inattention, hyperactivity, and internalizing behaviors, are also related to low language ability. It is likely that many of these behaviors exist in children, particularly in the studies that used representative samples and typically developing peers. It is possible that our estimate of externalizing behavior may have been influenced by these other scales within total composite

scores. Future research should conduct subgroup analyses and include these topographies of behavior to extend the literature.

Third, included studies seemed qualitatively heterogeneous, although moderator analyses were not significant for age or population. These results suggest that this relation is evident regardless of age and may also be a relevant consideration in schools for all children. However, it is important to note that a non-significant moderator effect for age or risk does not equate evidence of no effect; that is, we cannot conclude that age and risk status do not affect this relation. We simply may not have the statistical power to detect a significant relation.

Fourth, time points in longitudinal studies with repeated measures were selected randomly. This approach was to ensure that we did not systematically inflate (or deflate) the effect size of an individual study or skew the age distribution. However, it is possible that the effect sizes or ages at randomly selected time points were substantively different than the other time points in the study. Further, along the same logic, for we randomly selected studies that analyzed dependent data; that is, if a research group published follow-up papers over time, we randomly selected one for inclusion. This potentially limits our data when interpreting our age- and time-related moderator analyses. Future examinations should generate research-based hypotheses around age and design studies accordingly. For example, research could estimate and compare relations between language and behavior in young children to secondary students.

Fifth, and important in the context of this paper, the limitations of working with correlational data apply; that is, the individual study and overall mean effect sizes both represent bivariate correlations and do not represent modeled estimates that include covariates.

Finally, an important substantive limitation to this study is our inability to account for the association between behavior problems and socioeconomic disadvantage. Our theoretical model posits that language is an important fundamental proficiency that influences behavior and academic skill. It is possible that poor performance in literacy is compounded by low SES and, thus, precedes classroom entry. Studies have controlled for many student- and family-level characteristics, reporting that language still emerged a unique predictor of concurrent and later problem behavior (e.g., Bornstein et al. 2013; Peterson et al. 2013). However, it is important to consider this limitation because we are unable to tease out the role and the direction of SES in our model.

Future Directions

Federal funding agencies have recently stated that more research is needed to better understand the association between behavior problems and language deficits (Institute of Education Sciences 2016). While acknowledging its limitations, the present paper has several implications for future research aligned with this and other research initiatives. Our findings suggest that a meaningful relation exists across a wide range of students, but we cannot speak to subconstruct differences. Future studies should focus on the explicit relations between receptive language, expressive language, and different topographies of behavior to determine what types of behaviors are more or less related to each type of language construct. Studies should also compare the predictive value of norm-referenced measures to researcher-created, aligned measures of language and behavior.

Future studies should also consider whether externalizing rating scales tap into the same construct as the classroom behaviors teachers consider the most problematic. A measurement study might compare rating scales to direct observation of classroom behavior to determine if

rating scales are sufficient proxies of what is observed in the classroom. This measurement issue may be an important distinction to make, as observable problem behavior in the classroom is most likely an active ingredient in the cyclic reinforcement of negative student-teacher interactions (Myers and Pianta 2008; Wehby et al. 1995).

However, because the literature that points to a negative relation between low language and problem behavior is robust, the next step should be to determine how to successfully intervene. Two primary approaches to intervention are (1) intervention at the child level and (2) intervention at the contextual level. Intuitively, because language deficits predict later problem behavior, improving child language may help reduce problem behavior and support social and academic outcomes. Menting et al. (2011) found that children with higher receptive language exhibited decreases in problem behavior over time. A reduction of problem behavior may provide more opportunities for children to engage in academic instruction and also provide the skills to appropriately communicate her needs and reduce negative interactions. An explicit focus on improving child language, whether via pullout therapy, scaffolded classroom comprehension strategies, or providing corrective feedback, may be particularly important (and potentially missing) for some children with linguistic and behavioral deficits.

As discussed earlier, it is difficult to broadly consider behavioral and linguistic deficits without acknowledging the influential role of SES. Although some individual studies have controlled for participant-level SES, future studies related to these constructs should collect and present summary-level SES data, and future meta-analyses should include SES as a theorized moderator of effect size.

Another line of inquiry that may be effective in supporting these students is through contextual intervention. Specifically, researchers should consider teacher-level variables and how they impact student performance. If a mismatch between teacher instruction and student language promote negative teacher-student interactions, future research should identify ways to support teachers and their instruction. For example, exploring the effect of differential forms of instruction (e.g., feedback, visuals) on student outcomes in classroom settings may provide valuable information on how teachers can support students during instruction in a compensatory fashion. It is important that future research move from describing the problem to identifying malleable contextual factors that, through intervention, may help reduce problem behavior and increase the social, academic, and subsequent life outcomes for these children and youth.

Conclusions

Findings from this meta-analysis suggest a small but significant negative association between low language ability and problem behavior. At this summative stage, age and risk factor do not seem to be influencing factors of this association. However, these findings add to the literature by quantitatively synthesizing the magnitude of the association and broadening the scope to a wider population of schoolchildren. While considering study limitations, this extension of the literature corroborates previous literature and points to a continued, important focus on how language and behavior deficits are intertwined.

References

Achenbach, T. (1991). *Manual for the child behavior checklist/4–18 and 1991 Profile*. Burlington: University of Vermont Department of Psychiatry.

Aro, T., Eklund, K., Nurmi, J. E., & Poikkeus, A. M. (2012). Early language and behavioral regulation skills as predictors of social outcomes. *Journal of Speech, Language, and Hearing Research, 55*(2), 395–408. doi:10.1044/1092-4388(2011/10-0245).

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), 43–56. doi:10.1177/106342660201000105.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: Wiley.

Bornstein, M. H., Hahn, C. S., & Suwalsky, J. T. (2013). Language and internalizing and externalizing behavioral adjustment: developmental pathways from childhood to adolescence. *Development and Psychopathology, 25*(03), 857–878. doi:10.1017/s0954579413000217.

Botting, N., & Conti-Ramsden, G. (2008). The role of language, social cognition, and social skill in the functional social outcomes of young adolescents with and without a history of SLI. *British Journal of Developmental Psychology, 26*(2), 281–300. doi:10.1348/026151007x235891.

Bradley, R., Henderson, K., & Monfore, D. A. M. (2004). A national perspective on children with emotional disorders. *Behavioral Disorders*, 211–223

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*(1), 4–23. doi:10.1007/s10864-007.

Bretherton, L., Margot, P., Bauvin, E., Cini, E., Eadie, P., & Reilly, S. (2014). Developing relations in language and behavior in preschool children from the Early Language in Victoria Study: implications for intervention. *Emotional and Behavioural Difficulties, 19*(1), 7–27. doi:10.1080/13632752.2013.854956.

Briesch, A. M., Volpe, R. J., & Ferguson, T. D. (2014). The influence of student characteristics on the dependability of behavioral observation data. *School Psychology Quarterly, 29*(2), 171–181. doi:10.1037/spq0000042.

Burchinal, M., Roberts, J. E., Zeisel, S. A., Hennon, E. A., & Hooper, S. (2006). Social risk and protective child, parenting, and child care factors in early elementary school years. *Parenting: Science and Practice, 6*(1), 79–113.

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. doi:10.1044/0161-1461.1902.191.

Cantwell, D. P., & Baker, L. (1980). Academic failures in children with communication disorders. *Journal of the American Academy of Child Psychiatry, 19*(4), 579–591. doi:10.1016/s0002-7138(09)60963-8.

Cazden, C. B. (2001). *Classroom discourse: the language of teaching and learning*. Portsmouth: Heinemann.

Chow, J. C., & Hollo, A. H. (2014), *Language and Behavior Profiles of School Age Students with Emotional Disturbance*. Poster presented at the Council for Exceptional Children Annual Convention, Philadelphia, PA

Chow, J. C., & Wehby, J. H. (2015), *Language, Behavior, and Achievement: A Multivariate Profile Analysis*. Poster presented at the Pacific Coast Research Conference, Coronado, CA

Chow, J. C., & Jacobs, M. (2016). The role of language in fraction performance: A synthesis of literature. *Learning and Individual Differences, 47*, 252–257. doi:10.1016/j.lindif.2015.12.017.

Department for Education. (2014). *The national curriculum in England: framework document*. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-framework-for-key-stages-1-to-4.

Dickinson, D., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: why language is central for learning development. *Educational Researcher, 39*(4), 505–530. doi:10.3102/0013189x10370204.

Dockrell, J. E., Lindsay, G., & Connelly, V. (2009). The impact of specific language impairment on adolescents' written text. *Exceptional Children, 75*(4), 427–446. doi:10.1177/001440290907500403.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428. doi:10.1037/0012-1649.43.6.1428.

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. Circle Pines: American Guidance Service.

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. doi:10.1111/j.0006-341x.2000.00455.x.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629–634.

Ford, S., Farah, M. D., Shera, D. M., & Hurt, H. (2007). Neurocognitive correlates of problem behavior in environmentally at-risk adolescents. *Journal of Developmental and Behavioral Pediatrics, 28*(5), 376–385. doi:10.1097/dbp.0b013e31811430db.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology, 97*(3), 493. doi:10.1037/0022-0663.97.3.493.

Fujiki, M., Brinton, B., Morgan, M., & Hart, C. H. (1999). Withdrawn and social behavior of children with language impairment. *Language, Speech, and Hearing Services in Schools, 30*, 183–195. doi:10.1044/0161-1461.3002.183.

Gunter, P. L., Jack, S. L., Depaepe, P., Reed, T. M., & Harrison, J. (1994). Effects of challenging behaviors of students with EBD on teacher instructional behavior. *Preventing School Failure: Alternative Education for Children and Youth, 38*(3), 35–39.

Haak, J., Downer, J., & Reeve, R. (2012). Home literacy exposure and early language and literacy skills in children who struggle with behavior and attention problems. *Early Education & Development, 23*(5), 728–747. doi:10.1080/10409289.2011.565721.

Harmon, M. R., & Watson, M. J. (2012). *Beyond grammar: Language, power, and the classroom*. Mahwah, NJ: Lawrence Earlbaum Associates, Inc.

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*, 183–196.

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: causal relationships and underlying mechanisms. *Psychological Bulletin, 111*(1), 127–155. doi:10.1037//0033-2909.111.1.127.

Hinshaw, S. P., Han, S. S., Erhardt, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: correspondence among parent and teacher ratings and behavior observations. *Journal of Clinical Child Psychology, 21*(2), 143–150. doi:10.1207/s15374424jccp2102_6.

Hollo, A., & Chow, J. C. (2015). Supporting students with high-incidence disabilities: problem behavior as functional communication. *Beyond Behavior, 24*(3), 1–9.

Hollo, A., Wehby, J. H., & Oliver, R. M. (2014). Unidentified language deficits in children with emotional and behavioral disorders: a meta-analysis. *Exceptional Children, 80*(2), 169–186. doi:10.1177/001440291408000203.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004)

Institute of Education Sciences. (2016). Request for applications: special education research grants (CFDA Number: 84.324A). National Center for Special Education Research.

Kent, S., Wanzek, J., Petscher, Y., Al Otaiba, S., & Kim, Y. S. (2014). Writing fluency and quality in kindergarten and first grade: the role of attention, reading, transcription, and oral language. *Reading and Writing, 27*(7), 1163–1188. doi:10.1007/s11145-013-9480-1.

Kevan, F. (2003). Challenging behaviour and communication difficulties. *British Journal of Learning Disabilities, 31*(2), 75–80. doi:10.1046/j.1468-3156.2003.00226.x.

Lansing, A. E., Washburn, J. J., Abram, K. M., Thomas, U. C., Welty, L. J., & Teplin, L. A. (2014). Cognitive and academic functioning of juvenile detainees: implications for correctional populations and public health. *Journal of Correctional Health Care, 20*(1), 18–30. doi:10.1177/1078345813505450.

Lindsay, G., Dockrell, J. E., & Strand, S. (2007). Longitudinal patterns of behaviour problems in children with specific speech and language difficulties: child and contextual factors. *British Journal of Educational Psychology, 77*(4), 811–828. doi:10.1348/000709906x171127.

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

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

Menting, B., van Lier, P. A. C., & Koot, H. M. (2011). Language skills, peer rejection, and the development of externalizing behavior from kindergarten to fourth grade. *Journal of Child Psychology and Psychiatry, 52*(1), 72–79.

Myers, S. S., & Pianta, R. C. (2008). Developmental commentary: individual and contextual influences on student–teacher relationships and children's early problem behaviors. *Journal of Clinical Child & Adolescent Psychology, 37*(3), 600–608. doi:10.1080/15374410802148160.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research, 47*, 199–211. doi:10.1044/1092-4388(2004/017).

Nelson, N. W. (1985). Teacher talk and child listening: fostering a better match. Communication skills and classroom success: Assessment of language-learning disabled students, 65–104.

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*(2), 97–105. doi:10.1177/00224669050390020501.

Pagani, L. S., Fitzpatrick, C., Archambault, I., & Janosz, M. (2010). School readiness and later achievement: a French Canadian replication and extension. *Developmental Psychology, 46*(5), 984. doi:10.1037/a0018881.

Parker, F., & Riley, K. L. (2005). *Linguistics for non-linguists: a primer with exercises*. Boston: Pearson/Allyn and Bacon.

Patterson, G. R. (1982). Coercive family process. Eugene: Castalia.

Patterson, G. R., & Reid, J. B. (1984). Social interactional processes within the family: the study of the moment-by-moment family transactions in which human social development is imbedded. *Journal of Applied Developmental Psychology, 5*(3), 237–262. doi:10.1016/0193-3973(84)90021-2.

Peets, K. F. (2009). The effects of context on the classroom discourse skills of children with language impairment. *Language, Speech, and Hearing Services in Schools, 40*(1), 5–16. doi:10.1044/0161-1461(2008/07-0012).

Peterson, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A., Pettit, G. S., & Van Hulle, C. A. (2013). Language ability predicts the development of behavior problems in children. *Journal of Abnormal Psychology, 122*(2), 542–557. doi:10.1037/a0031963.

Qi, C. H., & Kaiser, A. P. (2004). Problem behaviors of low-income children with language delays: an observation study. *Journal of Speech, Language, and Hearing Research, 47*(3), 595–609. doi:10.1044/1092-4388(2004/046).

Redmond, S. M. (2011). Peer victimization among students with speech and language impairment, attention-deficit/hyperactivity disorder, at typical development. *Language, Speech, and Hearing Services in Schools, 42*, 520–535. doi:10.1044/0161-1461(2011/10-0078).

Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education, 38*(3), 130–143. doi:10.1177/00224669040380030101.

Scott, T. M., Alter, P. J., & Hirn, R. G. (2011). An examination of typical classroom context and instruction for students with and without behavioral disorders. *Education and Treatment of Children, 34*(4), 619–641. doi:10.1353/etc.2011.0039.

Spilt, J. L., Koomen, H. M. Y., & Harrison, L. J. (2015). Language development in early schools years: the importance of close relationships with teachers. *Developmental Psychology, 51*(2), 185–196. doi:10.1037/a0038540.

St Clair, M. C., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2011). A longitudinal study of behavioral, emotional and social difficulties in individuals with a history of specific language impairment (SLI). *Journal of Communication Disorders, 44*(2), 186–199. doi:10.1016/j.jcomdis.2010.09.004.

Stata. (2011). *Release 12. Statistical software*. College Station: StataCorp LP.

Stone, L. L., Otten, R., Engels, R. C., Vermulst, A. A., & Janssens, J. M. (2010). Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4-to 12-year-olds: a review. *Clinical Child and Family Psychology Review, 13*(3), 254–274. doi:10.1007/s10567-010-0071-2.

Sutherland, K. S., & Morgan, P. L. (2003). Implications of transactional processes in classrooms for students with emotional/behavioral disorders. *Preventing school failure: alternative education for children and youth, 48*(1), 32–37. doi:10.1080/1045988x.2003.10871077.

Titsworth, S., Mazer, J. P., Goodboy, A. K., Bolkan, S., & Myers, S. A. (2015). Two meta-analyses exploring the relationship between teacher clarity and student learning. *Communication Education, 64*(4), 385–418. doi:10.1080/03634523.2015.1041998.

Tomblin, B. T., 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. doi:10.1111/1469-7610.00632.

Warr-Leeper, G., Wright, N. A., & Mack, A. (1994). Language disabilities of antisocial boys in residential treatment. *Behavioral Disorders, 19*, 159–169.

Wehby, J. H., Symons, F. J., & Shores, R. E. (1995). A descriptive analysis of aggressive behavior in classrooms for children with emotional and behavioral disorders. *Behavioral Disorders, 20*, 87–105.

Wehby, J. H., Lane, K. L., & Falk, K. B. (2003). Academic instruction for students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 11*(4), 194–197. doi:10.1177/10634266030110040101.

Yew, S. G. K., & O'Kearney, R. (2013). Emotional and behavioural outcomes in childhood and adolescence with specific language impairments: meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry, 54*(5), 516–524. doi:10.1111/jcpp.12009.

Yew, S. G. K., & O'Kearney, R. (2015). The role of early language difficulties in the trajectories of conduct problems across childhood. *Journal of Abnormal Child Psychology, 43*, 1515–1527. doi:10.1007/s10802-015-0040-9.

Yoshimasu, K., Barbaresi, W. J., Colligan, R. C., Killian, J. M., Voigt, R. G., Weaver, A. L., & Katusic, S. K. (2011). Written-language disorder among children with and without ADHD in a population-based birth cohort. *Pediatrics, 128*(3), e605–e612. doi:10.1542/peds.2010-2581.