Language Disorders and Problem Behaviors: A Meta-analysis

Philip R. Curtis, MA, Jennifer R. Frey, PhD, Cristina D. Watson, Ed.M, Lauren H. Hampton, PhD, Megan Y. Roberts, PhD

CONTEXT: A large number of studies have shown a relationship between language disorders and problem behaviors; however, methodological differences have made it difficult to draw conclusions from this literature.

OBJECTIVE: To determine the overall impact of language disorders on problem behaviors in children and adolescents between the ages of birth and 18 years and to investigate the role of informant type, age, and type of problem behavior on this relationship.

DATA SOURCES: We searched PubMed, EBSCO, and ProQuest.

STUDY SELECTION: Studies were included when a group of children with language disorders was compared with a group of typically developing children by using at least 1 measure of problem behavior.

DATA EXTRACTION: Effect sizes were derived from all included measures of problem behaviors from each study.

RESULTS: We included 47 articles (63,153 participants). Meta-analysis of these studies revealed a difference in ratings of problem behaviors between children with language disorders and typically developing children of moderate size (g = 0.43; 95% confidence interval 0.34 to 0.53; P < .001). Age was entered as a moderator variable, and results showed that the difference in problem behavior ratings increases with child age (increase in g for each additional year in age = 0.06; 95% confidence interval 0.02 to 0.11; P = .004).

LIMITATIONS: There was considerable heterogeneity in the measures of problem behaviors used across studies.

CONCLUSIONS: Children with language disorders display greater rates of problem behaviors compared with their typically developing peers, and this difference is more pronounced in older children.
Between 13.4% and 19.1% of toddlers experience delayed language development,\(^1\) and between 6% and 8% of kindergartners have a developmental language disorder.\(^2\) These groups of children are defined as having delayed or disordered language development but intact nonverbal cognitive abilities, in the absence of other known genetic or neurodevelopmental disorders.\(^3\) Disorders in language development have been associated with a number of difficulties in academic and psychosocial development, including increased rates of problem behaviors.\(^4\)–\(^6\) Although many studies have revealed the association between language disorders and problem behaviors across development, there is considerable methodological heterogeneity between studies. This heterogeneity reflects differences in how language skills are assessed and the criteria used for diagnosis of language delay or disorder, the informant type used to measure problem behaviors (ie, parents, teachers, or researcher-coded observations of child behaviors), the age of children included in the study, as well as the types of problem behaviors that were assessed.

Measures of problem behaviors are often used to classify symptoms as either internalizing behaviors or externalizing behaviors.\(^7\) Internalizing behaviors include symptoms commonly associated with depression and anxiety, whereas externalizing behaviors include disruptive, hyperactive, and aggressive behaviors.\(^8\) Although this is only 1 system of classification, a majority of the behavioral and emotional assessments used in the existing literature investigating the relation between language disorders and problem behaviors use scales that reflect these dimensions (eg, the Child Behavior Checklist [CBCL],\(^9\) the Infant Toddler Social Emotional Assessment [ITSEA],\(^10\) the Social Competence and Behavior Evaluation,\(^11\) etc), so this classification system was used in the current study.

To quantitatively assess the associations between language delays and problem behaviors found in the literature, while taking into account the issues noted above, we conducted a meta-analysis used to address the following 3 questions:

1. Do children with language disorders display higher rates of problem behaviors compared with their typically developing peers?
2. Does informant type and/or age moderate the relation between language disorder status and problem behaviors?
3. Is language disorder status more strongly associated with either internalizing behaviors or externalizing behaviors?

### METHODS

#### Identification of Studies

Searches of PubMed, EBSCO, and ProQuest were performed for all dates until July 2017. The following search terms were used, restricted to the titles and/or abstracts within each database: “disruptive behavior*,” “behavior problems,” “problem behavior*,” “challenging behavior*,” “externalizing behavior*,” “internalizing behavior*,” “agress* behavior*,” or “behave*,” and “communication,” “language,” “vocabulary,” “semantics,” “syntax,” or “grammar” and “delay,” “disorder,” “impairment,” “disability,” or “late talkers.” In total, this search yielded 3128 unique abstracts. Additionally, reference lists of included studies were extracted from each article. During the data extraction process, an additional 29 abstracts were excluded from analyses. Reasons for exclusion of these articles are available in Supplemental Table 5.

#### Data Extraction

After digital or hard copies of each included study were obtained, data were extracted from each article by using a detailed coding protocol (this protocol can be obtained by contacting the first author). To test for bias within studies, a “quality of language assessment” variable was created to rate the rigor of the diagnostic methods used to classify children as typically developing or language disordered in each study. A 5-point scale was developed, and a code was assigned to every article (see Table 1 for a full explanation of this code). All articles were double-coded by 2 independent reviewers, and discrepancies were resolved through consensus.

Included in many studies were separate language-disordered groups. For instance, authors of
<table>
<thead>
<tr>
<th>Source, y</th>
<th>Control Group, N</th>
<th>Language-Delayed Group, N</th>
<th>Mean Age of Participants</th>
<th>Minimum Age</th>
<th>Maximum Age</th>
<th>Quality of Language Assessment</th>
<th>Race of the Sample</th>
<th>SES of the Sample</th>
<th>Excluded Children With ASD</th>
<th>Excluded Children With IQ &lt;70</th>
<th>Informant</th>
<th>Behavior Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asbell et al 1989</td>
<td>16</td>
<td>97</td>
<td>5.5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>NR</td>
<td>NR</td>
<td>NE</td>
<td>NE</td>
<td>Parent</td>
<td>CBCL 4–18</td>
</tr>
<tr>
<td>Beitchman et al 1989</td>
<td>58</td>
<td>82</td>
<td>NR</td>
<td>NR</td>
<td>NE</td>
<td>NE</td>
<td>Parent</td>
<td>CBCL 2–3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bretherton et al 2014</td>
<td>53</td>
<td>11</td>
<td>2.14</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>NR</td>
<td>Midhigh</td>
<td>NE</td>
<td>NE</td>
<td>Parent</td>
<td>CBCL 15–5</td>
</tr>
<tr>
<td>Carson et al 2007</td>
<td>80</td>
<td>93</td>
<td>2.49</td>
<td>1.19</td>
<td>3.52</td>
<td>1</td>
<td>W</td>
<td>Mixed</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent</td>
<td>Conners’ 10-Item Test</td>
</tr>
<tr>
<td>Carson et al 2002</td>
<td>27</td>
<td>25</td>
<td>6.38</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>NR</td>
<td>NR</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent</td>
<td>Conners’ Rating Scale–Revised</td>
</tr>
<tr>
<td>Fuji et al 1996</td>
<td>19</td>
<td>19</td>
<td>10.21</td>
<td>8</td>
<td>12</td>
<td>2</td>
<td>NR</td>
<td>NE</td>
<td>Excluded</td>
<td>Teacher</td>
<td>Social Skills Rating System</td>
<td></td>
</tr>
<tr>
<td>Fuji et al 2002</td>
<td>41</td>
<td>41</td>
<td>9.06</td>
<td>NR</td>
<td>NR</td>
<td>1</td>
<td>NR</td>
<td>NE</td>
<td>Excluded</td>
<td>Teacher</td>
<td>Emotion Regulation Checklist</td>
<td></td>
</tr>
<tr>
<td>Fuji et al 2004</td>
<td>43</td>
<td>43</td>
<td>8.84</td>
<td>5</td>
<td>12</td>
<td>2</td>
<td>NR</td>
<td>NE</td>
<td>Excluded</td>
<td>Teacher</td>
<td>Teacher Behavior Rating Scale</td>
<td></td>
</tr>
<tr>
<td>Goudsmit et al 1996</td>
<td>42</td>
<td>30</td>
<td>4.78</td>
<td>4.25</td>
<td>5.5</td>
<td>1</td>
<td>W</td>
<td>Mixed</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent</td>
<td>Teacher Behavior Rating Scale</td>
</tr>
<tr>
<td>Henrichs et al 2012</td>
<td>4169</td>
<td>687</td>
<td>1.55</td>
<td>NR</td>
<td>NR</td>
<td>1</td>
<td>Mixed</td>
<td>NE</td>
<td>NE</td>
<td>Parent</td>
<td>Behavior Problem Checklist (Quay)</td>
<td></td>
</tr>
<tr>
<td>Holme et al 1993</td>
<td>16</td>
<td>14</td>
<td>4.4</td>
<td>NR</td>
<td>NR</td>
<td>1</td>
<td>H</td>
<td>NR</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent</td>
<td>Behavior Problem Checklist (Quay)</td>
</tr>
<tr>
<td>Lemanek et al 1993</td>
<td>1047</td>
<td>191</td>
<td>4.08</td>
<td>NR</td>
<td>NR</td>
<td>2</td>
<td>NR</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent</td>
<td>Behavior Problem Checklist (Quay)</td>
</tr>
<tr>
<td>Lindholm et al 1979</td>
<td>2991</td>
<td>106</td>
<td>NR</td>
<td>NR</td>
<td>4</td>
<td>W</td>
<td>Mixed</td>
<td>NE</td>
<td>NE</td>
<td>Parent</td>
<td>Behavior Problem Checklist (Quay)</td>
<td></td>
</tr>
<tr>
<td>Malay 1995</td>
<td>15</td>
<td>16</td>
<td>4.2</td>
<td>3</td>
<td>5.5</td>
<td>1</td>
<td>NR</td>
<td>Mixed</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent</td>
<td>CBCL 2–3</td>
</tr>
<tr>
<td>McCabe 2005</td>
<td>22</td>
<td>116</td>
<td>4.32</td>
<td>3.25</td>
<td>5.25</td>
<td>2</td>
<td>NR</td>
<td>NR</td>
<td>NE</td>
<td>NE</td>
<td>Parent</td>
<td>Malay’s Observational Coding System</td>
</tr>
<tr>
<td>McCabe et al 2006</td>
<td>18</td>
<td>30</td>
<td>4.2</td>
<td>3.17</td>
<td>5.42</td>
<td>4</td>
<td>W</td>
<td>Mixed</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent</td>
<td>Parent-Child Rating Scale 3.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source, y</th>
<th>Control Group, N</th>
<th>Delayed Group, N</th>
<th>Minimum Age</th>
<th>Maximum Age</th>
<th>Quality of Language Assessmenta</th>
<th>Race of the Sampleb</th>
<th>SES of the Samplec</th>
<th>Excluded Children With ASD</th>
<th>Informant</th>
<th>Behavior Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCabe et al.2004</td>
<td>32261</td>
<td>1417</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Okuma et al.2015</td>
<td>14</td>
<td>14</td>
<td>9.59</td>
<td>11.25</td>
<td>1</td>
<td>NE</td>
<td>NR</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Nes et al.2015</td>
<td>32261</td>
<td>1417</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Q and Kaje et al.2004</td>
<td>2791</td>
<td>1307</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Paul et al.1990</td>
<td>33</td>
<td>33</td>
<td>2.5</td>
<td>3</td>
<td>3</td>
<td>NR</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Prior et al.2011</td>
<td>1179</td>
<td>651</td>
<td>4.14</td>
<td>6.25</td>
<td>4.5</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Raffa2004</td>
<td>20</td>
<td>17</td>
<td>5.98</td>
<td>7.98</td>
<td>4</td>
<td>NE</td>
<td>NR</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Redmond and Rice2004</td>
<td>822</td>
<td>467</td>
<td>2.55</td>
<td>4.25</td>
<td>3.5</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Roberts et al.2011</td>
<td>20</td>
<td>17</td>
<td>5.98</td>
<td>7.98</td>
<td>4</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Rotter2004</td>
<td>13</td>
<td>15</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Redmond and Rice2004</td>
<td>822</td>
<td>467</td>
<td>2.55</td>
<td>4.25</td>
<td>3.5</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Raffa2004</td>
<td>20</td>
<td>17</td>
<td>5.98</td>
<td>7.98</td>
<td>4</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>S. N. Chapman et al.2007</td>
<td>53</td>
<td>45</td>
<td>4.54</td>
<td>6.54</td>
<td>5</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Talpe et al.1989</td>
<td>49</td>
<td>81</td>
<td>4.25</td>
<td>8.25</td>
<td>8</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Tomblin et al.2000</td>
<td>417</td>
<td>164</td>
<td>7.98</td>
<td>11.98</td>
<td>11.98</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Whitehouse et al.2011</td>
<td>1290</td>
<td>43</td>
<td>2.17</td>
<td>4.17</td>
<td>4.17</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Willinger et al.2005</td>
<td>94</td>
<td>94</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
<tr>
<td>Zubrick2004</td>
<td>412</td>
<td>412</td>
<td>5.5</td>
<td>7.5</td>
<td>4</td>
<td>NE</td>
<td>NE</td>
<td>Excluded</td>
<td>NE</td>
<td>Parent Social Skills Rating System – Preschool Teacher</td>
</tr>
</tbody>
</table>
In the studies that were included in this meta-analysis, authors used a number of different measures of problem behaviors, including published standardized measures, researcher-created interviews or questionnaires, and coding of direct observations of children’s behaviors by researchers. Questionnaires and interviews were completed by parents, teachers, or both. Some studies divided children into receptive-expressive, expressive-only, and articulation-disordered groups. Because authors of studies varied considerably in how they defined language disorder subgroups, and no consistently defined groups could be extracted across studies, for the purposes of the current analyses, all language-disordered groups were combined to form a single language-disordered group for each study. Groups comprising only participants with articulation disorder, when reported separately, were excluded. Additionally, groups comprising only participants with pragmatic language impairment were excluded. Pragmatic language impairment, also referred to as social (pragmatic) communication disorder in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, is characterized by difficulties in the social use of language that is not better explained by deficits in grammar or word structure. Although these difficulties with pragmatic language may be associated with problem behaviors, the underlying mechanisms of that association may be different than the mechanisms linking deficits in language content and structure to problem behaviors. For this reason, we feel that articles in which authors investigate pragmatic language difficulties specifically warrant a separate study and so have been excluded in the current analyses.

Behavioral Measure Characteristics

In the studies that were included in this meta-analysis, authors used a number of different measures of problem behaviors, including published standardized measures, researcher-created interviews or questionnaires, and coding of direct observations of children’s behaviors by researchers. Questionnaires and interviews were completed by parents, teachers, or both. One complexity in measuring problem behaviors arises from the factor structures used when creating measures. Many measures, such as the CBCL, group items into lower-order “narrow-band” factors (ie, “aggression,” “anxious and/or depressed,” etc), as well as higher-order factors, typically labeled as “internalizing problems,” “externalizing problems,” or “total problem behavior” factor, in which all behaviors are combined. When measures that used such factor structures were included in studies, there was a great deal of heterogeneity in what scores authors reported. Authors of some studies reported only higher-order factors, such as “internalizing composite” or total problem behaviors, whereas other authors reported only subscales. In the current analyses, we were interested in the following 2 broad domains: total problem behaviors and a comparison of internalizing and externalizing problems. For this reason, all reported effect sizes were captured. When it was known that an author had not reported a certain scale (for instance, he or she reported the CBCL internalizing composite but not the externalizing composite), attempts were made to contact the author and obtain these data. Out of 7 data requests sent, 3 authors were able to provide us with missing data.

Data Synthesis

Study authors reported the included effect sizes as differences in means, percentages of each group meeting a “clinical cutoff” for problem behaviors, reported test results, and 2-group analyses of variance. Effect sizes and effect size variances were computed in the Comprehensive Meta-Analysis (version 3.3.070) software. Because some studies had small sample sizes, all effect sizes were converted to Hedges’ g, which is used to correct for small sample size.
In some studies, measures were available for only subsamples of the study participants, either because of measurement issues (e.g., Malay 1995) or because of missing data. When data were available, participant ages and sex ratios were calculated for each measure individually; when these data were not available, the overall ages and participant sexes for the whole study were used.

Effect sizes were classified as representing either internalizing or externalizing behaviors. Measures that were not classified by the measure itself (e.g., “CBCL 1.5–5 Internalizing Composite”), the authors of the study or other publications were independently classified by 2 of the authors with 91.8% agreement. Discrepancies were resolved by consensus.

Data Analysis

Conventional meta-analytic methods require that each study is used to contribute only 1 independent effect size. Because many of the study authors included in this meta-analysis reported >1 effect size that need to be included in the same analysis, these traditional meta-analytic methods are not appropriate for the current study. When multiple effect sizes are derived from the same participants, these effect sizes are not independent but are instead correlated. It is possible to create synthetic effect sizes for each study by averaging effect sizes from the same study; however, the synthetic effect size’s SEs are dependent on the covariance structure between the individual effect sizes from which they are computed, making this approach problematic.

To more accurately model these multiple, dependent effect sizes across studies, we employed the robust variance estimation method created by Hedges et al. This novel method of meta-analysis does not require the explicit covariance structure between effect sizes reported from the same study (which are rarely available) but instead uses the observed residuals to estimate the meta-regression coefficient estimates. A correction for small sample sizes was employed in the current analyses. These analyses allowed us to include multiple effect sizes from the same study (e.g., Malay), avoiding both the problems of excluding valid estimates of problem behaviors as well as biasing our effect size estimates.

Moderator Analyses

In addition to these strengths, robust variance estimation also allows researchers to include additional variables as a means of modeling observed heterogeneity across effect sizes, what is frequently called a moderator analysis. These analyses function much like typical linear regression analyses, with the study-derived effect sizes as the dependent variables and study-level covariates, such as average age of participants or informant type, as the independent moderator variables. Full details are given in Hedges et al. The method of ordinary least squares is used to solve the linear equation predicting individual effect sizes, modeled with an intercept (the average effect size across studies and measures) and any moderators the researcher chooses to include. Each regression coefficient within the meta-regression can be interpreted as in a typical linear regression (for a 1-unit increase in the moderator variable, what is the expected change in the observed effect size?). SEs, significance levels, and confidence intervals (CIs) are provided for each parameter estimate to aid in interpretation. Statistically significant moderator variables suggest that the differences in effect sizes across studies are associated with differences in that particular moderator variable in the meta-regression. It is important to note that moderator variables entered into these meta-regressions are used to predict the effect sizes from each study. That is, moderators such as age, informant type, or type of problem behavior are used to predict the standardized difference in problem behavior scores between children with typical development and children with language delays or disorders.

RESULTS

The first set of analyses were used to deal with total problem behaviors, the most broad and inclusive category of problem behaviors. These scores are derived by pooling all problem behaviors assessed within a given measure. However, some study authors failed to report a composite score for the total problem behaviors. For instance, Carson et al reported an internalizing composite score and externalizing composite score for the CBCL 2 to 3 but not a total problem behavior score. To ensure that all studies contributed at least 1 effect size for this analysis, preference was given in the following order: (1) total problem behavior composite scores were reported; (2) if a total problem behavior composite score was not reported, an internalizing and/or externalizing composite score was reported; and (3) if no composite scores were reported, individual subscale scores were reported. No overlapping effect sizes were included (i.e., if a total problem behavior composite score was reported, externalizing and internalizing composite scores were not also included, because these scales draw from the same items as total problem behaviors scores). This system was used to ensure that studies in which authors did not report total problem behavior composite scores were still included in these analyses.
Research Question 1: What Is the Difference in Rates of Problem Behaviors Between Children With Language Delays and Their Typically Developing Peers?

To address this question, we created an intercept-only model. Results are reported in Table 2, and a forest plot is available in the Supplemental Information. For this model, there were 47 studies included with a total of 128 effect sizes (minimum = 1; mean = 2.7; maximum = 18), for $\tau^2 = 0.05$. The intercept was significant (0.43; 95% CI 0.34 to 0.53; $P < .001$), indicating that, on the whole, children with language delays have problem behavior ratings 0.43 SDs higher than their typically developing peers. See Fig 1 for a forest plot of effect sizes included in this analysis.

Research Question 2a: Do Effect Sizes Differ on the Basis of Informant?

It is possible that ratings of problem behaviors may vary across settings (ie, home, school, or research laboratories) or that different informants may rate children’s problem behaviors differently. Estimates of effect sizes are given in Table 2 for each type of informant individually. Average effect sizes from teacher report were higher than both those derived from parent report, as well as from researcher observational coding (0.63 versus 0.37 and 0.43, respectively). To test whether these differences were statistically significant, a moderator analysis was run by using a “teacher report” dummy code. This variable was coded as 0 for parent or researcher observations and 1 for teacher reports. Because there were comparatively few effect sizes derived from researcher observation (5 studies, 14 effect sizes), and the effect sizes derived from parent reports and researcher observations were similar, no variable was entered to differentiate between parent and researcher observations. Results from this model are given in Table 2.

The intercept, representing the average standardized difference in problem behaviors between children with typical development and children with language delays or disorders, remained significant, indicating that children with language delays are rated by their parents and researcher observations as having significantly more problem behaviors than typically developing children. The unstandardized coefficient of the dummy code for teacher reports was statistically significant, indicating that, within the studies included in this meta-analysis, on average, teachers identified a larger difference between groups than do parents or research observations.

Research Question 2b: Does the Association Between Language Disorders and Problem Behaviors Vary on the Basis of Children’s Age?

To test whether the relation between language disorders and problem behaviors varies by children’s age, an additional analysis was run with the average child age from each study entered as a moderator variable. Again, the dependent variable in these models is individual effect sizes, representing the standardized difference between children with language delays or disorders and children with typical language development. The age variable was centered at the age of the youngest participant (Henrichs et al mean age = 1.5 years), so that the intercept would represent the average effect size for children 1.5 years of age, and the unstandardized regression coefficient on mean age would represent the increase in effect size predicted by a 1-year increase in children’s average age. Results from this model are given in Table 3.

Results revealed that even for children as young as 1.5 years of age, language disorder status was associated with higher rates of problem behaviors (unstandardized coefficient = 0.19; 95% CI 0.07 to 0.31; $P = .004$). The unstandardized coefficient for the mean age variable was also statistically significant (0.07; 95% CI 0.03 to 0.11; $P = .001$), meaning that the association between language disorder status and problem behaviors is larger in older children than in younger children.

It could be argued that age and number of effect sizes derived by teacher report may in fact be collinear with one another, confounding the relation between age and problem behaviors and between informant type and problem behaviors. Indeed, within the current sample of studies, the average age for teacher-reported outcomes was significantly older than the average age of parent-reported outcomes (mean parent or observer-rated reported age = 5.02 years; mean teacher reported age = 6.61; t(34.18) = −2.88; $P = .007$). When both mean age and the teacher report dummy code were included in the same model, the unstandardized coefficient for the dummy variable for teacher report no longer approached significance (0.19; 95% CI −0.12 to 0.49; $P = .21$). More importantly, the estimate of the difference between parent-reported or observer-rated effect sizes and teacher-reported effect sizes dropped from 0.35 to 0.19, after controlling for mean age. This suggests that the overall higher ratings of problem behaviors by teachers within this sample are strongly related to children’s age. Within this model, the unstandardized coefficient of mean age was again significant (0.06; 95% CI 0.01 to 0.10; $P = .01$), indicating that after controlling for informant type, each additional year in age was associated with a 0.06 SD increase in the difference in problem behavior scores between children with language disorders and their typically developing peers.
TABLE 2 Average Standardized Differences Between Typically Developing Children and Children With Language Delays or Disorders by Informant Type

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimatea (SE)</th>
<th>P</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All informants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.43 (0.05)</td>
<td>&lt;.001</td>
<td>0.34 to 0.53</td>
</tr>
<tr>
<td>Studies: 47</td>
<td>P² = 78.1</td>
<td>t² = 0.05</td>
<td></td>
</tr>
<tr>
<td>Parent only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.37 (0.04)</td>
<td>&lt;.001</td>
<td>0.29 to 0.48</td>
</tr>
<tr>
<td>Studies: 40</td>
<td>P² = 75.4</td>
<td>t² = 0.04</td>
<td></td>
</tr>
<tr>
<td>Observation coding only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.43 (0.13)</td>
<td>.03</td>
<td>0.05 to 0.80</td>
</tr>
<tr>
<td>Studies: 5</td>
<td>P² = 37.3</td>
<td>t² = 0.07</td>
<td></td>
</tr>
<tr>
<td>Teacher only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.63 (0.15)</td>
<td>.001</td>
<td>0.30 to 0.96</td>
</tr>
<tr>
<td>Studies: 14</td>
<td>P² = 81.9</td>
<td>t² = 0.25</td>
<td></td>
</tr>
<tr>
<td>All informants, controlling for teacher report</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.38 (0.05)</td>
<td>&lt;.001</td>
<td>0.28 to 0.47</td>
</tr>
<tr>
<td>Teacher report</td>
<td>0.53 (0.13)</td>
<td>.02</td>
<td>(0.06 to 0.65)</td>
</tr>
<tr>
<td>Studies: 47</td>
<td>P² = 76.24</td>
<td>t² = 0.04</td>
<td></td>
</tr>
</tbody>
</table>

In all analyses, p = 0.8.

*a Estimates are unstandardized regression coefficients.

Research Question 3: Is Language Disorder Status More Strongly Associated With Either Internalizing Behaviors or Externalizing Behaviors?

Several researchers have suggested that language more strongly impacts 1 type of behavior (internalizing versus externalizing) compared with the other.44,64 To test this possibility within the current sample of studies, a “contrast variable” was created and scored as −0.5 for internalizing and +0.5 for externalizing scales. This type of coding results in the intercept signifying the overall effect size for all scales, whereas the contrast variable represents the average difference between internalizing and externalizing effect sizes. Parameter estimates for this model are given in Table 4. For this model, there were 40 studies included with a total of 122 effect sizes (minimum = 1; mean = 3.05; maximum = 14), for t² = 0.05. When predicting the standardized mean difference between children with language delays or disorders and children with typical language development, the intercept remained significant, indicating that children with language disorders display significantly more problem behaviors than do their typical peers. The behavior type variable, used to differentiate between internalizing and externalizing effect sizes, was not significant (P = .50), indicating that language disorders are not significantly more associated with either internalizing or externalizing behaviors.

To test whether there may be a differential impact of language disorders on internalizing and externalizing behaviors that varies on the basis of the child’s age, we created an interaction term between the following 2 moderator variables: mean age (again centered at 1.5 years) and behavior type. These variables and the interaction term were used to predict individual effect sizes, or the standardized mean difference between children with language disorders or delays, and children with typical language development. This interaction term was not significant (unstandardized coefficient = 0.003; 95% CI = −0.07 to 0.07; P = .92), suggesting that, across development, the difference in rates of problem behaviors between children with language disorders or delays and children with typical language development does not differ between internalizing behaviors and externalizing behaviors, regardless of children’s ages.

Sensitivity Analysis and Publication Bias

To test the extent to which the quality of the language assessment might influence the effect sizes derived from each study, we performed a sensitivity analysis. A “potentially low-quality language assessment” binary indicator was created for studies receiving a rating of 3, 4, or 5 on our language assessment quality rating (see Table 1). When this indicator was entered into a meta-regression for total problem behaviors, the resulting coefficient was nonsignificant (−0.12; P = .22). These results reveal that the quality of language assessment did not significantly impact the results of our analyses. Additionally, a “leave-1-out” analysis was performed by systematically running the total problem behaviors analysis, leaving 1 study out each time to assess each study’s individual impact on the results. The exclusion of any 1 study did not significantly impact the results (minimum average effect size = 0.41; maximum average effect size = 0.45).

To test for the presence of publication bias, or the selective publication of only significant effect sizes, a publication bias analysis was performed in R by using the metafor package.65 This analysis was done separately for averaged total problem, internalizing, and externalizing measures from each study. For externalizing measures, there was no evidence of publication bias according to Egger’s linear regression test (z score = 1.32; P = .19). Conversely, there was evidence of publication bias for total problem behaviors and for internalizing effect sizes. Egger’s linear regression test revealed significant asymmetry in the total problem behavior funnel plot (z score = 2.62; P < .01) and the internalizing funnel plot.
Studies
Belitchman et al., 1989
CBCL, 4–18: total problem behaviors
Belitchman et al., 2014
Strengths and Difficulties Questionnaire: total difficulties
Carson et al., 1998
CBCL, 2–3: externalizing composite
CBCL, 2–3: internalizing composite
Carson et al., 2007
Temperament and Atypical Behavior Scale: dysregulated
Temperament and Atypical Behavior Scale: hyperactive
Cauflfeld et al., 1989
Cauflfeld observation of parent and child: negative behavior
Cauflfeld observation of parent and child: refusal
Cauflfeld Structured Parental Interview: difficulty separating from parents
Cauflfeld Structured Parental Interview: shy or fearful in new situations
Eyberg Child Behavior Inventory: composite
Curtis et al., 2017
CBCL, 1.5–5: total problem behaviors
ITSEA: externalizing composite
ITSEA: Internalizing composite
Fujiki et al., 1996
Social Skills Rating System (Teacher) (Elementary): problem behaviors
Fujiki et al., 1999
Teacher Behavioral Rating Scale: withdrawn (relicence)
Fujiki et al., 2001
Fujiki, M. (Observational Coding of Playground Interaction): withdrawal
Fujiki et al., 2002
Emotion Regulation Checklist: liability or negativity
Fujiki et al., 2004
Emotion Regulation Checklist: liability or negativity
Teacher Behavioral Rating Scale: withdrawn (relicence)
Goudsmit, n.d.
CBCL, 4–18: externalizing composite
CBCL, 4–18: internalizing composite
Guralnick et al., 1996
CBCL, 4–18: total problem behaviors
Harrich et al., 2012
CBCL, 1.5–5: externalizing composite
CBCL, 1.5–5: internalizing composite
Herzel, n.d.
CBCL, 4–18: externalizing composite
Holmes, n.d.
CBCL, 4–18: total problem behaviors
Horowitz et al., 2003
ITSEA: externalizing domain
ITSEA: Internalizing domain
Lemanek et al., 1993
Lemanek Observational Coding: ignore
Lemanek Observational Coding: leave
Malay, 1995
CBCL, 2–3: aggressive behavior
CBCL, 2–3: anxious and/or depressed
CBCL, 2–3: destructive behavior
CBCL, 2–3: sleep problems
CBCL, 2–3: somatic complaints
CBCL, 2–3: withdrawn
CBCL, 4–18: aggressive behavior
CBCL, 4–18: attention problems
CBCL, 4–18: delinquent behavior
CBCL, 4–18: social problems
CBCL, 4–18: somatic complaints
CBCL, 4–18: thought problems
CBCL, 4–18: withdrawn
Eyberg Child Behavior Inventory: composite
Malay Observational Coding: covert inattention
Malay Observational Coding: noncompliance
Malay Observational Coding: overt inattention
Malay Observational Coding: restlessness

FIGURE 1
Forest plot of total problem behaviors.

(z score = 2.97; P < .01) (funnel plots available in Supplemental Information), suggesting the presence of publication bias. A “trim-and-fill” analysis was conducted, wherein additional artificial effect sizes (in this case representing small or negative effect sizes) are added to balance the funnel plot. The resulting modified random effect size estimate continued to be significant for both types of analyses (total problem behaviors: unstandardized coefficient = 0.41; P < .001; internalizing behaviors: unstandardized coefficient = 0.27; P < .001), suggesting that even after accounting for publication bias, children diagnosed with language disorders have higher rates of total problem behaviors and internalizing behaviors compared with their typically developing peers.

Finally, a sensitivity analysis for the value of ρ, or the assumed within-study correlation value, was conducted as specified by Hedges et al.59. This value was systematically varied from 0 (no between-measure correlation) to 1 (perfect between-measure correlation), with little change to any parameter estimates, strengthening our confidence in the results of these analyses.

DISCUSSION
The results of this meta-analysis revealed that children with language disorders display greater rates of problem behaviors as compared with their typically developing peers. More nuanced patterns of associations also emerged, such that there is a greater association between language disorder status and problem behaviors in older children than in younger children, and that, although teachers’ ratings of problem behaviors were higher than parents’ or research observations overall, this difference was no longer significant once children’s age was accounted for. Furthermore, there was no difference between the associations of language disorders with externalizing as compared with internalizing behaviors.
There are important considerations to make when investigating associations between language and problem behaviors in cross-sectional studies, as was done in this meta-analysis. One is that early-identified language disorders may in fact reflect only mild language delays. Consequently, the increasing effect size over time that we observed in this sample of studies may reflect diagnostic inaccuracy in identifying children with language delays early in life. Another consideration is the potentially compounding influence of other life domains that are impacted by language disorders. Developmental language disorder has been associated with poor peer relationships, increased bullying by other children, and poor academic skills. These problems in turn have been linked with both internalizing and externalizing problems. The greater association between language and problem behaviors in older children observed in this sample of studies may be mediated through the detrimental effect that language disorders have on other areas of development. This possibility is especially important when thinking about intervention approaches for children with co-occurring language disorders and problem behaviors; it may be that addressing the use of language to improve social and academic functioning may improve problem behaviors. To address these questions, a longitudinal population-based study is needed, with dimensional associations between language abilities and problem behaviors tracked over time in all children, as well as measurements regarding the potential mediating roles of peer relations and academic skills.

These questions also point to the need to define clear mechanisms for the demonstrated association between language difficulties and problem behaviors. Although there have been several proposed models...
for this association, one suggested mechanism is that language acts as a tool to enhance emotion regulation, the ability to recognize and regulate one’s emotional state. Language delays or disorders may impair children’s ability to use language to regulate their emotions. Emotion regulation skills have been associated with both internalizing and externalizing behaviors in young children. Language skills have also been associated with executive functioning, another developmental domain that has been associated with problem behaviors. Further research is needed to elucidate the mechanistic pathways from language abilities to the presence of problem behaviors and how these pathways may change over the course of development. It is also possible that these mechanisms may differ for internalizing and externalizing behaviors and help to explain the publication bias noted for internalizing but not externalizing behaviors.

The results of the current meta-analysis have important clinical implications. Because language delays and disorders are associated with a greater rate of problem behaviors even at a young age, it is important to develop interventions to target these behaviors early in development for children with delayed language acquisition. Additionally, with these results, we highlight the importance of assessing both internalizing and externalizing behaviors in children with language disorders, because both types of behaviors were impacted by language disorders.

There were some methodological limitations in this meta-analysis. Authors of many studies failed to report nonsignificant findings. Although efforts were made to contact authors to obtain these data, 4 out of 7 authors contacted either did not respond or no longer had access to the original data. Another significant limitation is the heterogeneity of behavioral measures used by different researchers (see Table 1). Although we intended to do further analyses to examine the impact of language on narrow-band behaviors, such as attention-deficit/hyperactivity disorder–type behaviors as compared with oppositional defiant–type behaviors, classifying specific subscales as assessing only 1 type of behavior was problematic because of the differential item composition across measures.

**CONCLUSIONS**

Results from the included studies revealed that children with language disorders display higher rates of problem behaviors compared with their typically developing peers. The difference in rates of problem behaviors increases over time, but there was no observed difference between internalizing and externalizing behaviors. With these results, we suggest that pediatricians and clinicians should consider assessing for both internalizing and externalizing problem behaviors in children with language disorders and highlight the importance of early intervention.

**ABBREVIATIONS**

CBCL: Child Behavior Checklist
CI: confidence interval
ITSEA: Infant Toddler Social Emotional Assessment

---

**TABLE 3** Moderator Analysis of the Effect of Language Disorder on Problem Behaviors, Controlling for Mean Age of Participants

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate* (SE)</th>
<th>P</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect of mean age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.19 (0.06)</td>
<td>.004</td>
<td>0.07 to 0.31</td>
</tr>
<tr>
<td>Mean age, y</td>
<td>0.07 (0.02)</td>
<td>.001</td>
<td>0.03 to 0.11</td>
</tr>
<tr>
<td>Studies: 41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect sizes: 117</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I² = 70.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>τ² = 0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Estimates are unstandardized regression coefficients.

**TABLE 4** Moderator Analysis of the Differential Impact of Language Disorders on Internalizing Versus Externalizing Behaviors

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate* (SE)</th>
<th>P</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect of behavior type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.39 (0.05)</td>
<td>&lt;.001</td>
<td>0.28 to 0.49</td>
</tr>
<tr>
<td>Externalizing versus</td>
<td>−0.05 (0.07)</td>
<td>.10</td>
<td>−0.20 to 0.10</td>
</tr>
<tr>
<td>internalizing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies: 40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect sizes: 122</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I² = 70.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>τ² = 0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Estimates are unstandardized regression coefficients.
REFERENCES


34. Malay S. The Behaviour of Pre-school Children with Language Delays [doctoral thesis]. Halifax, NS: Mount Saint Vincent University; 1995


42. Raffa SG. Behavioral Problems Among Speech and Language Disordered Children in a Public School Setting [doctoral thesis]. Detroit, MI: Wayne State University; 1990

43. Redmond SM. Peer victimization among students with specific language impairment, attention-deficit/hyperactivity disorder, and typical development. Lang Speech Hear Serv Sch. 2011;42(4):520–535


50. Tam SHF. Speech and Language Impairment and Reading Disability in School Age Children: A Seven Year Follow-Up Study [doctoral thesis]. Toronto, ON: Department of Community Health, University of Toronto; 1995


56. Zubrick AM. Behavior Profiles of Language-Impaired Children Aged 4 Through 7 [doctoral thesis]. Ann Arbor, MI: Department of Psychology, University of Michigan; 1984


Language Disorders and Problem Behaviors: A Meta-analysis
Philip R. Curtis, Jennifer R. Frey, Cristina D. Watson, Lauren H. Hampton and Megan Y. Roberts
Pediatrics 2018;142;
DOI: 10.1542/peds.2017-3551 originally published online July 19, 2018;

Updated Information & Services
including high resolution figures, can be found at:
http://pediatrics.aappublications.org/content/142/2/e20173551

References
This article cites 54 articles, 1 of which you can access for free at:
http://pediatrics.aappublications.org/content/142/2/e20173551#BIBL

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
Developmental/Behavioral Pediatrics
http://www.aappublications.org/cgi/collection/development:behavioral_issues_sub
Cognition/Language/Learning Disorders
http://www.aappublications.org/cgi/collection/cognition:language:learning_disorders_sub
Psychosocial Issues
http://www.aappublications.org/cgi/collection/psychosocial_issues_sub

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
http://www.aappublications.org/site/misc/Permissions.xhtml

Reprints
Information about ordering reprints can be found online:
http://www.aappublications.org/site/misc/reprints.xhtml

American Academy of Pediatrics
DEDICATED TO THE HEALTH OF ALL CHILDREN®
Language Disorders and Problem Behaviors: A Meta-analysis
Philip R. Curtis, Jennifer R. Frey, Cristina D. Watson, Lauren H. Hampton and Megan Y. Roberts
*Pediatrics* 2018;142;
DOI: 10.1542/peds.2017-3551 originally published online July 19, 2018;

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/142/2/e20173551

Data Supplement at: