Article

Hammill Institute on Disabilities

Journal of Learning Disabilities
1–15

DOI: 10.1177/0022219419845070
journaloflearningdisabilities.sagepub.com
SAGE

Identifying Children With Clinical Language Disorder: An Application of Machine-Learning Classification

Laura M. Justice, PhD,[1] Woo-Young Ahn, PhD,[2] and Jessica A. R. Logan, PhD[1]

Abstract

In this study, we identified child- and family-level characteristics most strongly associated with clinical identification of language disorder for preschool-aged children. We used machine learning to identify variables that best classified children receiving therapy for language disorder among a sample of 483 3- to 5-year-old children (54% affected). Using a dichotomous outcome based on receipt of language therapy, we applied the least absolute shrinkage and selection operator (LASSO) classification approach to a range of background data available on the children, including teacher and caregiver ratings of communication and social skills. The sample was randomly split into a training (67% of children) and test sample (33% of children) to examine out-of-sample classification accuracy. The full model had excellent classification accuracy based on area under the curve (AUC) of .87 and .85 on the training and test sets, respectively, when utilizing all available background data. Variables most strongly contributing to accurate classification of language-therapy receipt were cognitive impairment, age, gender, and teacher- and parent-reported communication, social, and literacy skills. Use of machine-learning approaches to classify children receiving language services in school settings may provide a valuable approach for identifying those factors that best differentiate children with and without language disorders from a clinical perspective.

Keywords

disorders, language, early identification/intervention, identification/classification

Of the nearly 7,000,000 children with disabilities educated within our nation's school system through the Individuals with Disabilities Education Act (IDEA), about one-tenth—or 699,000—are preschool-aged children served through Part B (Digest of Education Statistics, 2015). The majority of these children qualify for special education services due to presence of impaired language skills (Digest of Education Statistics, 2015), which for many of these children represents a forerunner for future learning disabilities and/or language impairment in the primary grades. That is, longitudinal studies of young children with language disorders find that they are highly susceptible in their future for reading, spelling, and math disabilities (Young et al., 2002). In the present study, we explored child and family factors that may have strong explanatory power for early identification of language disorders in children, which may in turn result in improved intervention in the preschool and primary grades. We applied machine-learning techniques to determine the most salient, defining characteristics of children that differentiated those with clinically identified language disorders and their typically developing peers in an effort to determine factors that may be especially relevant to clinicians' identification practices.

A language disorder occurs when a child shows a persistent inability to acquire and use language skills as would be projected based on normative age-based expectations (American Psychiatric Association, 2013); this disorder is deemed "primary" or "specific" when there is no clear explanation for these lags in language skill. A majority of children identified with primary language disorder at school entry will continue to have significantly depressed language skills over time (Webster, Majnemer, Platt, & Shevell, 2004), show difficulties with kindergarten readiness (Pentimonti, Murphy, Justice, Logan, & Kaderavek, 2016), and have difficulties learning to read (Catts, Fey, Tomblin, & Zhang, 2001), the latter due in part to its effects on higher-level language skills (Hogan, Bridges, Justice, & Cain, 2011). Early childhood language disorder is also

[1]The Ohio State University, Columbus, OH, USA
[2]Seoul National University, Seoul, South Korea

Corresponding Author:
Laura M. Justice, The Ohio State University, 1945 North High, Columbus, OH 43210, USA.
Email: justice.57@osu.edu

linked to heightened risk for psychiatric concerns, attentional difficulties, social-behavioral problems, and learning disabilities in adolescence (Beitchman et al., 1996; Stanton-Chapman, Justice, Skibbe, & Grant, 2007; Young et al., 2002). Given the relatively high incidence of this childhood disability and its significant effects on numerous areas of development, there is great interest in ensuring accurate identification and early intervention for young affected children.

A defining characteristic of primary language disorder is that its cause is unknown (Bishop, 2014). Because of this, diagnosis of language disorder in young children cannot rely on presence of a known factor or condition aside from lags in acquisition of language skills relative to normative benchmarks. Thus, diagnosis of this developmental disability is based exclusively on the application of inclusionary and exclusionary criteria. This typically involves ruling out comorbidities, such as hearing loss and autism spectrum disorder, and potentially other linguistic issues (e.g., dialectical variations that may be conflated with language disorder), subsequent to which a cut-point is applied to scores on a comprehensive, norm-referenced assessment of language ability (Catts et al., 2001; Tomblin et al., 1997). Although there has been a tradition in the research literature to apply exclusionary criteria specific to nonverbal cognition, such that children with low nonverbal cognition are not considered to have primary language impairment, this practice has largely been discredited (see Bishop, 2014). Consequently, within the more recent research literature focused on children with primary language disorder, low nonverbal cognition is not used for inclusionary or exclusionary purposes (see Norbury et al., 2016). Clinicians practicing within the field may utilize a more holistic set of factors to inform their diagnostic decision making, to include clinical judgement derived from observations, interviews, and criterion-related probes; this approach is necessary for determining the extent to which a child's language limitations may affect his or her functioning and participation in everyday events. Recent efforts to build consensus on how best to identify language disorder in children emphasize the importance of drawing on information from multiple sources that reflect the child's language skills in various contexts (Bishop, Snowling, Thompson, Greenhalgh, & CATALISE Consortium, 2016).

Some evidence suggests that there is limited congruence between researcher- and clinician-identified samples of children with language disorder (Bishop & McDonald, 2009; Schmitt, Justice, Logan, Schatschneider, & Bartlett, 2014). Such research points to the considerable ambiguity as to the most appropriate and objective means by which primary disorder of language in children should be identified. In a recent review on this topic, experts noted that no "gold standard" in identifying primary language disorder can currently be applied as optimal diagnostic indicators have yet to be identified; they further note that once these optimal indicators are identified, a subsequent challenge is to identify which indicators contribute to an accurate diagnosis and which do not (Reilly et al., 2014). Relatedly, recent consensus studies focused on improving identification of children with language disorders noted that progress on this issue remains seriously constrained by a "lack of suitable tools" and valid assessment methods (Bishop et al., 2016; Bishop, Snowling, Thompson, Greenhalgh, & CATALISE-2 Consortium, 2017).

In the present study, we sought to advance our understanding of the characteristics of young children with clinically identified language disorder, with a focus on identifying those child- and family-level factors that best serve to differentiate these children from their typically developing classmates in early childhood special education (ECSE) settings. In this research, we used machine learning to generate a prediction model that would accurately differentiate between 3- to 5-year-old children receiving (n = 259) and not receiving language therapies in their preschool programs (n = 224). Our intent was to determine characteristics of children and their families that are associated with clinically identified language disorders. Our approach is similar to that of Morgan and colleagues (2016), who sought to identify factors predictive of children's receipt of speech-language services during the preschool years. In that study, researchers identified characteristics of children who received speech-language services during the preschool years in comparison to children who did not receive such services. That work was instrumental for demonstrating that racial, ethnic, and socioeconomic characteristics of children and their families were associated with access to services.

The children in the present study were participating in inclusive ECSE programs and represented children with and without clinically identified language disorders. A considerable number of background variables (26) were available for each child, such as maternal education, socioeconomic status, and parent and teacher ratings of children's functional communication skills within the home and classroom, among others. Some approaches to identifying language disorders focus on identifying significant group-mean differences between affected and nonaffected children (e.g., LaParo, Justice, Skibbe, & Pianta, 2004; Rice, Buhr, & Nemeth, 1990), such as prosocial skills or cognitive abilities, potentially as a means to identify markers or mechanisms associated with the disorder. However, these differences may not be effective at classification, especially because such work cannot fully represent and account for the multivariate patterns among predictor variables.

Machine learning is a promising approach for identifying multivariate patterns among variables in one data set, which are then tested for classification accuracy in new samples (data sets), which is a procedure that can increase

Table 1. Overview of Variables Used to Generate Four Prediction Models.

Model 1 (Results)	Model 2	Model 3	Model 4
Basic demographics	Basic demographics	Basic demographics	Basic demographics
Communication skills		Communication skills	Communication skills
Literacy skills		Literacy skills	Literacy skills
Social skills			
Problem behaviors			
Parent literacy supports			Parent literacy supports
Print interest			Print interest
Home literacy environment			Home literacy environment
Household chaos			Household chaos

its generalizability. While conventional univariate methods compare (multiple) groups in an univariate way (comparing each measure one at a time), machine-learning (supervised learning) methods search for multivariate predictive patterns of data that optimize prediction accuracy in new samples.

While traditional approaches such as forward/backward stepwise regression can be useful when dealing with a small set of independent variables, they may be inappropriate for a large number of independent (predictor) variables, or its dimension is large compared to the number of samples. Machine-learning methods with cross-validation can handle the dimensionality issue, and certain machine-learning methods can automatically perform variable selection among a number of variables and provide a parsimonious model for the classification, which is challenging with conventional univariate methods.

The first goal of this study was to identify the most parsimonious set of variables that served to reliably classify children with clinically identified language disorders from nonaffected children, drawing on data available representing child- and family-level factors. Model 1, the full model, used all available data for children and families and represents the best test of classification accuracy (see Table 1).

The full model was then followed by three decreasingly parsimonious models to explore alternatives approaches to classification. Model 2 examined classification accuracy relying solely on basic background data that would be readily available for most young children, namely, age, gender, income, and maternal education as an index of socioeconomic status. These background variables are associated with children's language skills in the early years of life (LaParo et al., 2004); we explored the extent to which these would allow for accurate classification of clinically identified language disorders in children. Model 3 added parent and teacher report of children's communication skills at home and in the classroom and teacher report of children's literacy skills; here, we sought to determine the extent to which basic caregiver and teacher report instruments may augment basic background data in accurate classification of clinically identified language disorders. Some work has suggested that caregiver report of children's language skills coupled with standardized language assessments is better able to discriminate between children with and without language impairment than either parent report or standardized assessment alone (Bishop & McDonald, 2009). Consequently, we assessed whether a high degree of classification accuracy could be obtained using a combination of four basic background variables plus parent and teacher report of children's skills at home and school. Finally, Model 4 added contextual information representing the children's home environment, specifically, parents' literacy beliefs, reading practices, and household chaos. The role of the environment for shaping children's early language abilities is well established (Hayiou-Thomas, Dale, & Plomin, 2012), and there are documented differences in the home-learning environments of children with language impairment relative to their peers (Justice, Logan, Işıtan, & Saçkes, 2016). This final model thus considered the extent of classification accuracy when considering basic background data, caregiver and teacher report instruments, and contextual information related to children's home caregiving environments.

Methods

Study Population

This study involved secondary analysis of data collected during a randomized controlled trial (RCT) of an early-literacy intervention conducted in 83 inclusive ECSE classrooms in one state over three consecutive academic years. Each cohort involved a unique, nonoverlapping set of participating teachers, children, and their families. The classrooms generally utilized a 50:50 model such that they served approximately six children with disabilities alongside six typically developing peers. A description of the parent study plus the database used in this study are publicly available (see Acknowledgment for data link).

As part of the RCT, teachers were consented into the study, at which point caregiver consent was solicited for all

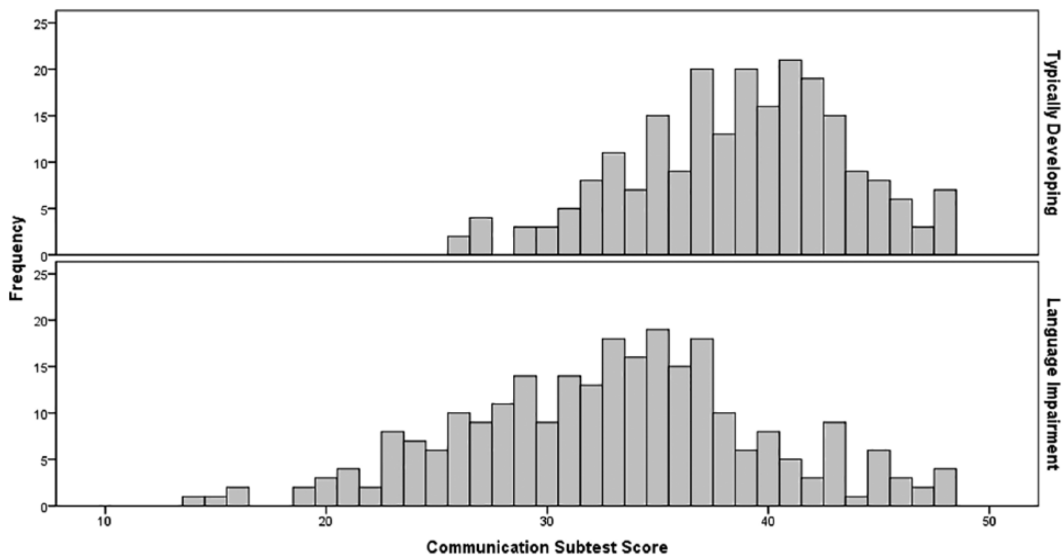

Figure 1. Students' scores on the communication subtest of the Descriptive Pragmatics Profile as rated by parents, for the two samples.

children in each classroom. The average class size was about 13 children (*SD* = 4, range = 5–21), and consent rates averaged about 85% across the 83 classrooms, with consents received for 794 children in total. Of these children, 53% had an individualized education plan (IEP; *n* = 420), which in the United States is a legal record stipulating a child's eligibility for special education services. While we do not have access to documents or data leading to eligibility determination, in the state in which the study was located, statutory requirements for identifying children with disabilities were based on federal legislation. This required use of a complete, individualized evaluation involving multiple technically sound assessment tools and determination of disability by qualified professionals. Eligibility for special education services was to be based on a significant deficit in 1 of 13 categories, including speech or language impairment, or a significant deficit transcending 2 or more areas, referred to generically as *developmental delay* (which could be used in place of speech or language impairment). Eligibility was based on performing 2 standard deviations below the mean in one area or 1.5 standard deviations below the mean in two or more areas on appropriate assessments.

For the children with IEPs enrolled in the larger study (*n* = 420), additional information regarding the children was also gathered, to include determining whether (a) the child had been clinically identified as having a speech/language disorder (88% affirmative) and (b) the child was seen by a speech-language therapist at school (83% affirmative). These data show that the majority of children receiving special education services in preschool settings do so due to speech or language impairment.

For the present purpose, an analysis database representing 483 children was created that omitted any child with missing data on key variables of interest and any child with an IEP who did not have goals related to speech/language and were not receiving speech/language services. The resulting sample was then divided into two subgroups: (a) children with clinically identified language disorder (*n* = 259), all of whom had speech/language goals on their IEP being addressed by an SLP, and (b) children who are typically developing (*n* = 224). Although the approach used to identify membership in the former group potentially could include children with *speech-only* disorders, this is unlikely as a speech-only disorder in the state in which the study was conducted does not typically result in school-based provision of speech-language treatment as it is not perceived to have adverse educational impact.

To further describe the differences between the two samples, we also examined the distributions of the communication subtest of the Descriptive Pragmatics Profile as rated by parents, which is presented in Figure 1 (the same subtest as rated by teachers showed a similar distribution). This subtest shows that children with a language disorder have a wider distribution of skills compared to those children who are typically developing. While some children with a language disorder have very low scores, some children are also had very high scores on this subtest. This demonstrates that while there are some substantial differences between the two samples, the samples are not distinguishable only by mean differences in scores.

As part of the larger study, children with clinically identified language disorders were administered standardized assessments of language skills. These were not administered to children without IEPs. Using the Comprehensive Evaluation of Language Fundamentals- Preschool 2 (CELF; Wiig, Secord, & Semel, 2004), these children scored on average about 1.5 standard deviation below the mean (M = 78.27, SD = 18.42). This same subsample of children were also given the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 2004), a nonverbal IQ test. Children's standard scores ranged from 53 to 124 (M = 85, SD = 18) and showed a relatively normal distribution. This demonstrates that this sample had a wide range of skills.

Measures

In the fall of the academic year, a battery of questionnaires and indirect-report assessments comprising more than 400 individual items were completed by children's primary caregivers and preschool teachers to contextualize the sample and gather information on children's background, experiences, behaviors, and skills. Demographic background provided by caregivers included children's age, gender, home language, parental education (maternal and paternal), child's race and ethnicity, and annual household income. Parents also provided information about their children's home-literacy experiences. Teachers also reported some information about the children, including whether the child has severe/profound cognitive disorder (responding to the question "Does this child have severe or profound cognitive impairment [low IQ and significant functional limitation]?").

Information about children's behaviors and skills was collected via both teacher and parent report for each child using identical indirect report instruments. There is evidence that children's communication, literacy, and social behaviors can vary significantly between the school and home contexts; thus, use of multiple informants can be beneficial (Dinnebeil et al., 2013). An overview of all available measures is included in Table 2, and a comparison of the two samples on all measures is provided in Table 3, which examines differences between samples based on analyses of variance. As can be seen, the two groups significantly differed on most descriptive variables. Critical to this investigation, children who are identified as language impaired are rated significantly lower on their communication skills (nonverbal, conversation, and informal conversation) by both parents and teachers compared to their peers, lending credibility to the idea that they truly have a language impairment.

Statistical Analysis

To classify children with language disorder versus typically developing children, we used the least absolute shrinkage and selection operator (LASSO) machine-learning approach (Tibshirani, 1996). The LASSO is a penalized regression method that automatically selects important variables for classification/prediction and shrinks the coefficients of unimportant variables toward zero. This is performed by imposing L1 penalty, which means the sum of absolute values of coefficients is constrained. We applied the LASSO to each of the four models (Models 1–4) described previously.

For the present analyses, the dependent variable was a dichotomous variable based on whether the child has a clinically identified language disorder (1 or 0), whereas the independent variables are those listed in Table 2. For out-of-sample predictions, we randomly split the entire data set into a training (67%; 322 children) and a test set (33%; 161 children). We estimated the LASSO model using 10-fold cross-validation (CV) using the training set only and then made predictions on the test set. We examined classification accuracy of the LASSO model on the training set for completeness. Finally, we examined the performance of the model on a randomly generated 1,000 sets of training/test sets to make sure that the model performance is robust regardless of how we divided the entire data set into training/test sets. To estimate beta coefficients of the LASSO model, we used 10-fold CV across the whole data set to identify predictors that were the most robust across all samples (Ahn & Vassileva, 2016). We fit the LASSO model using the *easyml* package (Ahn, Hendricks, & Haines, 2017), which provides a wrapper function for the *glmnet* R package (Friedman, Hastie, & Tibshirani, 2010).

As an index of classification accuracy, we used the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. An AUC value of 1 represents a perfect classification model, whereas a value of 0.5 represents a random model. AUC values between 0.9 and 1 are considered outstanding, and AUC values between 0.8 and 0.9 are considered excellent (Hosmer, Lemeshow, & Sturdivant, 2013).

Results

Table 1 (see also Table 2 for descriptions) shows the variables available for the participants and used in the machine-learning application for each of the four models. Using an adjusted p value of $<.002$ as the threshold for statistical significance, given multiple comparisons, children with clinically identified language disorders differed significantly from nonaffected peers on gender, presence of severe cognitive impairment, annual household income (see Table 3). Specifically, children with clinically identified language disorders were more likely to be male; have severe cognitive impairment; have poorer parent- and teacher-rated functional communication skills, early literacy skills, and social skills; have higher levels of problem behaviors per teacher report; have less interest in print; and live in households that are more chaotic, have lower annual household

Table 2. Measures and Corresponding Variables Used to Create Prediction Models.

Focus	Measures	Variables
Basic demographics	Caregiver questionnaire providing information on child age, gender, parental education (maternal and paternal), annual household income, and whether the child has severe/profound cognitive disorder	Age Gender Cognition Gender MomEd Income
Communication skills	Descriptive Pragmatics Profile (DPP; Wiig, Secord, & Semel, 2004) completed by parents and teachers; contains 26 items divided into three subscales: (1) nonverbal skills (comm nonverbal), (2) conversational skills (comm conversation), and (3) asking for, giving, and responding to information (comm information). Respondents rate the child's performance on a 4-point scale for each item. For each subscale, a total score was calculated by summing the scores for each item.	Comm nonverbal – P Comm conversation – P Comm information – P Comm nonverbal – T Comm conversation – T Comm information – T
Literacy skills	Preschool Literacy Rating Scale (PLRS; Wiig et al., 2004) completed by parents and teachers; contains 26 items, with 8 items comprising a subscale of emergent reading behaviors. Respondents rate the child's performance on a 4-point scale for each item. A total score was calculated (literacy skills) as well as a subscale for emergent reading (parents only) by summing the scores for each items.	Literacy skills – P Emergent reading – P Literacy skills –T
Social skills and problem behaviors	Social Skills Rating System (SSRS: Gresham & Elliott, 1990) completed by parents and teachers; contains 40 items focused on cooperation, assertion, and self-control and 10 items focused on problem behaviors (internalizing and externalizing behaviors). A standardized score for social skills and problem behaviors was computed using norm tables.	Social skills – P Social skills – T Prob behaviors – P Prob behaviors - T
Parent literacy supports and home-literacy environment (HLE)	Home-literacy questionnaire compiled from multiple sources (e.g., Bennett Weigel, & Martin, 2002; Fritjers, Barron, & Brunello, 2000; Griffin & Morrison, 1997) completed by parents with multiple items capturing the frequency with which parents engage in basic home literacy practices and hold certain beliefs. HLE is a total score across all items (home literacy practices). Home reading represents how often parents read with their children. Print interest represents how interested children are in print activities (e.g., how often they look at books on their own). Lit teaching represents the extent to which parents directly teach their children about print, lit beliefs focuses on whether parents believe reading to their children is important and that they have a role to play in cultivating their child's skills. The Title Recognition Test (Home literacy-TRT, Cunningham & Stanovich, 1990), examines parent familiarity with children's storybooks, thought to be a proxy for parent-child home reading frequency.	Home literacy practices Home reading Print interest Lit teaching Lit beliefs Home literacy – TRT
Household chaos	Parents completed a subset of items from the Confusion, Hubbub, and Order Scale (CHAOS; Matheny, Wachs, Ludwig, & Phillips, 1995).	Chaos

income, and provide fewer reading experiences (per the parent Title Recognition Test). These analyses show that the sample of children with clinically identified language disorders differed in key ways from their ECSE classmates. However, these analyses do not demonstrate which variables most contribute to classifying children with clinically identified language disorders, which was addressed using the machine-learning approach described previously.

Machine learning was applied to four different models (see Table 1). Model 1, the full model, included all available variables as potential contributors to classification accuracy. Figure 2 shows the multivariate patterns of the available background variables used for the potential classification of children with language disorder. To identify those variables most useful for classification, based on statistical significance and effect size, we consider those with a value of .25 or higher (conversely, −.25 or lower) to be most useful. Seven variables were important to classification of language disorder: presence of cognitive impairment, gender, age, pragmatic skills (teacher report and parent report), social skills (teacher report), and literacy skills (teacher report). Figure 3 shows the distribution of the AUC values on the (A) training and (B) test sets over 1,000 random divisions of training/test sets. The mean AUCs were .87 and .84 for the training and test sets, which is considered to be excellent in terms of classification accuracy.

Table 3. Demographic, Household, and Indirect-Report Data for Children With Clinically Identified Language Impairment (LI) and Typical Classmates.

Variable	LI (n = 259)	Typical (n = 244)	Test Statistic (F)	Significance
Age (in months)	51.8 (7.3)	51.5 (6.3)	0.2	.639
Gender (% female)	25	49	29.5	<.001
Cognition (% severe cognitive impairment)	10	0	24.9	<.001
MomEd	5.8 (2.5)	6.5 (2.4)	8.9	.003
Income	10.5 (6.1)	12.7 (5.2)	18.1	<.001
Comm nonverbal – P	24.4 (4.0)	26.2 (2.7)	31.6	<.001
Comm conversation – P	32.9 (6.8)	38.7 (4.8)	112.9	<.001
Comm information – P	18.8 (4.6)	22.8 (3.4)	112.5	<.001
Comm nonverbal – T	22.2 (5.0)	25.6 (3.6)	70.9	<.001
Comm conversation – T	30.4 (8.0)	39.2 (6.7)	170.1	<.001
Comm information – T	16.8 (5.3)	22.2 (4.6)	142.3	<.001
Literacy skills – P	58.3 (16.0)	71.7 (15.9)	84.9	<.001
Emergent reading – P	32.0 (8.3)	37.4 (6.7)	61.7	<.001
Literacy skills –T	50.8 (15.5)	65.6 (16.7)	102.2	<.001
Social skills – P	86.0 (18.1)	100.7 (13.9)	116.5	<.001
Social skills – T	90.4 (18.0)	106.0 (15.9)	133.6	<.001
Prob behaviors – P	100.0 (14.8)	96.5 (11.1)	13.7	<.001
Prob behaviors - T	98.8 (11.5)	92.1 (10.2)	62.7	<.001
Home literacy practices	1.4 (0.6)	1.5 (0.5)	5.3	.021
Home reading	14.1 (6.1)	15.2 (5.4)	4.0	.045
Print interest	6.1 (5.4)	8.5 (4.9)	25.2	<.001
Lit teaching	6.1 (4.8)	6.7 (4.8)	1.9	.174
Lit beliefs	2.3 (0.4)	2.2 (0.4)	2.5	.113
Home literacy-TRT	3.7 (2.3)	4.7 (2.6)	18.0	<.001
Chaos	1.6 (0.6)	1.4 (0.6)	9.5	.002

Note. MomEd based on an ordinal variable with 11 ordered categories representing highest degree earned ranging from 1 = < high school diploma to 11 = doctoral degree; income based on an ordinal variable with 18 ordered categories representing annual household income ranging from 1 = < $5,000 to 18 = >$85,000; for prob behaviors, lower scores represent fewer problem behaviors; for chaos, lower scores represent less household chaos.

Models 2, 3, and 4 were used to determine whether the level of classification accuracy found with all variables could be achieved using a smaller set of variables. Model 2 included only basic background information, namely, age, maternal education, household income, and child gender. Each also contributed significantly to classification (see Figure 4), but prediction accuracy was poor (mean AUC = .68 and mean AUC = .66 for the training and test sets, respectively). This finding shows that classification of clinically identified language disorders draws on much more information than simply background factors related to age, gender, and socioeconomic status.

Model 3 incorporated teacher and parent reports of children's communication skills (Descriptive Pragmatics Profile; Wiig et al., 2004) and teacher report of children's literacy skills (Preschool Literacy Rating Scales; Wiig et al., 2004). Though it included only eight variables, this model had excellent classification accuracy (mean AUC = .86 and mean AUC = .85 for the training and test sets, respectively; see Figure 5). Clinically identified language disorder was classified by older age, lower maternal education and income, being male, lower literacy, and lower functional communication skills. The results of this model suggest that accurate classification of clinically impaired language skills may rely heavily on functional skills (per teacher and parent report) in literacy and communication.

In Model 4, we added measures representing the child's home environment, including home-literacy activities. As shown in Figure 6, Model 4 performed similarly to Model 3 on the test set (mean AUC = .87 and mean AUC = .85 for the training and test sets, respectively). Thus, incorporating information, the home context children experienced did not contribute in any way to accurate classification.

Discussion

Language disorder is one of the most common developmental disabilities to affect young children, with recent estimates showing a prevalence rate of nearly 10% (Norbury et al., 2016). As many as 40% and 50% of these youngsters

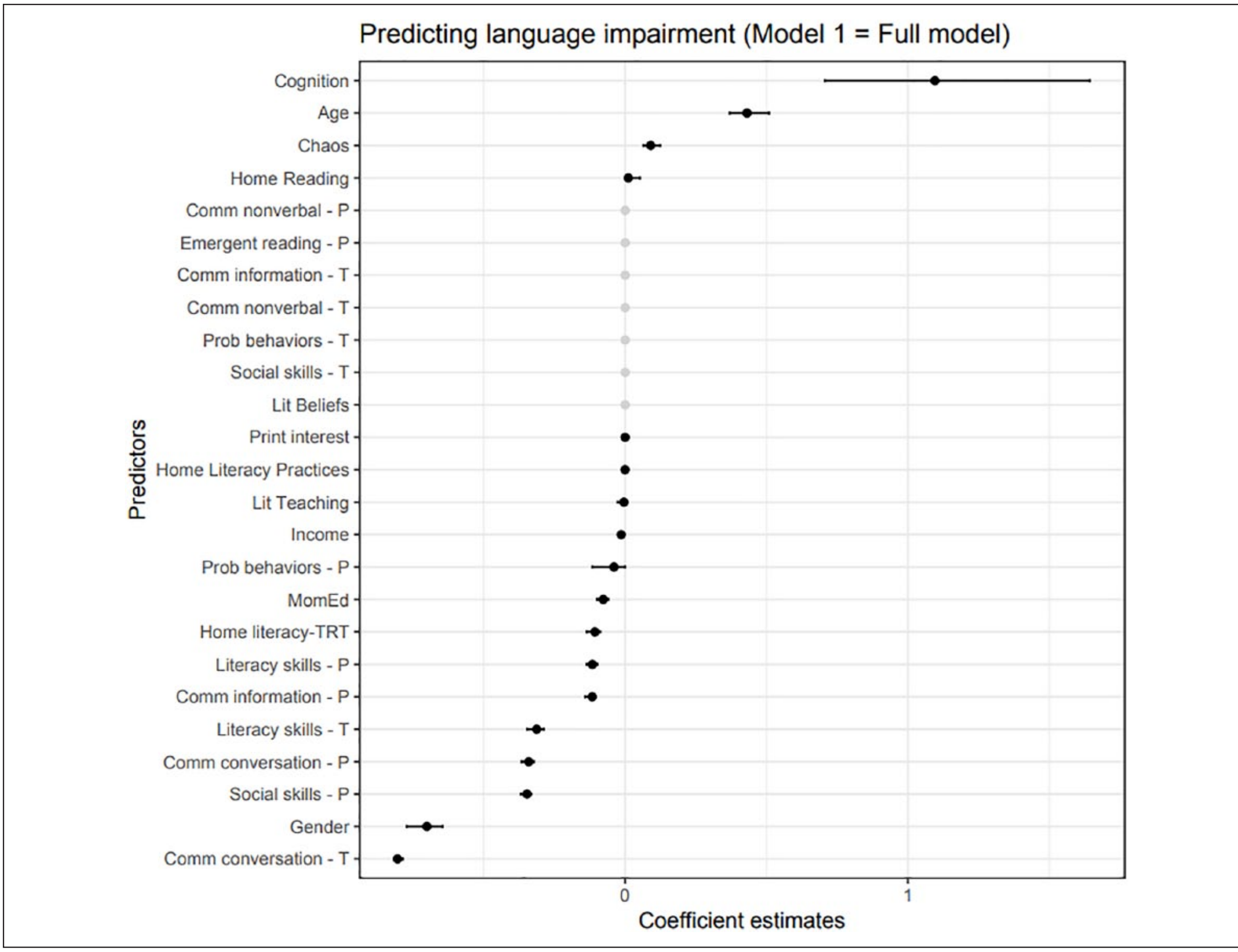

Figure 2. Multivariate patterns of background variables classifying children with language disorder (see Table 2 for variable names). Variables with gray color: Those whose effects are shrunk to 0. Coefficient estimates indicate beta estimates of variables in a LASSO model.

will meet criteria for reading and math disabilities, respectively (Young et al., 2002). For most of these children, the cause is unknown (Norbury et al., 2016); thus, clinicians who diagnose and treat language disorders in young children rely on holistic assessments of children's background (e.g., household characteristics), functional skills in communication and related areas (e.g., literacy), and children's performance on norm-referenced measures of language skill. Statutory practices for identifying children for speech-language services in the schools allow for considerable flexibility in how identification occurs, including the types and nature of assessment tools used. There is a significant need to better understand those characteristics that differentiate children with clinically identified language disorders from nonaffected peers as this may serve to identify how affected children are distinguishable from their peers (see Bishop et al., 2016, 2017). Although the children served by speech-language pathologists within the public schools are a heterogeneous group in terms of the language skills affected (Tambyraja, Schmitt, Farquharson, & Justice, 2015), the present study was predicated on the premise that some general factors, such as nonverbal cognition and social competencies, may serve to distinguish children with clinically impaired language skills from their peers.

Given the need for advances in this area, the present paper represents an innovative effort to apply machine learning to the classification of language disorders. There is increased momentum surrounding the use of machine learning to increase accurate classification of various diseases and disabilities (e.g., Ahn & Vassileva, 2016). Here, machine learning was used to identify factors that accurately classified young children with language disorders, representing children identified via clinical practices in the schools, from their peers. In considering these results, it is

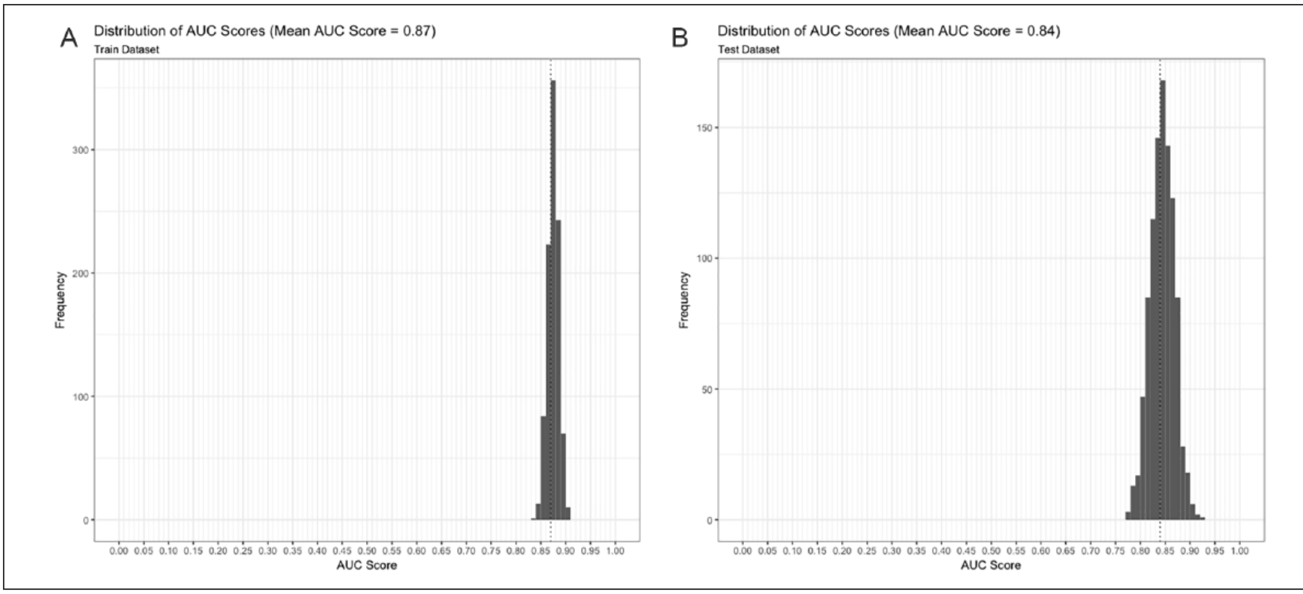

Figure 3. Distributions of the area under the curve of the receiver operating characteristic curves on (A) training and (B) test sets over 1,000 random divisions of training/test sets.

important to note that our population of interest—children with language disorders—may not be perfectly reflected by our available sample in that there is imperfect overlap between clinically identified children and the population of children with this disorder. This occurs for multiple reasons, including under-identification of children with language disorders by current clinical practices and lack of uptake of clinical services by affected children and their families. Consequently, in considering the results of this study, our results must be refined to reflect only those children identified to receive school-based speech-language services, with generalizability to the more general population of children with language disorders presently unknown.

Of particular importance is the finding that a relatively parsimonious set of seven variables provided excellent classification accuracy, correctly identifying approximately 85% of children into their assigned group. Most prominently, knowledge of basic background factors reflecting child age, gender, and socioeconomic status indicators coupled with teacher and parent report of children's functional language and early literacy skills had classification accuracy of .85 for our sample. In fact, we found that seven variables have excellent classification accuracy for determining presence of language disorder: Being a boy, being older in age (i.e., closer to 5 years), having poorer parent- and teacher-reported functional communication and literacy skills, and being lower socioeconomic status have strong discriminatory value in differentiating children with clinically impaired language skills from nonaffected peers. All such factors have previously been referenced in the literature as contributors to language disorders, yet this application of machine learning supports several key prior findings.

First, being older in age among this 3- to 5-year old sample was a strong contributor to classification accuracy. This may reflect a tendency for diagnosis of language disorder to occur later rather than sooner during the preschool years, reflecting a "wait and see" approach; in this regard, the age effect may reflect clinical practices. Alternatively, the role of age in classification accuracy could possibly represent the transient nature of early childhood language disorders, with a significant portion of young affected children resolving their language problems from age 3 to 5 years (Dale, Price, Bishop, & Plomin, 2003; LaParo et al., 2004). Language acquisition is a highly dynamic process, and language skills become increasingly stabilized as children grow older (Bornstein, Hahn, & Putnick, 2016). It is important to consider whether this result suggests that diagnosis of language disorder should not occur until near kindergarten entry as some work has suggested that this is an effective period to identify children with more stable (vs. transient) forms of this disability (Justice, Bowles, Pence Turnbull, & Skibbe, 2009). However, this should not be interpreted to suggest withholding language supports for young children with or at risk for lags in language development; a number of evidence-based strategies are available to enhance the early language growth of children exhibiting lags in this area of development (Roberts & Kaiser, 2011).

Second, being a boy was highly associated with classification accuracy, confirming a large body of extant research indicating a higher prevalence of language disorders in young males (Tomblin et al., 1997). It is unclear whether

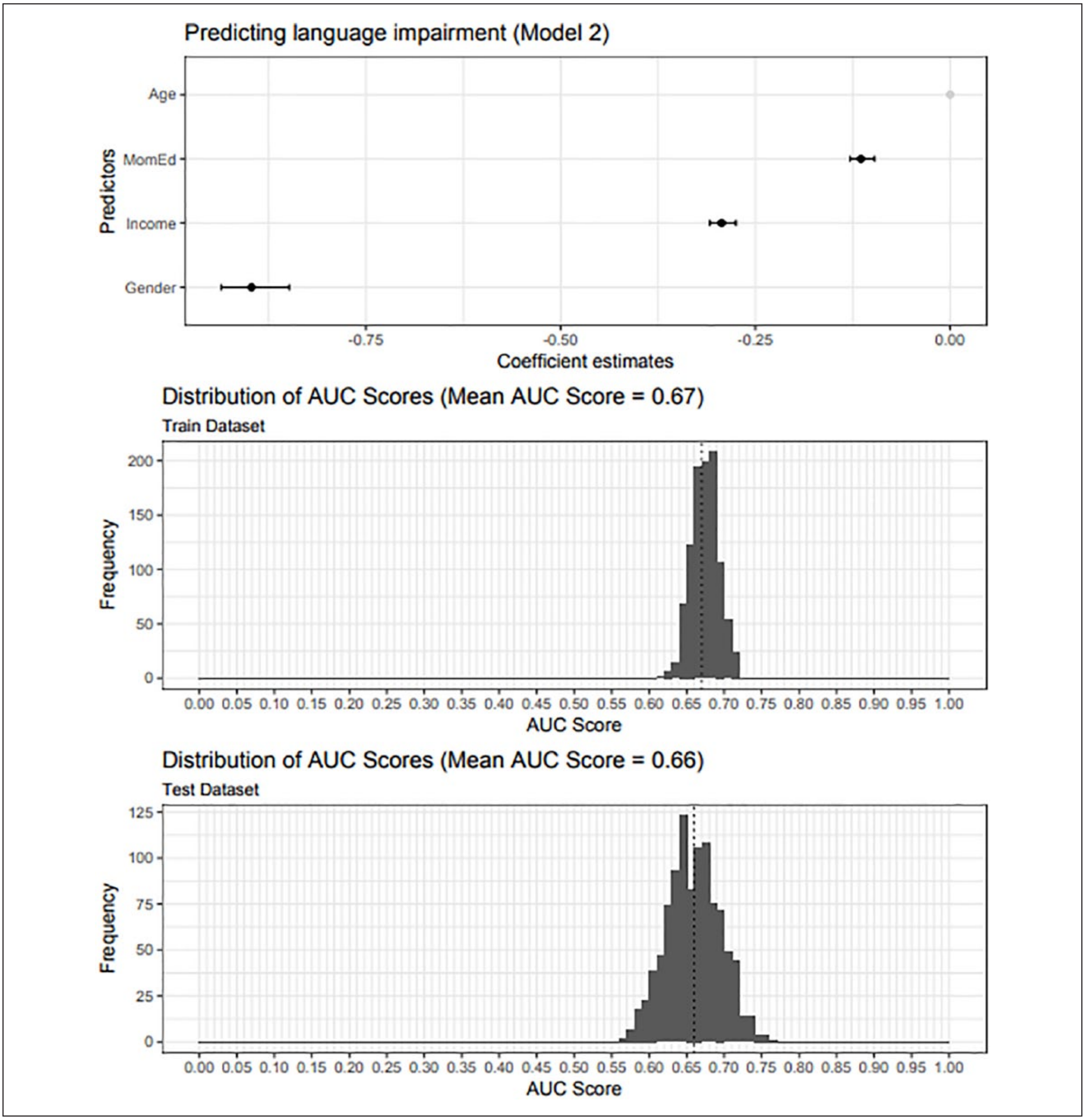

Figure 4. Model 2 results for classifying children with language disorder using four basic background variables (see Table 2 for variable names). Coefficient estimates indicate beta estimates of variables in a LASSO model.

the increased prevalence of language disorders among young boys reflects some type of sex-linked heritable genetic transmission or whether this represents differences in the rate at which key language skills are acquired across the genders. Further, recent evidence suggests that girls with LI have relatively better prosocial behaviors than boys with LI, potentially suggesting that the more positive prosocial behaviors of girls can buffer the consequences of LI within social settings (Toseeb, Pickles, Durkin, Botting, & Conti-Ramsden, 2017). Nonetheless, the present findings show that being male is a significant factor in classification of clinically impaired language skills, thus raising attention toward the specific vulnerability of boys for this developmental disability.

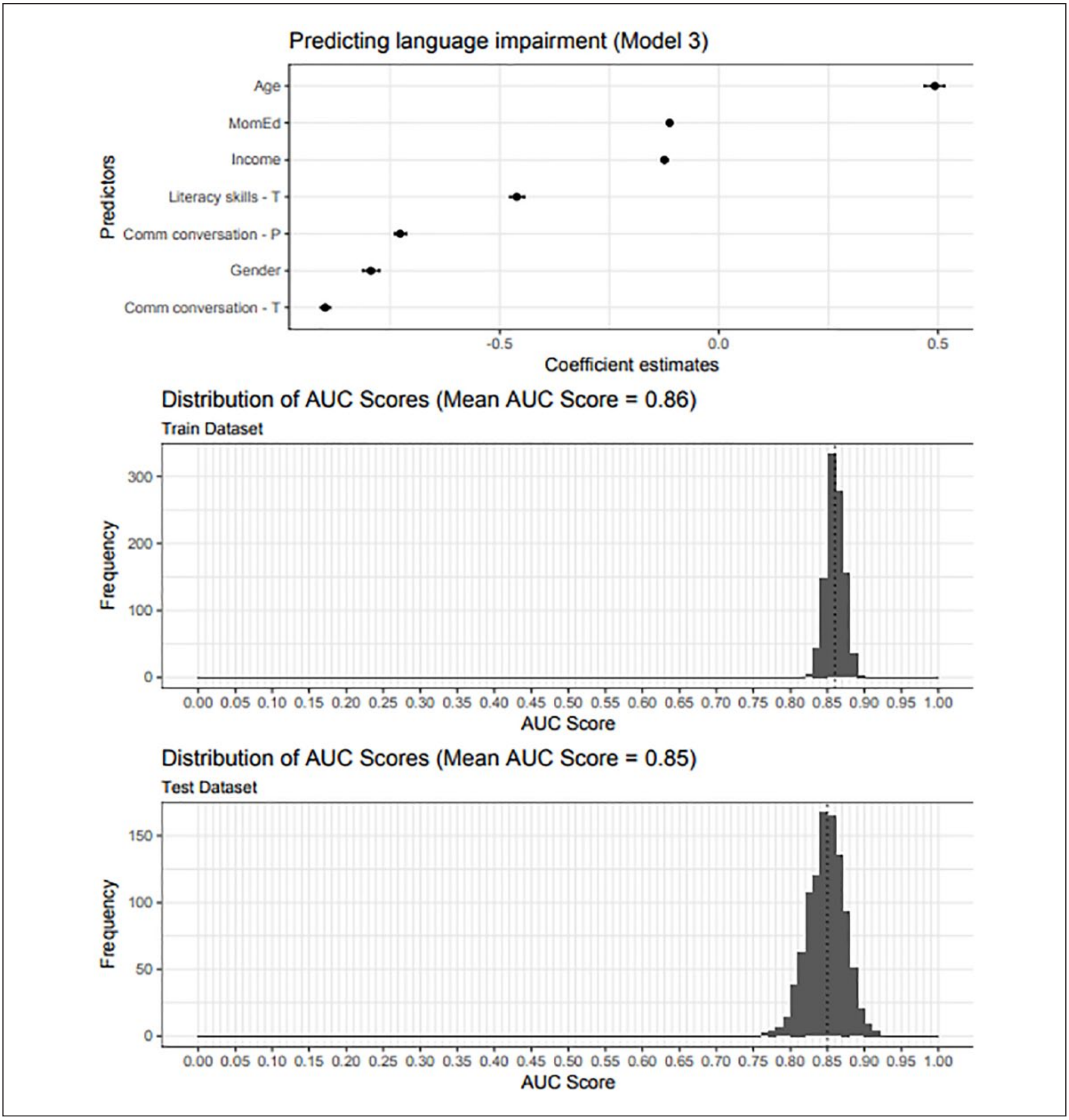

Figure 5. Model 3 results for classifying children with language disorder using basic background variables plus teacher and parent report tools (see Table 2 for variable names). Coefficient estimates indicate beta estimates of variables in a LASSO Model.

Third, we find that parent and teacher reports of children's functional communication and early literacy skills contributed strongly to classification accuracy: Children with language disorders were rated as much poorer than nonaffected children on their abilities to use language as a tool to communicate in the home and classroom and had poorer early literacy skills based on teacher report. This finding is a particularly compelling one as recent consensus statements have emphasized the importance of (a) using information on children's language skills from multiple informants (b) and ensuring that attention is paid to how children use language in pragmatic and social contexts

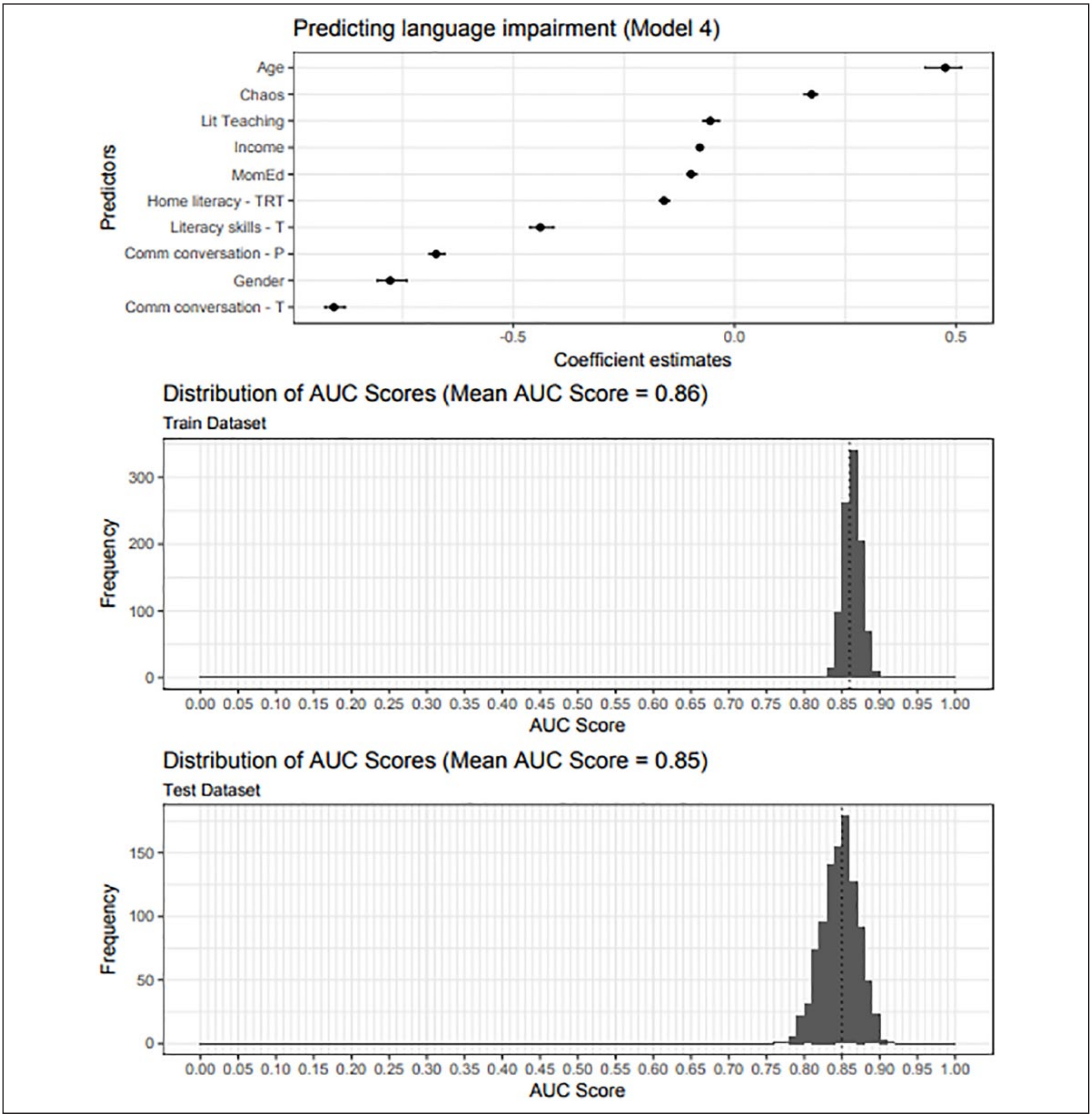

Figure 6. Model 4 results for classifying children with language disorder using basic background variables plus teacher and parent report tools and home environment measures (see Table 2 for variable names). Coefficient estimates indicate beta estimates of variables in a LASSO model.

(Bishop et al., 2016, 2017). Such perspectives reflect long-standing theoretical perspectives proposing that an individual's language skill is best represented by one's agility in using language instrumentally to meet one's communicative needs and interact with others (e.g., Tomasello, 2009). In the present work, information about how children use language functionally was captured with relatively simple teacher- and parent-report checklists and therefore do not necessarily require extensive observations in various contexts by trained clinical professionals. This finding converges with prior work indicating that inclusion of parent report of children's communication skills improves classification accuracy between referred and nonreferred cases of childhood language disorder (Bishop & McDonald, 2009).

The significant role of teacher-reported early literacy skills to classification accuracy is especially interesting. Over the past several decades, a substantial body of research has emerged to show the strong association between language disorders and reading disabilities in children (Bishop & Adams, 1990; Skibbe et al., 2008). To our knowledge, no study has shown that early-literacy skills are associated with classification of language disorders in young children; thus, the present findings suggest a new research direction for early and accurate identification of language disorders. Specifically, the present results suggest that children's early-literacy skills may be relatively important in identifying language disorders in young children, potentially because of the robust associations between early language skill and children's literacy development. In the present sample, children clinically identified as having a language disorder had much poorer literacy skills, based on teacher report, than their typically developing classmates; on average, the former group of children had literacy skills about one standard deviation below their typical peers. While similar results have been reported elsewhere (Cabell, Justice, Zucker, & McGinty, 2009; Justice, Bowles, & Skibbe, 2006), we are aware of no work indicating that low levels of literacy skill are used diagnostically for identification of language impairment in children.

On the other hand, information concerning children's home-literacy environment, household organization and chaos, and parent beliefs appeared to have limited utility for improving classification accuracy. While it is well recognized that features of home environment are instrumental in providing children with enriching opportunities to acquire language skills (Hoff, 2003; Landry, Smith, & Swank, 2006), we find no evidence that these opportunities contribute to the identification of language disorders in children or, potentially, serve as causal contributors to this condition. This is an important finding as historically there have been efforts to differentiate language problems in children that are a product of environmental deprivation (e.g., limited caregiver input) versus those that are biologically based; as discussed by Bishop (2014), however, it is often impossible to disentangle the causes of a child's language disorder. The present findings suggest that clinical identification of language disorders in children is not influenced by characteristics of children's home learning environment even while several such characteristics did significantly differ across the two samples of children. That is, children with clinically identified language disorders had less educated mothers, had homes with higher levels of chaos, and their mothers read less frequently to them (based on the Title Recognition Test) than their typically developing classmates. While these household factors differentiated the two groups of children in key ways, these factors were not influential to diagnostic classification.

Conclusion and Limitations

Language disorders are one of the most commonly occurring disabilities to affect young children. Despite their high prevalence, affecting about 1 in 10 young children, there are unresolved issues regarding how to effectively identify these children to ensure they receive the treatments they need. This study, through use of a machine-learning approach, provides evidence that a relatively small set of key factors, including basic background variables as well as teacher and parent report of functional communication and literacy skills, allow for an excellent level of classification accuracy for clinically identified childhood language disorders. We intend for the present findings to lead to new, innovative efforts to improve clinical practices and research endeavors focused on this problem and other disorders.

Several limitations warrant note and can help guide future research on this and related topics. First, there are limitations regarding the sample of children with language disorders. These children were identified via clinical practices within the public school system. It is unclear whether the same variables that led to high levels of classification accuracy in the present sample would occur for a sample of children with language disorders identified using a different set of criteria. Similarly, this work was conducted with a subset of all students for whom data were available, and therefore our conclusions apply to a relatively smaller sample that may not be representative of all students with language disorders. Also, neither the severity of children's language problems nor the possibility of concomitant diagnoses was considered in the models. Classification accuracy that attends to these issues should be explored in future work. Second, the machine-learning results represent those available from the present set of available variables. While our models had a high degree of classification accuracy, potentially greater accuracy would be observed with a different set of variables. Third, while the use of machine-learning methods for classification of numerous health conditions is on the rise, it is important to critically examine automated decisions when using artificial intelligence. Thus, the results of the present study must be replicated with other samples and using other methodological approaches. Finally, we also must highlight that this work suggests but does not confirm a one-way prediction model, with a set of independent variables (e.g., parent rating of children's communication skills) predictive of the dependent variable (language status). However, the predictors themselves may be the symptoms of the disorder. Given such complexities, application of emerging methodologies, including artificial intelligence, may have great promise for improved methods of identification and intervention for children with communication disorders.

Acknowledgments

The data used in this study are publicly available at https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36738. The authors are grateful to the teachers, families, and children who participated in this research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Grant #R324A080037 from the Institute of Education Sciences to the first author.

References

Ahn, W.-Y., Hendricks, P., & Haines, N. (2017). Easyml: Easily build and evaluate machine learning models. *bioRxiv*, 137240.

Ahn, W.-Y., & Vassileva, J. (2016). Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence. *Drug and Alcohol Dependence*, *161*, 247–257.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing, Inc.

Beitchman, J. H., Wilson, B., Brownlie, E., Walters, H., Inglis, A., & Lancee, W. (1996). Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(6), 815–825.

Bennett, K. K., Weigel, D. J., & Martin, S. S. (2002). Children's acquisition of early literacy skills: Examining family contributions. *Early Childhood Research Quarterly*, *17*(3), 295–317.

Bishop, D. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language & Communication Disorders*, *49*(4), 381–415.

Bishop, D., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry*, *31*(7), 1027–1050.

Bishop, D., & McDonald, D. (2009). Identifying language impairment in children: Combining language test scores with parental report. *International Journal of Language and Communication Disorders*, *44*(5), 600–615.

Bishop, D., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE Consortium. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS One*, *11*(7), e0158753.

Bishop, D., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE-2 Consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068–1080.

Bornstein, M. H., Hahn, C.-S., & Putnick, D. L. (2016). Long-term stability of core language skill in children with contrasting language skills. *Developmental Psychology*, *52*(5), 704–716.

Cabell, S. Q., Justice, L., Zucker, T. A., & McGinty, A. S. (2009). Emergent name-writing abilities of preschool-age children with language impairment. *Language, Speech, and Hearing Services in Schools*, *40*(1), 53–66.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. (2001). Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation. *Language, Speech, and Hearing Services in Schools*, *32*(1), 38–50. doi:10.1044/0161-1461(2001/004)

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology*, *82*(4), 733–740.

Dale, P. S., Price, T. S., Bishop, D., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research*, *46*(3), 544–560.

Digest of Education Statistics. (2015). *Children 3 to 21 years old served under Individuals with Disabilities Education Act (IDEA), Part B, by type of disability: Selected years, 1976-77 through 2013-14*. Retrieved from https://nces.ed.gov/programs/digest/d15/tables/dt15_204.30.asp

Dinnebeil, L. A., Sawyer, B. E., Logan, J., Dynia, J. M., Cancio, E., & Justice, L. M. (2013). Influences on the congruence between parents' and teachers' ratings of young children's social skills and problem behaviors. *Early Childhood Research Quarterly*, *28*(1), 144–152.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22.

Frijters, J. C., Barron, R. W., & Brunello, M. (2000). Direct and mediated influences of home literacy and literacy interest on prereaders' oral vocabulary and early written language skill. *Journal of Educational Psychology*, *92*(3), 466–477.

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system: Manual*. Circle Pines, MN: American Guidance Service.

Griffin, E. A., & Morrison, F. J. (1997). The unique contribution of home literacy environment to differences in early literacy skills. *Early Child Development and Care*, *127*(1), 233–243.

Hayiou-Thomas, M. E., Dale, P. S., & Plomin, R. (2012). The etiology of variation in language skills changes with development: A longitudinal twin study of language from 2 to 12 years. *Developmental Science*, *15*(2), 233–249.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368–1378.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). New York, NY: John Wiley & Sons.

Justice, L. M., Bowles, R. P., Pence Turnbull, K. L., & Skibbe, L. E. (2009). School readiness among children with varying histories of language difficulties. *Developmental Psychology*, *45*(2), 460–476.

Justice, L. M., Bowles, R. P., & Skibbe, L. E. (2006). Measuring preschool attainment of print-concept knowledge: A study of typical at-risk 3- to 5-year-old children using item response theory. *Language, Speech, and Hearing Services in Schools*, *37*(3), 224–235.

Justice, L. M., Logan, J. A., Işıtan, S., & Saçkes, M. (2016). The home-literacy environment of young children with disabilities. *Early Childhood Research Quarterly*, *37*, 131–139.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test*. New York, NY: Wiley Online Library.

Landry, S., Smith, K. E., & Swank, P. R. (2006). Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills. *Developmental Psychology*, *42*(4), 627–642.

LaParo, K. M., Justice, L. M., Skibbe, L. E., & Pianta, R. C. (2004). Relations among maternal, child, and demographic factors and the persistence of preschool language impairment. *American Journal of Speech-Language Pathology*, *13*(4), 291–303.

Matheny, A. P. Jr., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Journal of Applied Developmental Psychology*, *16*(3), 429–444.

Morgan, P. L., Hammer, C. S., Farkas, G., Hillemeier, M. M., Maczuga, S., Cook, M., & Morano, S. (2016). Who receives speech/language services by 5 years of age in the United States? *American Journal of Speech-Language Pathology*, *25*(2), 183–199.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., . . . Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*(11), 1247–1257.

Pentimonti, J. M., Murphy, K. A., Justice, L. M., Logan, J. A., & Kaderavek, J. N. (2016). School readiness of children with language impairment: Predicting literacy skills from pre-literacy and social-behavioural dimensions. *International Journal of Language & Communication Disorders*, *51*(2), 148–161.

Reilly, S., Tomblin, J., Law, J., McKean, C., Mensah, F. K., Morgan, A., . . . Wake, M. (2014). Specific language impairment: A convenient label for whom? *International Journal of Language and Communication Disorders*, *49*(4), 416–451.

Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech, Language, and Hearing Research*, *55*(1), 33–42.

Roberts, M. Y., & Kaiser, A. P. (2011). The effectiveness of parent-implemented language interventions: A meta-analysis. *American Journal of Speech-Language Pathology*, *20*(3), 180–199.

Schmitt, M. B., Justice, L. M., Logan, J. A., Schatschneider, C., & Bartlett, C. W. (2014). Do the symptoms of language disorder align with treatment goals? An exploratory study of primary-grade students' IEPs. *Journal of Communication Disorders*, *52*, 99–110.

Skibbe, L. E., Grimm, K. J., Stanton-Chapman, T. L., Justice, L. M., Pence, K. L., & Bowles, R. P. (2008). Reading trajectories of children with language difficulties from preschool through fifth grade. *Language, Speech, and Hearing Services in Schools*, *39*(4), 475–486.

Stanton-Chapman, T. L., Justice, L. M., Skibbe, L. E., & Grant, S. L. (2007). Social and behavioral characteristics of preschoolers with specific language impairment. *Topics in Early Childhood Special Education*, *27*(2), 98–109. doi:10.1177/02711214070270020501

Tambyraja, S. R., Schmitt, M. B., Farquharson, K., & Justice, L. M. (2015). Stability of language and literacy profiles of children with language impairment in the public schools. *Journal of Speech, Language, and Hearing Research*, *58*(4), 1167–1181.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Tomasello, M. (2009). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomblin, J., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research*, *40*(6), 1245–1260.

Toseeb, U., Pickles, A., Durkin, K., Botting, N., & Conti-Ramsden, G. (2017). Prosociality from early adolescence to young adulthood: A longitudinal study of individuals with a history of language impairment. *Research in Developmental Disabilities*, *62*, 148–159.

Webster, R. I., Majnemer, A., Platt, R. W., & Shevell, M. I. (2004). The predictive value of a preschool diagnosis of developmental language impairment. *Neurology*, *63*(12), 2327–2331.

Wiig, E. H., Secord, W., & Semel, E. M. (2004). *Clinical Evaluation of Language Fundamentals-Preschool* (2nd ed.). Boston, MA: Pearson Education.

Young, A. R., Beitchman, J. H., Johnson, C., Douglas, L., Atkinson, L., Escobar, M., & Wilson, B. (2002). Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. *Journal of Child Psychology and Psychiatry*, *43*(5), 635–645.