Journal of Communication Disorders 78 (2019) 57–68

Contents lists available at ScienceDirect

Journal of Communication Disorders

journal homepage: www.elsevier.com/locate/jcomdis

Differences in the relation between temperament and vocabulary based on children's stuttering trajectories

Cara M. Singer[a,*], Tedra A. Walden[b], Robin M. Jones[a]

[a] Department of Hearing and Speech Sciences, Vanderbilt University, 1215 21st Avenue South, Suite 8310 MCE South Tower, Nashville, TN 37232-8242, United States

[b] Department of Psychology and Human Development, Peabody College, Vanderbilt University, 230 Appleton Place, Nashville, TN 37203-5721, United States

ARTICLE INFO

Keywords:
Stuttering
Temperament
Vocabulary
Children
Chronicity
Longitudinal
Preschool

ABSTRACT

Purpose: The purpose of this study was to assess the relation between temperament and vocabulary development for children who stutter and persist, children who stutter and recover and children who do not stutter.

Methods: Participants, aged 3;0–4;11 at the start of the study, were followed for two years. They were classified as persisting (n = 10), recovered (n = 26), and non-stuttering (n = 24) based on multiple assessments of stuttering spaced across study participation. Groups were balanced for age and gender ratios. At each visit, participants completed the Peabody Picture Vocabulary Test, 4th edition, and the Expressive Vocabulary Test, 2nd edition; caregivers completed the Children's Behavior Questionnaire.

Results: For both persisting and recovered groups, higher negative emotionality was associated with lower receptive vocabulary. These associations were both significantly more negative than the non-stuttering group's association.

Conclusion: These findings suggest that receptive vocabulary development is differentially linked to temperament based on a child's stuttering status. As others have speculated (Conture & Walden, 2012), it appears as though there are salient associations between temperament, speech-language development, and childhood stuttering.

1. Introduction

The Dual-Diathesis Stressor Model (Conture & Walden, 2012) proposes that language and temperament contribute to stuttering. One commonly investigated linguistic factor in early childhood stuttering is vocabulary (for meta-analytic review, see Ntourou, Conture, & Lipsey, 2011). In recent years, temperament has also received attention as an important emotional contributor to stuttering (Jones, Choi, Conture, & Walden, 2014). Current evidence outside the area of stuttering suggests that temperament is associated with vocabulary acquisition in typically developing children (e.g., Dixon & Smith, 2000; Salley & Dixon, 2007). To date, temperament and language have been investigated in isolation of one another in early childhood stuttering, despite speculation that the language skills of children who stutter may be importantly associated with aspects of their temperament (e.g., Ntourou et al., 2011). The present study explored the relation of vocabulary and temperament in childhood stuttering and their possible connection

* Corresponding author at: Department of Hearing and Speech Sciences, Vanderbilt University, 10221 MCE South Tower, 1215 21st Avenue South, Nashville, TN 37232, United States.

E-mail address: cara.m.singer@vanderbilt.edu (C.M. Singer).

https://doi.org/10.1016/j.jcomdis.2019.01.004
Received 19 April 2018; Received in revised form 4 January 2019; Accepted 21 January 2019
Available online 25 January 2019
0021-9924/

to stuttering chronicity over a two-year period of early childhood development.

1.1. Trajectories of stuttering development

The onset of developmental stuttering typically occurs between 2 to 5 years of age, with 5 to 8% of preschool children affected (e.g., Felsenfeld et al., 2000; Månsson, 2000). Approximately 75% of these children eventually drop below diagnostic criteria for stuttering (for review, see Yairi & Ambrose, 2013). Longitudinal studies to identify differences between preschool children who persist and children who recover (Yairi, Ambrose, Paden, & Watkins, 2005; Spencer & Weber-Fox, 2014) represent a significant advance in stuttering research with the potential for considerable impact on clinical practice. Risk factors for stuttering persistence, such as a family history of persistent stuttering and male gender, have been identified from longitudinal studies (e.g., Ambrose, Cox, & Yairi, 1997; Månsson, 2000). Accurately identifying the likelihood a child will recover or persist and why is an integral piece of understanding the onset and development of stuttering in young children.

Longitudinal studies to date have primarily compared the performance of children who persist and children who recover on single measures within domains believed to be involved with the etiology, occurrence and development of stuttering (e.g., PPVT scores; Ambrose, Yairi, Loucks, Seery, & Throneburg, 2015; Ryan, 2001; and TACL scores; Spencer & Weber-Fox, 2014). Cognitive, speech-language, emotional, genetic, and motoric domains are often investigated (for an example of a multifactorial model of stuttering, see Smith & Weber, 2017). Longitudinal studies have established that speech-language abilities (e.g., articulation) and nonlinguistic characteristics (e.g., gender, time since onset, temperament, family history) may be associated with developmental trajectories of childhood stuttering (Ambrose et al., 1997, 2015; Paden & Yairi, 1996; Paden, Yairi, & Ambrose, 1999; Ryan, 2001; Spencer & Weber-Fox, 2014; Watkins, Yairi, & Ambrose, 1999; Yairi & Ambrose, 1992, 1999). Presently there is a gap in our understanding of the joint role of multiple domains, such as vocabulary and temperament, in stuttering chronicity.

1.2. Vocabulary and developmental stuttering

Measures of receptive and expressive vocabulary reflect a child's semantic knowledge. Receptive vocabulary refers to the words a child understands. It is commonly evaluated using norm-referenced tests, such as the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007). The PPVT requires the child to demonstrate understanding of a word by selecting one of four pictures that best matches the word's meaning. Expressive vocabulary refers to the words a child is able to verbally label. Expressive vocabulary can be assessed using norm-referenced tests, such as the Expressive Vocabulary Test (EVT; Williams, 2007). The EVT requires the child to demonstrate accurate word use by labeling pictures or providing synonyms.

1.2.1. Vocabulary performance of children who do and do not stutter

Many investigations have focused on whether children who stutter differ from children who do not stutter in their performance on measures of receptive and expressive vocabulary. For example, Ntourou and colleagues' (2011) meta-analytic review included studies of children up to eight years old. Children who stutter performed within normal limits on receptive and expressive vocabulary measures, but approximately one-half standard deviation below children who do not stutter. The researchers suggested that one possible explanation for this lower performance stems from temperamental differences between children who do and do not stutter. For instance, they speculated that children who stutter may perform slightly worse on linguistic measures because they have more difficulty adapting to the testing environment than children who do not stutter.

1.2.2. Vocabulary performance of children who persist and recover

Between group differences comparing children who persist and recover on vocabulary measures have been inconsistent. Some studies have shown no significant difference in receptive vocabulary (e.g., Chow & Chang, 2017; Kloth, Kraaimaat, Janssen, & Brutten, 1999; Ryan, 2001) or expressive vocabulary (e.g., Chow & Chang, 2017; Kloth et al., 1999), but Ambrose et al. (2015) found that children who persist performed significantly below children who recover on measures of receptive and expressive vocabulary at different points of their longitudinal study. Differences in receptive vocabulary were found when children were 2–3 years post onset, whereas differences in expressive vocabulary occurred within a year of onset. Although these findings do not indicate that vocabulary performance is a stable indicator of risk for stuttering persistence, Ambrose et al. (2015) suggested that they provide evidence for continued use of language tests in the assessment of childhood stuttering. Specifically, the authors argued that the study's relatively consistent findings that persisting children are behind their peers in linguistic skills, indicate that standardized language tests uncover subtle group differences that may be relevant to understanding trajectories of stuttering persistence and recovery. Temperament is one factor that has been speculated to impact the speech-language abilities of children who stutter (e.g., Ntourou et al., 2011) and may relate to between group differences in these domains.

1.3. Temperament and stuttering

Temperament has been defined using multiple approaches. Approaches, in general, speculate that temperament characteristics can provide insights into a child's behavior. For example, Rothbart, Ahadi, Hershey, and Fisher, (2001) defined it as "an individual's constitutionally-based proclivity for emotional reactivity and regulation" (p. 22). Whereas temperament characteristics are strongly correlated from childhood to adulthood, they develop across life and are sensitive to environmental influences (Neppl et al., 2010). Temperament has been investigated as an emotional contributor to stuttering (for review, see Jones, Choi et al., 2014) using a variety

of methods: (a) caregiver reports (e.g., Anderson, Pellowski, Conture, & Kelly, 2003; Eggers, De Nil, & Van den Bergh, 2010; Karrass et al., 2006), (b) behavioral observations (e.g., Jones, Buhr et al., 2014; Choi, Conture, Walden, Lambert, & Tumanova, 2013), and (c) physiological measures (e.g., Jones, Buhr et al., 2014, 2017; Zengin-Bolatkale, Conture, & Walden, 2015). Caregiver reports enable parents to contribute a "long-term" and cross-situational perspective on a child's temperament. Behavioral observations and physiological measures, on the other hand, provide a "snap shot" of the child in a particular situation.

Two frequently used caregiver report measures in investigations of childhood stuttering and temperament include the Behavioral Style Questionnaire (BSQ; McDevitt & Carey, 1978) and the Children's Behavior Questionnaire (CBQ: Rothbart et al., 2001). The BSQ characterizes temperament on nine dimensions (i.e. activity level, rhythmicity, approach-withdrawal, adaptability, intensity of reactions, quality of mood, attention-span, distractibility and threshold of responsiveness). The CBQ investigates temperament using 15 subscales (i.e. positive anticipation, high intensity pleasure, smiling/laughter, activity level, impulsivity, shyness, discomfort, fear, anger/frustration, sadness, soothability, inhibitory control, attentional focusing, low intensity pleasure, and perceptual sensitivity). These subscales form three composite scores: effortful control, negative affectivity, and extraversion/surgency. Effortful control reflects the child's ability to regulate emotions and attention. Negative affectivity refers to the tendency to exhibit negative emotions, such as discomfort, fear anger/frustration, and sadness. Finally, extraversion/surgency refers to the tendency to exhibit positive emotions, such as positive anticipation, low intensity pleasure, and smiling/laughter.

1.3.1. Temperament in children who do and do not stutter

Previous studies using caregiver report measures indicated that young children who stutter are less adaptable (Anderson et al., 2003; Embrechts, Ebben, Franke, & van de Poel, 2000; Howell et al., 2004), more negative in mood (Howell et al., 2004), less rhythmic (Anderson et al., 2003), more emotionally reactive (Eggers et al., 2010; Embrechts et al., 2000; Karrass et al., 2006), and less able to regulate their emotions and attention (Eggers et al., 2010; Embrechts et al., 2000; Felsenfeld, van Beijsterveldt, & Boomsma, 2010; Karrass et al., 2006) than children who do not stutter. Empirical studies employing behavioral and physiological measures have supported such conclusions (e.g., Johnson, Walden, Conture, & Karrass, 2010; Piispala, Kallio, Bloigu, & Jansson-Verkasalo, 2016; Schwenk, Conture, & Walden, 2007).

1.3.2. Temperament in children who persist and recover

To date, one longitudinal study on developmental stuttering has investigated temperament. Parent responses on the CBQ indicated that the persisting group exhibited greater negative affectivity than the recovered group (Ambrose et al., 2015). Ambrose et al. (2015) concluded that children who persist exhibit fundamental temperament differences than children who recover and children who do not stutter. However, the nature of this difference in relation to stuttering remains unclear; for instance, it is unknown whether temperament influences the development of stuttering and/or if the experience of stuttering impacts temperament.

1.3.3. Temperament and language development

Similarly, it has both been speculated that perhaps temperament impacts vocabulary development and that vocabulary development could also influence temperament. Based on previous evidence that temperament predicts later vocabulary development in children as well as the relative stability of temperament (McClelland et al., 2007), we speculate that temperament likely impacts vocabulary development. Temperament may impact language development either by influencing a child's acquisition of vocabulary or performance during assessments. In line with theories that temperament may impact language acquisition, Bloom, Beckwith, and Capatides, (1988) suspected that "processes related to emotional expression compete for cognitive resources required for language learning" (p. 170). Further, Bates suggested that a child's emotionality may impact the child's communication and opportunities for learning with the caregiver. In other words, a child with more negative emotionality may be provided fewer opportunities for learning than a child with a more moderate or positive temperament. Similarly, a child who is better able to regulate emotions may also receive more opportunities for learning than a child less able to regulate emotions. In contrast, Slomkowski, Nelson, Dunn, and Plomin, (1992) suggested children less able to regulate emotions may exhibit lower language due to testing performance. For instance, a child may perform poorly because he is unable to sustain attention during the test. No evidence to date indicates one theory (i.e., temperament impacts acquisition vs. performance) is more plausible than the other, but there is ample evidence to support a correlation between language outcomes and temperament (e.g. Bloom et al., 1988; Dixon & Smith, 2000; Salley & Dixon, 2007).

Empirical studies on the association between language and temperament have indicated that greater positive and negative emotionality is associated with delayed receptive vocabulary in early childhood (Bloom et al., 1988; Coplan, Barber, & Lagacé-Séguin, 1999; Noel, Peterson, & Jesso, 2008; Salley & Dixon, 2007). In contrast, greater emotionality has been found to be associated with higher *expressive* vocabulary in toddlers (Bloom & Beckwith, 1989) and three-year olds (Slomkowski et al., 1992). In addition, emotional regulation measures, such as better effortful control, inhibitory control and attention span relate to better receptive and expressive vocabulary scores (Dixon & Smith, 2000; Dixon, Salley, & Clements, 2006; Slomkowski et al., 1992; Wolfe & Bell, 2003).

Further, evidence from Clark, Conture, Walden, and Lambert, (2015) suggests that the relation between temperament and language outcomes may be differentially impacted by stuttering status. The authors investigated distractibility in preschoolers who not only stutter but who also exhibit receptive-expressive language dissociations, which the authors defined as "imbalances between subcomponents of speech-language planning and production" (p. 481). For children who stutter, but not children who do not, lower distractibility was associated with more speech-language dissociations (Clark et al., 2015). This cross-sectional finding suggests that temperament and speech-language may have joint effects and impact the development of children who stutter and those who do not differentially. Despite this promising relation between stuttering, temperament and language, no longitudinal investigations have considered whether language development may be impacted by *both* a child's temperament and also by the child's stuttering

status–whether the child stutters and recovers, stutters and persists, or does not stutter.

1.4. Present study

The purpose of this study was to extend previous findings on linguistic and temperamental associates of stuttering persistence by considering the joint contributions of vocabulary and temperament. As a preliminary step, this study investigates whether the vocabulary of children who persist is differentially associated with temperament compared to children who recover and children who do not stutter. Although temperament is considered relatively stable, children's temperament characteristics, along with vocabulary abilities, were repeatedly assessed across two years of development, to assess differential development in these domains based on talker group. Specifically, this study assessed between-group differences based on the association between (a) vocabulary and emotional regulation and (b) vocabulary and negative emotional reactivity.

For our first hypothesis we predicted that higher receptive vocabulary would be (a) positively associated with emotional regulation and (b) negatively associated with negative emotionality. For our second hypothesis we predicted that higher expressive vocabulary would be (a) positively associated with emotional regulation and (b) negatively associated with negative emotionality. For both hypotheses, we predicted that the associations of children who persist would be stronger than the associations exhibited by children who recover and those who do not stutter. Further, we predicted that the disparities between the strength of the associations of children who persist and the associations of children who recover and those who do not stutter would become greater over the course of the study (e.g., as children who persist continue to stutter and as children who recover discontinue stuttering).

2. Methods

Vanderbilt University's Institutional Review Board approved the study protocol. Informed consent by parents and assent by children were obtained. Data were collected as part of a large-scale, longitudinal investigation of linguistic and emotional contributions to childhood stuttering at Vanderbilt University (Clark et al., 2015; Jones, Buhr et al., 2014; Zengin-Bolatkale et al., 2015). This is the first publication to report longitudinal CBQ and vocabulary data from these participants. The research questions were developed retrospectively.

2.1. Participants

Participants were informed of the study through: (1) advertisements in a free, monthly parent-oriented magazine available in Middle Tennessee, (2) referrals from the Vanderbilt Bill Wilkerson Center and other health professionals, and (3) self-referrals. They voluntarily participated and received monetary compensation. All participants were naïve to the purposes and design of the study. They had no known hearing, neurological, developmental, academic, intellectual or emotional problems, or speech and/or language disorders other than stuttering.

Children were eligible to participate from the large-scale study cohort of 206 children if they met specific inclusion criteria. Fig. 1 depicts inclusion/exclusion criteria of participants. Based on the number of children who did not participate for at least 24 months due to volitional reasons (e.g., moved, could not be contacted, and asked to exit study), there was an attrition rate of 32%. Ultimately, 98 monolingual, Standard American English-speaking children, aged 3;0 to 4;11, were eligible to participate in the present study.

2.1.1. Children who stutter

At the initial visit, children participated in a diagnostic visit. Children were classified as stuttering if they (a) exhibited at least 3 stuttered disfluencies per 100 words in a 300-word conversational play sample and (b) scored at least an 11 on the Stuttering Severity Instrument (SSI; Riley, 1994, 2009).[1] The SSI was used to provide additional information regarding stuttering behavior (i.e., physical concomitants and duration of stuttered disfluencies) to ensure that children who had low stuttering frequency (almost precisely 3% stuttering), but other salient features and characteristics of stuttering (i.e., increased duration, tension, etc.; Sawyer & Yairi, 2001), were not misidentified. The threshold of 11 on the SSI has previously been used to categorize children who do and do not stutter (e.g., Anderson, Pellowski, & Conture, 2005; Tumanova, Conture, Lambert, & Walden, 2014). Based on these criteria, 40 children met criteria for stuttering at the initial visit.

The children who stutter were further divided into two groups based on their speech samples collected across the two-year duration of the study: a persisting group (n = 10; 9 M) and a recovered group (n = 26; 21 M). Children were classified as ***recovered*** if (a) they exhibited below 3 stuttered disfluencies per 100 words in two play samples spaced 1–2 weeks apart approximately two years after study entry, (b) they scored 10 or below on the SSI and (c) there was no parental concern of continued stuttering at his/her final visit. Children were classified as ***persisting*** if they continued to exhibit (a) at least 3 stuttered disfluencies per 100 words in a play sample and (b) a score of at least 11 on the SSI at each visit. Based on these criteria, three children who stutter were unclassifiable

[1] The Stuttering Severity Instrument is a diagnostic tool for stuttering based on frequency of syllables stuttered during speech samples, duration of stuttered disfluencies, and concomitant behaviors. The lowest severity category, very mild, corresponds to ratings of 10 or below. To minimize overlap between CWNS and CWS in this study, children were unclassifiable if their categorization based on stuttering frequency and SSI score were incongruent. The 3rd and 4th editions of the SSI were both used during the course of the longitudinal study, but do not affect the identification criteria, as they measure stuttering severity using the exact same methodology for non-readers in this age group.

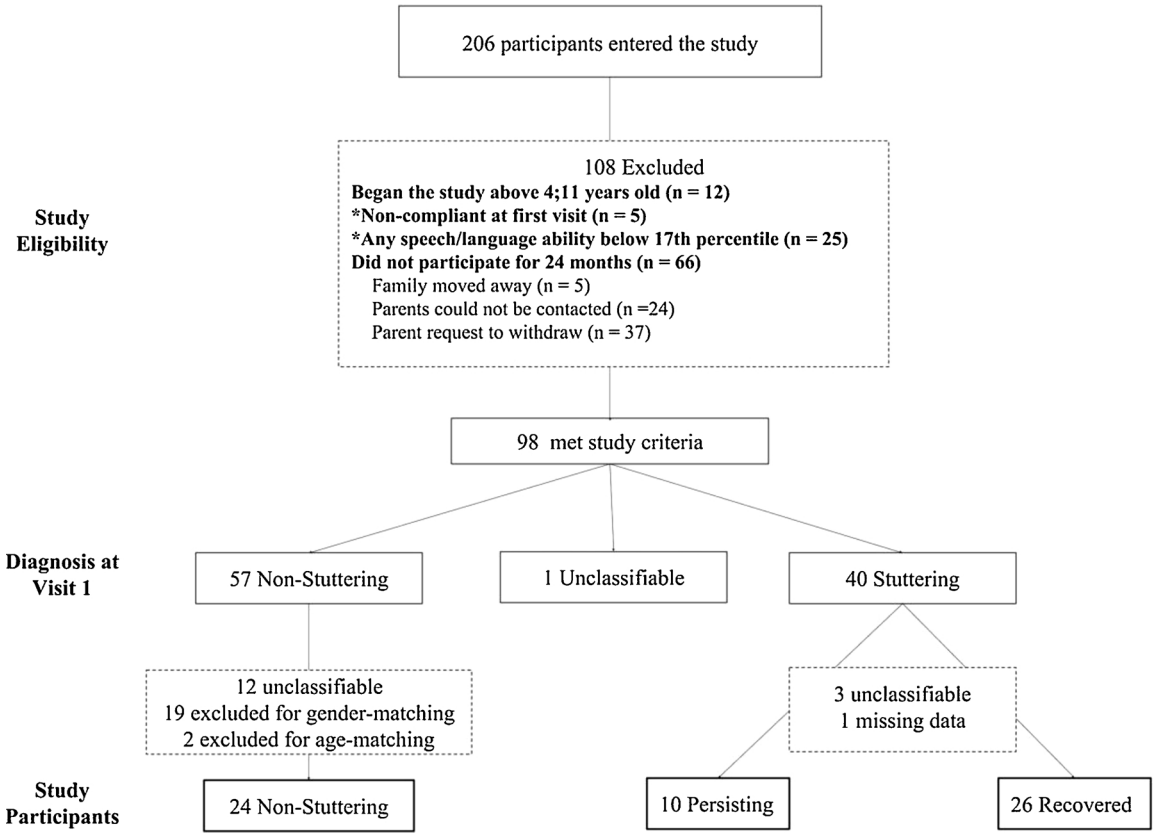

Fig. 1. The flow chart depicts study participation including exclusion criteria and final participant classification. The dashed boxes report reason's participants were excluded from the final participant groups.

(i.e., did not meet both criteria [stuttering frequency and SSI score] to be classified as recovered or persisting). One child was excluded for not having complete temperament assessments. Per parent report, 23% of the recovered group (n = 6) and 20% of the persisting group (n = 2) received treatment for stuttering at some point during the study.

2.1.2. Children who do not stutter

Children were classified as non-stuttering if they (a) exhibited below 3 stuttered disfluencies per 100 words in a play sample and (b) scored 10 or below on the SSI. Children who do not stutter were also administered the SSI to rule out the possibility that a participant may produce fewer than 3 stuttered disfluencies per 100 words, but exhibit increased tension and/or physical concomitants associated with any stuttered disfluencies. Fifty-seven children met criteria for classification as non-stuttering at their first visit. Twelve children were excluded for either exhibiting elevated disfluent speech or SSI scores at a subsequent second or third visit before falling below the diagnostic criteria for stuttering again by their final visit. Based on the young age of participants at the initial visit and time between visits, emergence of stuttering at the second or third visit is unsurprising as later onset would still fall within developmental norms (e.g., Yairi & Ambrose, 1999). Due to the high number of girls classified as non-stuttering (n = 25), an age-matched control was identified and randomly selected for each girl classified as persisting and recovered (n = 6) and the rest were excluded (n = 19). Two males were excluded for not age-matching with a persisting or recovered child (i.e., there was no persisting or recovered child who was within 4 months of age). The non-stuttering group (n = 24) was thus selected to control for age and gender with the participants classified as persisting and recovered.

2.2. Procedure

Participants and their parents visited Vanderbilt University every 7–10 months (when available) spanning a period of 24–32 months. At each visit, testing was completed in a clinical environment. One examiner engaged the child in play and collected a 300-word conversational speech sample before administering a standardized test battery (i.e. PPVT-4, EVT-2, Goldman Fristoe Test of Articulation –Second Edition [Goldman & Fristoe, 2000] and the Test of Early Language Development –Third Edition [Hresko, Reid, & Hammill, 1999]). In an observation room, a second examiner conducted the parent interview. Parents completed the CBQ ahead of every visit. Socioeconomic status data were collected during the parent interviews and calculated using the Hollingshead Four-Factor Index of Social Positions (Hollingshead, 1975). This measure ranges from 8 to 66, with higher scores reflecting higher levels of

occupation and education.

2.3. Measures

2.3.1. Vocabulary measures

Participants' receptive and expressive vocabularies were assessed using the PPVT-4 and EVT-2, respectively, at every visit. For both tests, higher standard scores indicate better skills relative to age. The PPVT-4 and EVT-2 have internal consistencies of $\alpha = .95$ and test-retest reliabilities of $\alpha = .92$ and $\alpha = .84$, respectively.

2.3.2. Emotion measures

Temperamental characteristics were evaluated at every visit using the CBQ, as have other studies investigating temperament in children who stutter (e.g. Ambrose et al., 2015; Eggers et al., 2010). The questionnaire is used to assess parent report of temperament in 3- to 7-year-old children. Parents were asked to rate whether each statement (i.e., item) was true of their child using a 7-point Likert scale, rated from 1 = "extremely untrue of your child" to 7 = "extremely true."

Over the course of the study two forms of the CBQ were administered: a full 237-item form and a shorter, 70-item form. Shorter forms were given to 41% of parents at the first visit and 30% of parents at the second visit (see Appendix A for included items). The use of these forms was spread across the three talker groups (i.e., 40% of the persisting children, 30% of the recovered children, and 40% of the non-stuttering children). The effortful control composite is based on 47 items in the full version and 22 items in the shorter form. Calculations of this scale based on both versions are highly correlated, $r\,(476) = .872$. The negative emotionality scale is based on 62 items in the full version and 23 items in the short version. Calculations of this scale based on both forms are also highly correlated, $r\,(476) = .872$. However, calculated measures of positive emotionality based on the full and short forms are much lower, $r\,(468) = .417$. Based on these correlations, the present study measured emotional regulation and negative emotionality using the effortful control and negative reactivity scales, respectively and surgency was excluded. Table 1 provides example questions for effortful control and negative reactivity. Rothbart et al. (2001) reported that the correlations between the two scales was close to zero (p. 1400).

2.4. Data analysis

2.4.1. Descriptive analyses

Descriptive analyses were included to provide a description of group characteristics. To assess possible between-group differences in age, SES, speech disfluencies, SSI scores, time since stuttering onset, vocabulary scores, and temperament scores, one-way ANOVAs were conducted. When there was a main effect of talker group, post-hoc contrasts of the differential effects of talker group were conducted, and coefficients and p-values are reported. To assess between-group differences in gender and four levels of race: Asian, Black, White, and Multi-racial, chi-square analyses were conducted.

2.4.2. Statistical modeling of primary hypothesis

Repeated measure linear mixed-effects models were conducted to investigate the hypotheses (LMM; Pinheiro & Bates, 2010). As described by Jones, Buhr et al. (2014), LMMs were selected to better control for autoregressive errors that occur with repeated measures of an individual (Nich & Carroll, 1997). Our statistical approach was employed for the precise assessment of between-group differences regarding the association of temperament and vocabulary across the developmental time period of the study. The specific parameters of two separate models, one for each vocabulary type (i.e., receptive and expressive) are described below.

The LMMs were run using Statistical Package for the Social Sciences Statistics version 24 (SPSS Statistics). The vocabulary measure (i.e., PPVT-4 or EVT-2), measured at each of four time points was the dependent variable. Fixed factors for each model included talker group (3: persisting, recovered, and non-stuttering), visit number (4: visit 1, visit 2, visit 3, and visit 4), and gender (2: male and female). Temperament (i.e., emotional regulation and negative emotionality) measured at each of four time points was a continuous predictor variable. Gender, age and SES were included as fixed effects to explicitly model their known impact on vocabulary skills (e.g., Hart & Risley, 1995; Hoff & Tian, 2005; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). For each term in the

Table 1
Examples of items from subscales of the Children's Behavior Questionnaire.

Effortful Control Composite	
Attention Focusing	"When drawing or coloring in a book, shows strong concentration"
Inhibitory Control	"Can easily stop an activity when s/he is told 'no'"
Low-Intensity Pleasure	"Enjoys looking at picture books"
Perceptual Sensitivity	"Is quickly aware of some new item in the living room"
Negative Affectivity Composite	
Anger/Frustration	"Becomes easily frustrated when tired"
Discomfort	"Is quite upset by a little cut or bruise"
Fear	"Is very frightened by nightmares"
Sadness	"Sometimes appears downcast for no reason"
Soothability	"Calms down quickly following an exciting event"

Table 2
Demographic, vocabulary, and temperament, and speech fluency characteristics of the Non-stuttering group, Recovered group, and the Persisting group.

Variable	Non-Stuttering (n = 24) M (SD)	Recovered (n = 26) M (SD)	Persisting (n = 10) M (SD)	*F*(df)	Wald χ^2(df)	*p*	Effect size[a]
Demographic Characteristics at Initial Visit							
Chronological age (mos.)	45.31 (6.7)	43.2 (5.4)	46.6 (4.5)	1.57 (2, 59)		.216	.052
Gender	18 boys	21 boys	9 boys		.568 (2)	.753	.097
Race					9.20 (6)	.163	.389
SES	46.8 (10.8)	44.9 (11.8)[c]	45.2 (10.4)	.20 (2, 58)		.822	.007
Vocabulary Skills at Initial Visit							
EVT	120.3 (9.0)	116.0 (9.8)	117.4 (10.8)	1.25 (2, 59)		.295	.042
PPVT	119.4 (11.5)	114.8 (12.4)	121.1 (14.7)	1.29 (2, 59)		.284	.043
Temperament Scores[b] at Initial Visit							
Emotional Regulation	4.9 (0.7)	4.6 (0.6)	4.6 (0.8)	2.02 (2, 59)		.143	.066
Negative Emotionality	3.7 (0.7)	3.9 (0.5)	3.8 (0.6)	.90 (2, 59)		.412	.031
Speech Fluency Measures at Initial Visit							
Stuttering Frequency (%)	1.3 (0.8)[c]	8.0 (3.7)	9.9 (6.5)	28.3 (2, 58)		< .001	.503
SSI total score	6.9 (1.0)[c]	18.8 (5.8)	19.3 (6.8)	43.5 (2, 57)		< .001	.613
Time since onset (mos.)	–	8.6 (5.1)	11.2 (5.3)	1.73 (1, 30)		.198	.056
Speech Fluency Measures at Final Visit							
Stuttering Frequency (%)	.88 (.7)	1.1 (.7)	4.8 (2.2)	46.6 (2, 57)		< .001	.629
SSI total score	5.4 (2.4)	6.5 (1.1)	14.1 (3.2)	61.7 (2, 58)		< .001	.692

Note. SES – socioeconomic status, SSI = Stuttering Severity Instrument; EVT = Expressive Vocabulary Test; PPVT = Peabody Picture Vocabulary Test.

[a] Effect sizes for gender and race are measured using d; all other effect sizes are measured using η_p^2.

[b] Temperament was measured using the Effortful Control and Negative Affectivity subscales of the Children's Behavior Questionnaire.

[c] Data was missing from one participant.

model an effect size, Cohen's d, was calculated using the values of the F-statistic and degrees of freedom from the mixed models as described by Rosenthal and DiMatteo (2001). The main effect of Temperament, the Temperament x Talker Group interaction, and the Temperament x Talker Group x Time interaction were the primary foci of the analyses to assess the study's hypotheses.

3. Results

3.1. Descriptive analyses

Table 2 shows participants' demographic, speech fluency, vocabulary, and temperament characteristics at the initial visit. Between-group differences were analyzed using ANOVAs and chi-square tests.

When comparing the stuttering characteristics of the three talker groups at the initial visit, there was a main effect of talker group for both stuttering frequency ($p < .001$, $\eta_p^2 = .503$) and total SSI scores ($p < .001$, $\eta_p^2 = .613$). Post hoc analyses were conducted to follow-up the main effect of talker group for both stuttering frequency and total SSI scores at the initial visit. These analyses indicated that the non-stuttering group exhibited fewer disfluencies than both the persisting ($B = -8.604$, $p < 0.001$) and recovered ($B = -6.654$, $p < .001$) groups as expected. No significant difference in stuttering frequency was detected between the persisting and recovered groups ($B = 1.950$, $p = .156$). Similarly, the non-stuttering group exhibited lower SSI scores than both the persisting ($B = -11.429$, $p < .001$) and recovered ($B = -10.491$, $p < .001$) groups. No significant difference in total SSI score was detected between the persisting and recovered groups ($B = 0.938$, $p = 0.708$). Last, there was no between-group difference in time since onset of stuttering for the recovered group compared to the persisting group ($p = .118$, $\eta_p^2 = .056$).

When comparing the stuttering characteristics of the three talker groups at the final visit, there was a main effect of talker group for both stuttering frequency ($p < .001$, $\eta_p^2 = .629$) and total SSI scores ($p < .001$, $\eta_p^2 = .692$). Post hoc analyses were conducted to follow-up the main effect of talker group for both stuttering frequency and total SSI scores at the final visit. These analyses indicated that the persisting group exhibited significantly higher stuttering frequency than the non-stuttering ($B = 3.883$, $p < .001$) and recovered ($B = 3.673$, $p < .001$) groups. There were no between-group differences in stuttering frequency between the recovered and non-stuttering groups ($B = .209$, $p = .523$). The persisting group also exhibited higher total SSI scores than the non-stuttering ($B = 8.683$, $p < .001$) and recovered ($B = 7.600$, $p < .001$) groups. Further, there were no differences in total SSI scores between the recovered and non-stuttering groups, ($B = 1.083$, $p = .084$).

The three groups were found to be equivalent in gender, age and SES at the initial visit. Due to balancing gender ratios when selecting participants, there were no between-group differences for gender ($p = .753$, $d = 0.097$). No statistically significant differences in race were found for the non-stuttering group (4.2% Asian, 4.2% Black, 8.3% Multi-racial, and 83.3% White), recovered group (3.7% Black, 11.1% Multi-racial, and 85.2% White), and persisting group (30% Black, 10% Multi-racial, and 60% White; $p = 0.163$, $d = 0.389$). Similarly, there was no main effect of talker group for SES ($p = .822$, $\eta_p^2 = .007$). Despite non-significant between-group differences in SES and gender, both were included in the statistical analyses for the main hypotheses to account for their

Table 3
Mixed Model Results.

	F (df)	*p*	*d*
Hypothesis 1: Receptive Vocabulary and Temperament			
Talker Group	**5.67 (2, 157)**	**.004**	**.380**
Time	0.45 (3, 90)	.715	.141
Emotional Regulation	0.01 (1, 72)	.932	.228
Negative Emotionality	**5.24 (1, 113)**	**.024**	**.431**
SES	**8.28 (1, 164)**	**.005**	**.449**
Age	**4.82 (1, 154)**	**.030**	**.354**
Gender	**13.02 (1, 160)**	**< .001**	**.571**
Talker Group x Time	1.14 (6, 99)	.223	.215
Emotional Regulation x Talker Group	0.87 (2, 93)	.422	.193
Negative Emotionality x Talker Group	**3.92 (2, 132)**	**.022**	**.345**
Emotional Regulation x Time	1.49 (3, 80)	.223	.273
Negative Emotionality x Time	0.20 (3, 89)	.898	.095
Emotional Regulation x Talker Group x Time	0.78 (6, 87)	.585	.193
Negative Emotionality x Talker Group x Time	0.63 (6, 97)	.708	.161
Hypothesis 2: Expressive Vocabulary and Temperament			
Talker Group	0.14 (2, 147)	.870	.062
Time	0.30 (3, 95)	.896	.112
Emotional Regulation	1.69 (1, 79)	.197	.293
Negative Emotionality	0.28 (1, 120)	.599	.097
SES	0.36 (1, 153)	.143	.097
Age	**6.45 (1, 153)**	**.012**	**.411**
Gender	2.17 (1, 166)	.550	.229
Talker Group x Time	0.94 (6, 99)	.470	.139
Emotional Regulation x Talker Group	0.18 (2, 98)	.840	.086
Negative Emotionality x Talker Group	0.07 (2, 134)	.935	.046
Emotional Regulation x Time	0.41 (3, 85)	.744	.139
Negative Emotionality x Time	0.37 (3, 89)	.776	.129
Emotional Regulation x Talker Group x Time	0.86 (6, 93)	.530	.192
Negative Emotionality x Talker Group x Time	0.70 (6, 97)	.658	.170

Note. Bolded values are significant at $p < .05$.

effects on vocabulary.

Further, the groups were found to exhibit equivalent vocabulary skills and temperament characteristics at the initial visit. There was no main effect of talker group for either receptive vocabulary ($p = .284$, $\eta_p^2 = .043$) or expressive vocabulary ($p = .295$, $\eta_p^2 = .042$) at the initial visit. Similarly, there was no main effect of talker group for either emotional regulation ($p = .143$, $\eta_p^2 = .066$) or negative emotionality ($p = .412$, $\eta_p^2 = .031$) at the initial visit.

3.2. Relation between vocabulary and temperament in childhood stuttering

Results pertaining to the two a priori hypotheses are presented below and are based on data collected across two years for each participant (i.e., vocabulary, temperament scores, age and SES from all visits were included in the repeated measures analyses). Full results from the linear mixed models, including group and related interactions, are in Table 3.

3.2.1. Receptive vocabulary

As expected there was a significant main effect of talker group on receptive vocabulary ($p = .004$, $d = 0.380$) when comparing the non-stuttering (M = 117.5, SE = 1.37), persisting (M = 115.6, SE = 2.4) and recovered (M = 115.7, SE = 1.35) groups. The main effect of Time was non-significant ($p = .715$, $d = 0.141$). As expected, there were main effects of Gender ($p < .001$, $d = 0.571$), SES (p = .005, $d = 0.449$), and Age ($p = .030$, $d = .358$). The Talker Group x Time interaction was non-significant ($p = .342$, $d = 0.215$).

3.2.1.1. Receptive vocabulary and emotional regulation (Hypothesis 1a). Contrary to predictions, we did not detect a significant main effect of Emotional Regulation on receptive vocabulary, ($p = .932$, $d = 0.228$) nor a significant Talker Group x Emotional Regulation interaction ($p = .422$, $d = 0.193$). Further, neither the Emotional Regulation x Time interaction ($p = .223$, $d = 0.273$) or the Emotional Regulation x Talker Group x Time interaction ($p = .585$, $d = 0.193$) were statistically significant.

3.2.1.2. Receptive vocabulary and negative emotionality (Hypothesis 1b). As expected, the main effect of Negative Emotionality was statistically significant ($p = .024$, d = .431), with greater Negative Emotionality associated with lower receptive vocabulary scores ($B = -23.458$, S.E. = 10.678). Also, as expected, the Negative Emotionality x Talker Group interaction was statistically significant ($p = .022$, $d = .235$); however, we did not detect statistically significant Negative Emotionality x Time ($p = .898$, $d = 0.095$) or Negative Emotionality x Talker Group x Time ($p = .708$, $d = 0.161$) interactions.

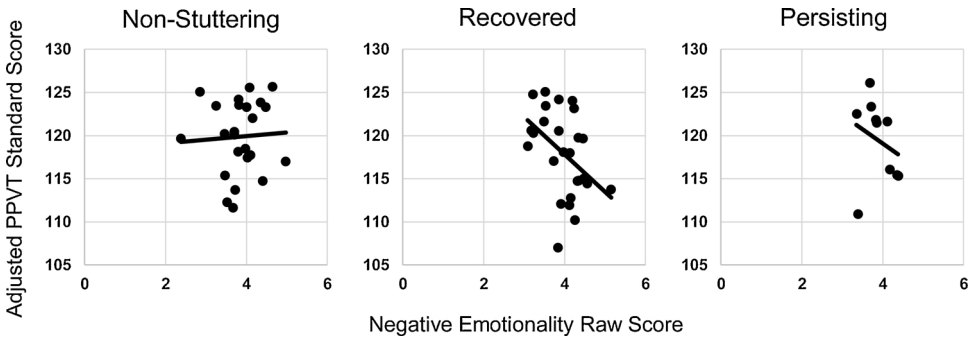

Fig. 2. The line graphs depict the association between Negative Emotionality and adjusted Peabody Picture Vocabulary Test (PPVT) standard scores for the non-stuttering (B = 0.020, p = .981), recovered (B = -3.759, p < .001), and persisting (B = -3.611, p = .018) groups. The lowest PPVT score earned by a participant was 107, so y-axis starts at 105.

Post-hoc contrasts of the differential effects of negative emotionality on receptive vocabulary among the talker groups were conducted to delineate the Negative Emotionality x Talker Group interaction on receptive vocabulary. Fig. 2 graphically depicts the varying patterns of associations within each group. There was no statistically significant association detected in the non-stuttering group (B = 0.020, p = .981) as depicted by the essentially flat regression line for that group. There was a significant negative association between negative affect and receptive vocabulary for the persisting (B = -3.611, p = .018) and the recovered (B = -3.759, p < .001) groups. Relative to the effects of negative emotionality on receptive vocabulary within the non-stuttering group, there was a statistically significantly negative effect of negative emotionality in the other two talker groups (persisting: B = -5.188, p = .008; recovered: B = -4.311, p = .001). The strength and direction of the association was not statistically significantly different for the persisting and recovered groups (B = .877, p = .658) (see Fig. 2).

3.2.2. Expressive vocabulary

We did not detect statistically significant main effects of Talker Group (p = 0.870, d = 0.062) or Time (p = .826, d = 0.112) on expressive vocabulary. There was a statistically significant main effect of Age (p = 0.012, d = 0.411), but no main effect of either Gender (p = .143, d = .571) or SES (p = .550, d = .449). The Talker Group x Time interaction was also not significant (p = .470, d = 0.215).

3.2.2.1. Expressive vocabulary and emotional regulation (Hypothesis 2a). Contrary to predictions, we did not detect a significant main effect of Emotional Regulation on expressive vocabulary (p = .197, d = 0.293) nor a significant Talker Group x Emotional Regulation interaction (p = .840, d = 0.086). Neither the Emotional Regulation x Time interaction (p = .744, d = 0.139) nor the Emotional Regulation x Talker Group x Time interaction (p = .530, d = 0.086) were statistically significant.

3.2.2.2. Expressive vocabulary and negative emotionality (Hypothesis 2b). We did not detect a significant main effect of Negative Emotionality on expressive vocabulary either (p = .599, d = 0.097). The Negative Emotionality x Talker Group interaction was also non-significant (p = .935, d = 0.046). Lastly, we did not detect that either the Negative Emotionality x Time interaction (p = .776, d = 0.129) or the Negative Emotionality x Talker Group x Time interaction were significant (p = .658, d = 0.170).

4. Discussion

4.1. Negative emotionality and receptive vocabulary

The first main finding was a negative association between negative emotionality and receptive vocabulary for children who persisted and children who recovered. This finding is consistent with investigations of typically developing children, which found that higher emotional reactivity is related to lower vocabulary performance (e.g., Noel et al., 2008; Salley & Dixon, 2007). Salley and Dixon (2007) suggested that:

> …when children are very high in negative affectivity, a relative greater burden is placed on their behavioral control systems (i.e., fundamental tendencies to approach or avoid stimuli), which must regulate this negative affectivity. The end result is fewer resources available for linguistically relevant activities, such as paying attention to word-referent associations when learning novel labels. (pp. 131–132)

Similarly, Conture and Walden (2012) speculated that, when stressed, more reactive children "retreat from or refuse to engage with the listener…" (pp. 115). Conture, Kelly, & Walden, 2013 also provided an alternative possibility that delayed language may result in more negative affect. Without identifying the directionality of this relation, we cannot assume that temperament influences vocabulary rather than vocabulary influencing temperament.

Interestingly, the second main finding indicated that the children who stutter (i.e., persisting and recovered groups) exhibited

receptive vocabulary that seems to be more sensitive to negative emotionality than children who do not stutter. Perhaps the speech-language systems, and relatedly—receptive vocabulary, of children who stutter, at least for a relatively brief period, are more sensitive to the "burden" placed on behavioral control systems by negative emotionality as described by Salley and Dixon (2007) than children who do not stutter. In addition, these findings support speculation by Ntourou et al. (2011) that lower language skills of children who stutter, compared to children who do not stutter, may be related to temperament-related differences between the groups.

4.2. Temperament and expressive vocabulary

Findings from the present study indicate that temperament is not associated with expressive vocabulary. This is consistent with Noel et al. (2008), which found that emotionality was not associated with EVT scores for children 2- to 4-years old. Temperament was associated with other aspects of expressive language, however—utterance length, use of descriptors, cohesiveness and informativeness. Perhaps we can attribute this finding to the testing measure. Salley and Dixon (2007) used parent report of child's word usage when they found an association between expressive vocabulary and temperament. Words the child produces independently may be more influenced by temperament than words the child is prompted to say during the administration of the EVT. Alternatively, as a child matures, there may be a trade-off in temperamental influences on speech-language abilities. For example, as expressive vocabulary increases, temperament may influence other aspects of expressive language. Therefore, given the age of the children in the present study, it is possible that their temperament influenced utterance length, grammatical complexity, or the words they selected rather than expressive vocabulary.

4.3. Caveats

The present study assessed temperament and vocabulary, but not the directionality of the relation. Future empirical investigations are warranted to determine which domain, if either, drives the relation. Additionally, the small number of participants, particularly the number of children classified as persisting, limited the present study's power to detect significant effects. Due to these small sample sizes, findings are preliminary in nature and need to be replicated prior to drawing firm conclusions. Future studies, particularly those with larger sample sizes, may consider physiological or behavioral measures of temperament and emotion to determine whether similar associations found between temperament and vocabulary are present. Investigating additional language measures, such as MLU, will also further our understanding of the relation between temperament and language.

5. Conclusion

The present study provides evidence of a differential association between temperament and receptive vocabulary for children who do and do not stutter. The association between receptive vocabulary and negative emotionality was more negative for children who stutter than children who do not stutter. No differences were noted between children who stutter and persist and children who stutter and recovered.

Findings have both clinical and theoretical implications. From a clinical point of view, these findings further support the inclusion of measures of temperament as part of a comprehensive assessment of stuttering in young children. A child's temperament may provide further insight to a his or her vocabulary. These findings support previous speculation that children who stutter are differentially impacted by temperament (Conture & Walden, 2012) and that their speech-language system are more "vulnerable" (Ntourou et al., 2011) to temperament than children who do not stutter. Last, current results underscore that speech-language and temperament are jointly associated with stuttering, and future empirical studies should attempt to consider nuanced associations between multiple domains (e.g., speech-motor, linguistic, emotional) thought to contribute to the onset and/or developmental trajectory of childhood stuttering.

CRediT authorship contribution statement

Cara M. Singer: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Visualization. **Tedra A. Walden:** Conceptualization, Methodology, Investigation, Writing - review & editing, Visualization, Supervision, Funding acquisition. **Robin M. Jones:** Conceptualization, Methodology, Validation, Investigation, Resources, Data curation, Writing - original draft, Visualization, Supervision, Project administration, Funding acquisition.

CRediT authorship contribution statement

Cara M. Singer: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Visualization. **Tedra A. Walden:** Conceptualization, Methodology, Investigation, Writing - review & editing, Visualization, Supervision, Funding acquisition. **Robin M. Jones:** Conceptualization, Methodology, Validation, Investigation, Resources, Data curation, Writing - original draft, Visualization, Supervision, Project administration, Funding acquisition.

Acknowledgements

This research was supported in part by National Institutes of Health (NIH) grants from the National Institute on Deafness and Other Communication Disorders (NIDCD) to Vanderbilt University (5R01DC000523-19, 2R56DC000523-20A1), and Vanderbilt University Medical Center (R21DC016723), and the National Center for Research Resources, a CTSA grant (1 UL1 RR024975-01) to Vanderbilt University that is now at the National Center for Advancing Translational Sciences (UL1 TR000445-06). The research reported herein does not reflect the views of the NIH, NIDCD, or Vanderbilt University. We also extend sincere appreciation to the young children and caregivers whose participation and cooperation made this project possible.

Appendix A

Questions from the full 237-item Children's Behavior Questionnaire included in the 70-item shorter form: 4, 6, 14,16, 20, 27, 29, 32, 38, 40, 42, 44, 47, 53, 55, 61, 62, 63, 64, 68, 75, 78, 85, 92, 93, 95, 101, 103, 108, 116, 118, 125, 130, 134, 136, 144, 147, 150, 160, 162, 167, 168, 171, 177, 180, 184, 185, 186, 190, 195, 198, 200, 202, 206, 208, 210, 214, 215, 217, 220, 221, 223, 226, 227, 229, 230, 232, 233, 235, and 237

References

Ambrose, N. G., Cox, N. J., & Yairi, E. (1997). The genetic basis of persistence and recovery in stuttering. *Journal of Speech Language and Hearing Research, 40*, 567–580.
Ambrose, N. G., Yairi, E., Loucks, T. M., Seery, C. H., & Throneburg, R. (2015). Relation of motor, linguistic and temperament factors in epidemiologic subtypes of persistent and recovered stuttering: Initial findings. *Journal of Fluency Disorders, 45*, 12–26.
Anderson, J. D., Pellowski, M. W., & Conture, E. G. (2005). Childhood stuttering and dissociations across linguistic domains. *Journal of Fluency Disorders, 30*, 219–253.
Anderson, J. D., Pellowski, M. W., Conture, E. G., & Kelly, E. M. (2003). Temperamental characteristics of young children who stutter. *Journal of Speech Language and Hearing Research, 46*, 1221–1233.
Bloom, L., Beckwith, R., & Capatides, J. B. (1988). Developments in the expression of affect. *Infant Behavior & Development, 11*, 169–186.
Bloom, L., & Beckwith, R. (1989). Talking with feeling: Integrating affective and linguistic expression in early language development. *Cognition & Emotion, 3*, 313–342.
Choi, D., Conture, E. G., Walden, T. A., Lambert, W. E., & Tumanova, V. (2013). Behavioral inhibition and childhood stuttering. *Journal of Fluency Disorders, 38*, 171–183.
Chow, H. M., & Chang, S.-E. (2017). White matter developmental trajectories associated with persistence and recovery of childhood stuttering. *Human Brain Mapping, 13*, 266–315.
Clark, C. E., Conture, E. G., Walden, T. A., & Lambert, W. E. (2015). Speech-language dissociations, distractibility, and childhood stuttering. *American Journal of Speech-language Pathology, 24*, 480–503.
Conture, E. G., & Walden, T. (2012). *Dual diathesis-stressor model of stuttering. Theoretical issues of fluency disorders.* Moscow, Russia: National Book Centre94–127.
Conture, E. G., Kelly, E. M., & Walden, T. A. (2013). Temperament, speech and language: An overview. *Journal of Communication Disorders, 46*, 125–142.
Coplan, R. J., Barber, A. M., & Lagacé-Séguin, D. G. (1999). The role of child temperament as a predictor of early literacy and numeracy skills in preschoolers. *Early Childhood Research Quarterly, 14*, 537–553.
Dixon, W. E., Jr., & Smith, P. H. (2000). Links between early temperament and language acquisition. *Merrill-Palmer Quarterly, 46*, 417–440.
Dixon, W. E., Salley, B. J., & Clements, A. D. (2006). Temperament, distraction, and learning in toddlerhood. *Infant Behavior & Development, 29*, 342–357.
Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). Bloomington, MN: Pearson Assessments.
Eggers, K., De Nil, L. F., & Van den Bergh, B. R. H. (2010). Temperament dimensions in stuttering and typically developing children. *Journal of Fluency Disorders, 35*, 355–372.
Embrechts, M., Ebben, H., Franke, P., & van de Poel, C. (2000). Temperament: A comparison between children who stutter and children who do not stutter. In E. C. Healey, & H. M. F. Peters (Eds.). *Stuttering: Proceedings of the third world congress on fluency disorders* (pp. 557–561). Nijmegen, The Netherlands: University Press Nijmegen.
Felsenfeld, S., Kirk, K., Zhu, G., Statham, D., Neale, M., & Martin, N. (2000). A study of the genetic and environmental etiology of stuttering in a selected twin sample. *Behavioral Genetics, 30*, 359–366.
Felsenfeld, S., van Beijsterveldt, C. E. M., & Boomsma, D. I. (2010). Attentional regulation in young twins with probable stuttering, high nonfluency, and typical fluency. *Journal of Speech Language and Hearing Research, 53*, 1147–1166.
Goldman, R., & Fristoe, M. (2000). *Goldman-fristoe test of articulation* (2nd ed.). Circle Pines, MN: American Guidance Service, Inc.
Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H Brookes Publishing.
Hoff, E., & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of Communication Disorders, 38*, 271–278.
Hollingshead, A. (1975). *Four factor index of social status. Unpublished manuscript.* New Haven, CT: Yale University.
Howell, P., Davis, S., Patel, H., Cuniffe, P., Downing-Wilson, E., Au-Yeung, J., et al. (2004). *Fluency development and temperament in fluent children and children who stutter. theory, research and therapy in fluency disorders. Proceedings of the 4th World Congress on Fluency Disorders*250–256.
Hresko, W., Reid, D., & Hammill, D. (1999). *Test of early language development* (3rd ed.). Austin, TX: Pro-Ed, Inc.
Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236–248.
Johnson, K. N., Walden, T. A., Conture, E. G., & Karrass, J. (2010). Spontaneous regulation of emotions in preschool children who stutter: Preliminary findings. *Journal of Speech, Language, and Hearing Research : JSLHR, 53*, 1478–1495.
Jones, R. M., Walden, T. A., Conture, E. G., Erdemir, A., Lambert, W. E., & Porges, S. W. (2017). Executive functions impact the relation between respiratory sinus arrhythmia and frequency of stuttering in young children who do and do not stutter. *Journal of Speech Language and Hearing Research, 60*, 2133–2150.
Jones, R. M., Buhr, A. P., Conture, E. G., Tumanova, V., Walden, T. A., & Porges, S. W. (2014). Autonomic nervous system activity of preschool-age children who stutter. *Journal of Fluency Disorders, 41*, 12–31.
Jones, R., Choi, D., Conture, E., & Walden, T. (2014). Temperament, emotion, and childhood stuttering. *Seminars in Speech and Language, 35*, 114–131.

Karrass, J., Walden, T. A., Conture, E. G., Graham, C. G., Arnold, H. S., Hartfield, K. N., et al. (2006). Relation of emotional reactivity and regulation to childhood stuttering. *Journal of Communication Disorders, 39*, 402–423.
Kloth, S. A. M., Kraaimaat, F. W., Janssen, P., & Brutten, G. J. (1999). Persistence and remission of incipient stuttering among high-risk children. *Journal of Fluency Disorders, 24*, 253–265.
Månsson, H. (2000). Childhood stuttering: Incidence and development. *Journal of Fluency Disorders, 25*, 47–57.
McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43*, 947–959.
McDevitt, S. C., & Carey, W. B. (1978). The measurement of temperament in 3-7 year old children. *Journal of Child Psychology and Psychiatry, 19*, 245–253.
Neppl, T. K., Donnellan, M. B., Scaramella, L. V., Widaman, K. F., Spilman, S. K., Ontai, L. L., et al. (2010). Differential stability of temperament and personality from toddlerhood to middle childhood. *Journal of Research in Personality, 44*, 386–396.
Nich, C., & Carroll, K. (1997). Now you see it, now you don't: A comparison of traditional versus random-effects regression models in the analysis of longitudinal follow-up data from a clinical trial. *Journal of Consulting and Clinical Psychology, 65*, 252–261.
Noel, M., Peterson, C., & Jesso, B. (2008). The relationship of parenting stress and child temperament to language development among economically disadvantaged preschoolers. *Journal of Child Language, 35*, 823–843.
Ntourou, K., Conture, E. G., & Lipsey, M. W. (2011). Language abilities of children who stutter: A meta-analytical review. *American Journal of Speech-language Pathology, 20*, 163–179.
Paden, E. P., & Yairi, E. (1996). Phonological characteristics of children whose stuttering persisted or recovered. *Journal of Speech Language and Hearing Research, 39*, 981–990.
Paden, E. P., Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering II: Initial status of phonological abilities. *Journal of Speech Language and Hearing Research, 42*, 1113–1124.
Piispala, J., Kallio, M., Bloigu, R., & Jansson-Verkasalo, E. (2016). Delayed N2 response in Go condition in a visual Go/Nogo ERP study in children who stutter. *Journal of Fluency Disorders, 48*, 16–26.
Pinheiro, J., & Bates, D. (2010). *Mixed-effects models in S and S-PLUS.* New York, NY: Springer Science & Business Media.
Riley, G. D. (1994). *Stuttering Severity Instrument for children and adults* (3rd ed.). Austin, TX: Pro-Ed, Inc.
Riley, G. D. (2009). *Stuttering Severity Instrument for children and adults* (4th ed.). Austin, Texas: Pro-Ed, Inc.
Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*, 59–82.
Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development, 72*, 1394–1408.
Ryan, B. P. (2001). A longitudinal study of articulation, language, rate, and fluency of 22 preschool children who stutter. *Journal of Fluency Disorders, 26*, 107–127.
Salley, B. J., & Dixon, W. E. (2007). Temperamental and joint attentional predictors of language development. *Merrill-Palmer Quarterly, 53*, 131–154.
Schwenk, K. A., Conture, E. G., & Walden, T. A. (2007). Reaction to background stimulation of preschool children who do and do not stutter. *Journal of Communication Disorders, 40*, 129–141.
Slomkowski, C. L., Nelson, K., Dunn, J., & Plomin, R. (1992). Temperament and language: Relations from toddlerhood to middle childhood. *Developmental Psychology, 28*, 1090–1095.
Smith, A., & Weber, C. (2017). How stuttering develops: The multifactorial dynamic pathways theory. *Journal of Speech Language and Hearing Research, 60*, 2483–2505.
Spencer, C., & Weber-Fox, C. (2014). Preschool speech articulation and nonword repetition abilities may help predict eventual recovery or persistence of stuttering. *Journal of Fluency Disorders, 41*, 32–46.
Tumanova, V., Conture, E. G., Lambert, E. W., & Walden, T. A. (2014). Speech disfluencies of preschool-age children who do and do not stutter. *Journal of Communication Disorders, 49*, 25–41.
Watkins, R. V., Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering III. *Journal of Speech Language and Hearing Research, 42*, 1125–1211.
Williams, K. T. (2007). *Expressive vocabulary test* (2nd ed.). Circle Pines, MN: Pearson, Inc.
Wolfe, C. D., & Bell, M. A. (2003). Working memory and inhibitory control in early childhood: Contributions from physiology, temperament, and language. *Developmental Psychobiology, 44*, 68–83.
Yairi, E., & Ambrose, N. (1992). A longitudinal study of stuttering in children: A preliminary report. *Journal of Speech and Hearing Research, 35*, 755–760.
Yairi, E., & Ambrose, N. G. (1999). Early childhood stuttering I: Persistency and recovery rates. *Journal of Speech Language and Hearing Research, 42*, 1097–1112.
Yairi, E., Ambrose, N. G., Paden, E. P., & Watkins, R. V. (2005). *Early childhood stuttering for clinicians by clinicians.* Austin, TX: Pro-ed.
Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders, 38*, 66–87.
Zengin-Bolatkale, H., Conture, E. G., & Walden, T. A. (2015). Sympathetic arousal of young children who stutter during a stressful picture naming task. *Journal of Fluency Disorders, 46*, 24–40.