Journal of Communication Disorders xxx (xxxx) xxxx

Contents lists available at ScienceDirect

Journal of Communication Disorders

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

Exploring gender as a potential source of bias in adult judgments of children with specific language impairment and attention-deficit/hyperactivity disorder

Alison Shimko*, Sean Redmond, Amy Ludlow, Andrea Ash

Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, Utah, USA

ARTICLE INFO

Keywords:
Implicit bias
Gender bias
Specific language impairment
Developmental language disorder
Attention-deficit/hyperactivity disorder

ABSTRACT

The purpose of this follow-up study to Ludlow (2013) was to examine potential sources of variability within attributional ratings adults (age range: 21–73) assigned to child speakers affected by either Specific Language Impairment (SLI) or Attention-Deficit/Hyperactivity Disorder (ADHD). Factors considered were rater's gender [Rater Male (RM) or Rater Female (RF)], the reported gender of the speakers [Speaker Male (SM) or Speaker Female (SF)], and the type of neurodevelopmental disorder involved (SLI or ADHD). Eighty participants (40 male and 40 female) rated brief, transcribed, narratives previously produced in Ludlow (2013) by boys affected by either SLI, ADHD, or who had typical neurodevelopment (TN). Narratives were presented to raters as having been generated by either a boy or a girl. After reading each narrative, participants provided ratings in response to 15 questions about the narrative, the child speaker's attributes, and family background. Analyses revealed a significant main effect for speaker group, such that raters assigned more pejorative attributes to children with a disorder (ADHD = SLI < TN and ADHD < SLI < TN) across all dimensions. Significant speaker gender main effects (SM < SF) were limited to questions targeting the speaker's behavioral attributes. Results obtained in this study with transcription stimuli replicated previous reports that had used audio stimuli. These findings contribute to a growing body of research documenting the presence of robust, multidimensional, implicit, negative biases among most individuals towards children displaying language differences associated with common neurodevelopmental disorders.

1. Introduction

Specific language impairment (SLI) and attention-deficit/hyperactivity disorder (ADHD) are two common neurodevelopmental disorders that affect the communication and behavior of millions of school-age children. SLI is a subtype of the broader developmental language disorder (DLD) designation. SLI refers to pediatric cases of language disorder that occur in the absence of concomitant perceptual, cognitive, socioemotional, or other disorders (Johnson et al., 1999; Tomblin et al., 1997). Epidemiological reports indicate that SLI represents approximately 7–8% of the student population (Johnson et al., 1999; Norbury et al., 2016; Tomblin et al., 1997) – a rate that is higher than autism, stuttering, and traumatic brain injury combined. Children affected by SLI have been shown to be at increased risk for poorer social, academic, and vocational outcomes (Tomblin & Nippold, 2014).

* Corresponding author.
E-mail addresses: ashimko@uw.edu (A. Shimko), sean.redmond@hsc.utah.edu (S. Redmond), amy.ludlow@jordandistrict.org (A. Ludlow), andrea.ash@health.utah.edu (A. Ash).

https://doi.org/10.1016/j.jcomdis.2019.105910
Received 7 February 2018; Received in revised form 6 May 2019; Accepted 14 May 2019

ADHD is a behavioral disorder characterized by elevated levels of inattention, hyperactivity, and impulsivity and according to epidemiological reports affects 3%–5% of the school-age population (National Institutes of Health [NIH], 2009; Scahill & Schwab-Stone, 2000). The presence of ADHD has also been associated with significant elevations in individual social, academic, and vocational risk. Some studies suggest possible links between ADHD, SLI, and other communication disorders (For a review, see Redmond, 2016). Even though both SLI and ADHD contribute to children's communication difficulties, the two conditions appear to be associated with very different verbal symptoms. For example, Redmond (2004) collected spontaneous conversational language samples from children with SLI, children with ADHD, and children with typical neurodevelopment (TN) and found that tense marking and other grammatical deficits were only associated with samples from the SLI group. In contrast, elevated speaking rates, false starts, and revisions were primarily characteristic of the spontaneous utterances produced by children with ADHD.

1.1. Social risks of communication disorders

Part of the social risk associated with having a communication disorder is the way in which peers, parents, teachers, and other adults interpret and respond to atypical verbal behaviors. A notable body of research suggests that symptoms associated with communication disorders in general, and language disorders in particular, trigger strong negative attributional biases and stereotypes in listeners (Allard & Williams, 2008; Bebout & Arthur, 1992; DeThorne & Watkins, 2001; Overby, Carrell, & Bernthal, 2007; Rice, Hadley, & Alexander, 1993; Wenker, Wegener, & Hart, 1996). Additionally, studies indicate that capturing these negative biases does not require extensive exposures. Listeners from a variety of backgrounds readily extend their snap judgments on to a multitude of personal attributes and traits that cannot be determined from brief samples of someone's speech. For example, Allard and Williams (2008) found that when 445 undergraduate students were presented with audio samples of the same paragraph read by an adult actor portraying different communication disorders (articulation, fluency, voice, language) and a non-disordered comparison, the non-disordered condition was perceived more favorably than the disordered conditions. More specifically, across the types of communication disorders under consideration, the sample representing a language disorder was perceived the most negatively across each area evaluated (intelligence, self-esteem, decisiveness, reliability, emotional stability, social adjustment, stress, employability, and ambition). In an earlier report, Rice et al. (1993) observed similar biases in the judgements provided by undergraduate students, kindergarten teachers, and speech-language pathologists based on 90-second audio samples collected from six kindergarten children with different profiles representing speech impairments, language impairments, and typical neurodevelopment (one boy and one girl for each profile). Results indicated that the samples collected from the children with language impairments were consistently rated the most pejoratively across several attributes: intelligence, leadership abilities, social maturity, academic success, parental education, and parental social status. No significant rater effects were found. In other words, despite their training and work experience, speech-language pathologists in the Rice et al. (1993) study sample were as likely as the other raters to assign pejorative ratings to children with speech and language impairments. Potential effects of gender on the ratings provided across either the raters or the children who provided the samples were not examined.

DeThorne and Watkins (2001) replicated and extended the results of Rice et al. (1993). In addition to teachers, undergraduate students, and speech-language pathologists, a rater group of sixth-grade students was included in this study's design. Open-ended questions were also incorporated into DeThorne and Watkins' questionnaire (e.g. *please describe any characteristics of this child's talking that you noticed*). Raters were also given the option in the DeThorne and Watkins study to select an *uncertain* value. Evaluations were based on slightly longer 2.5 min samples collected on a boy with SLI, an age-matched boy with TN, and a younger boy with TN matched to the boy with SLI on the basis of mean length of utterance (MLU). Despite these methodological differences, results strongly paralleled Rice et al. (1993). All four rater groups assigned their lowest ratings to the boy with SLI. Across all groups, the majority of raters elected not to use the *uncertain* value, although more raters from the speech-language pathologist group did choose the *uncertain* option. This finding indicated awareness among a minority of speech-language pathologists that making judgments based on brief audiotaped samples was inappropriate. Responses to the open-ended questions were primarily negative for the boy with SLI and they extended to comments regarding poor vocal quality. In other words, global judgments provided by the raters appeared to incorporate speech elements other than what was coming across from children's grammatical and vocabulary symptoms. The effects of rater gender on the evaluations provided was not examined.

Some investigations have used brief narrative samples collected on children affected and unaffected by SLI as the basis for eliciting adult judgments. The use of narratives provides some control over the content of children's language samples. Newman and McGregor (2006) focused on the extent to which teachers and laypersons could detect qualitative differences between 4 min narratives produced by 10 children with SLI and 10 TN children matched to the children with SLI on the basis of age. The number of boys and girls within each group was not reported, but Newman and McGregor did state that gender was balanced between their groups (see p. 1026). To control for possible rater gender and parental status effects, all raters in the study were mothers. Significant rating differences, favoring the unaffected group, were found on the amount of information conveyed, the ease in understanding the story, and the fluidity on the part of the speakers. Similar to Rice et al. (1993) and DeThorne and Watkins (2001); Newman and McGregor (2006) found no differences between layperson and teacher ratings. Newman and McGregor (2006) interpreted their results as evidence that the functional impact of SLI could be reliably detected by untrained listeners when directed at narratives produced by these children.

Ludlow (2013) also used narrative samples and mothers in her study, but she combined the narrative quality ratings utilized in Newman and McGregor (2006) with the personal attributions ratings from Rice et al. (1993); DeThorne and Watkins (2001); Overby et al. (2007); Wenker et al. (1996), and others. Narrative samples were taken from 3 boys who had participated in Redmond, Thompson, and Goldstein (2011) representing a case of SLI, a case of ADHD, and a case of TN. The inclusion of a case who had an

independent diagnosis of ADHD and overall language abilities within then normal range (as measured by the CELF-4) allowed for an examination of the specific impact of communication characteristics associated with ADHD on adults' perceptions (e.g. false starts and revisions). To control for the content of children's narratives and potential order effects, Ludlow (2013) used the story retell subtest from the *Test of Narrative Language* (Gillam & Pearson, 2004) and a Latin square design in which subgroups of the study sample evaluated different versions of the protocol. This allowed each speaker to appear in each serial position (first, second, third). Another important adjustment in the study's design over previous studies of pediatric language impairments was the use of visual analog scales rather than Likert values to encourage raters to provide more differentiated responses (see Hasson & Arnetz, 2009). Visual analog scales (VAS) are similar to semantic differential scales in that both ask respondents to indicate the magnitude of their evaluations by providing a tick mark on a continuous line. Raters were 31 mothers of children with disabilities and 29 mothers of children without disabilities. Ludlow hypothesized that mothers of children with disabilities would be less inclined to assign pejorative ratings to the speakers with SLI and ADHD because of their experiences raising children with disabilities.

Results of Ludlow (2013) provided replication of previous studies. For example, several significant differences were found between the VAS ratings of the narrative provided by the boy with SLI and the boy with TN favoring the boy with TN across a wide range of attributes (narrative quality, academic attributes, behavioral attributes, social attributes, and family attributes). Ludlow (2013) also provided new findings. The narrative provided by the TN boy was rated more favorably than the narrative provided by the boy with ADHD. These differences were of a similar magnitude and across the same dimensions as the differences observed between the narratives provided by the boy with SLI and the boy with TN (SLI = ADHD < TN). Furthermore, there were significant differences between narrative quality ratings assigned to the narrative provided by the boy and the narrative from the boy with ADHD favoring the boy with SLI (ADHD < SLI < TN). This finding was interesting because even though the narrative provided by the boy with SLI contained more grammatical errors, raters evaluated the overall narrative quality associated with the narrative provided by the boy with ADHD more poorly. The narrative provided by the boy with ADHD contained several false starts and revisions that were not present in the other narratives suggesting these communicative differences were at least as penalizing as grammatical errors. Results, however, did not support Ludlow's central hypothesis. No significant differences were found in the ratings provided by mothers with children who had a disability and by mothers whose children did not have a disability.

In sum, research into listener reactions indicates robust negative biases. These represent the implicit perceptual default for most listeners when they encounter communication differences of the kind associated with either SLI or ADHD. Effect sizes (Cohen's *d*) associated with reported differences between children with these common neurodevelopmental disorders as compared to TN children have ranged from medium to very large (.78–3.5). Some studies indicate very little overlap between the ratings assigned to samples collected from affected and unaffected individuals. Variation across individual raters has been reported in each study but attempts to identify the source of variation, for the most part, have not been successful. Differences in rater's levels of age, professional training, work experience, and personal exposures to disabilities have not had a substantial impact on their ratings. Collectively, the research record covers more than a twenty-year span, suggesting that these biases have been durable despite changes that have occurred in pre-professional training, educational policies, awareness initiatives (e.g. anti-bullying campaigns), and service provisions. One aspect of attributional bias however, that has not been systematically considered in the research on attributional biases linked to either SLI or ADHD symptoms is the contribution of gender to these biases. This gap is interesting given the presence of gender effects reported in a parallel line of studies examining the public's general perceptions of disability.

1.2. Gender differences

Research documents the existence among able-bodied people of pervasive, robust, negative perceptions of people with a variety of physical and mental disabilities (e.g., Bogdan & Biklen, 1993; Gartner, Lipsky, & Turnbull, 1991; Nelson, 1994; Tang, Davis, Wu, & Oliver, 2000). Investigations into the effects of gender on these perceptions have identified rater effects, target effects (i.e. the person with the disability being rated, referred to as "speaker" in the current study), and interactions between raters and the targets (i.e., "speaker"). Studies focusing on rater gender effects have generally found that relative to men, women's attitudes towards people with different kinds of disabilities tend to be more favorable (Panek & Jungers, 2008; Rice, 2009; Vilchinsky, Werner, & Findler, 2010). For example, Panek and Jungers (2008) examined the effects of gender on the perceptions of individuals with intellectual disabilities. One hundred and sixteen undergraduate students read brief text descriptions of three target persons with intellectual disabilities and were asked to use semantic differential scales to rate the individual described across several attributes (e.g. good-bad, valuable-worthless, pleasant-unpleasant, and weak–strong). A Latin-square design was used to manipulate the target's gender across descriptions and control for possible order effects. Results indicated that overall female raters provided more favorable ratings than male raters. In this study, no significant differences were found between genders of the person with a disability.

In contrast, other investigations have found that women with disabilities tend to be viewed more negatively than their male counterparts (Fine & Asch, 1981; Weisel & Florian, 1990). Weisel and Florian (1990) examined perceptions of disabilities across 286 11^{th}-grade Israeli adolescents. One hundred and forty-six participants were males while 139 participants were females. Participants responded to two questionnaires describing a person with a disability (male and female). Factors that were evaluated included distressed identification, rejection of intimacy, inferred emotional consequences, authoritarian virtuousness, generalized rejection, interaction strain, and imputed functional limitation (Weisel & Florian, 1990). In three of the six factors (generalized rejection, imputed functional limitation, and rejection of intimacy), ratings towards females with disabilities were less positive than ratings towards males with disabilities. Additionally, on two factors (generalized rejection and imputed functional limitation) negative attitudes were only expressed by male students.

Vilchinsky et al. (2010) examined the interaction of rater gender effects and target gender effects. Two hundred and six males and

198 females without disabilities were asked to react to text-based scenarios involving males and females using or not using a wheelchair. Between raters, women experienced less distancing behaviors towards the target person with the physical disability. Distancing behaviors include feelings of wanting to separate from the person with a physical disability. Additionally, women had more positive beliefs and thoughts towards the target person with the physical disability. However, male raters exhibited distancing behaviors more frequently towards male targets than toward female targets. Both male and female raters had more positive beliefs towards male targets than female targets. These researchers also examined potential interactions of rater and target gender with having a disability. Observed gender effects in rater and target were smaller than the effect disability status had on targets' reported behaviors, beliefs, and emotions.

1.3. Study purpose

Studies of the attributional biases that adults readily extend to people with communication disorders have been based predominately on brief audio samples of running speech. The impact of gender within this line of research has largely gone unexamined. In contrast, studies of attributional biases associated with other disabilities have relied more heavily on brief vignettes and text descriptions. In this line of research, rater gender effects, target gender effects, and their interactions have been widely reported – although not consistently. In the current study, we sought to merge these two lines of inquiry to arrive at a more complete picture of the attributional biases associated with the linguistic symptoms of SLI and ADHD. This study expands upon the inquiry of Ludlow (2013) by using orthographic transcriptions of these narratives instead of audio samples, permitting the manipulation of the gender of the person providing the narrative whilst keeping constant the symptoms displayed, as well as the content and stylistic components of each narrative. We then asked both men and women to read and rate these transcriptions using the same questions used by Ludlow (2013). These adjustments allowed us to address five specific research questions:

1) Are the ratings collected from our participants using a text vignette of children's narratives similar to those obtained by Ludlow (2013) using the same narratives but presented in that study to participants as audio samples?
2) Do raters perceive narratives produced by children with SLI and ADHD more negatively than a child with TN?
3) Are there significant differences between the ratings provided by male raters and female raters?
4) Are there significant differences between the ratings assigned to child speakers when they are identified as either male or female?
5) Are there significant interaction effects among the speaker's clinical status (speaker group), speaker's reported gender, and the rater's gender?

We hypothesized that even with volume, rate, intonation, voice, and other speech parameters removed from our narrative samples, orthographic representations of the grammatical and utterance formulation symptoms of SLI and ADHD would be sufficient to trigger the kinds of attributional biases found in previous reports. Specifically, we predicted the same patterns found in Ludlow (2013) would emerge. We expected our raters to respond to the orthographic transcriptions with broad-based preferences for the speaker with typical neurodevelopment over the speakers with SLI or ADHD. It was not clear based on previous studies however, whether the magnitude of group differences would be comparable across audio samples and orthographic transcriptions. The elimination of speech, voice, and other potential cues in the orthographic transcriptions could either exaggerate or mitigate these effects. For example, audio samples provide raters with additional cues of potential clinical features (e.g. DeThorne & Watkins, 2001) that cannot be captured with an orthographic transcription. On the other hand, reading orthographic transcriptions might highlight some symptoms (e.g. verb inflection errors) that appear briefly in the audio sample. Based on the findings of Panek and Jungers (2008); Vilchinsky et al. (2010), and Rice (2009), we expected that on average male raters would provide more pejorative ratings than female raters of the speakers with SLI and ADHD. We also expected a main speaker gender effect such that when the speaker was identified as female they would be rated more pejoratively. Based on the results of Fine and Asch (1981), Vilchinsky et al. (2010) and Weisel and Florian (1990), we anticipated these main speaker gender effects would be moderated by the presence of an interaction effect such that when the speaker with SLI or ADHD was identified as female, the male raters, in particular, would assign more pejorative ratings.

2. Methods

Approval from the University of Utah IRB was secured prior to recruitment, consent, and testing. Participants entered into a drawing for a $100 gift card as compensation for their time.

2.1. Participants

Participants were 80 adults (40 males 40 females). All participants were at least 21 years old and, according to self-report collected at the time of the study, were not enrolled in university coursework. The exclusion of college students was deliberate and meant to encourage a more representative study sample. Average age associated with the study sample was 44.51 (*Range* = 21–73, *SD* = 14.91). Demographic characteristics of the male and female rater groups are presented in Table 1. Participants were primarily recruited through community flyers or solicited onsite in a variety of public spaces (coffee shops, airports, libraries). Most questionnaires were completed in participants' homes. Additional locations included coffee shops, airplanes, airports, and libraries. A chain-referral sampling technique was used whereby participants referred acquaintances to the researcher, who might be interested

Table 1
Demographic Information.

		Male	Female	Total
Gender				
	Male	40	–	40
	Female	–	40	40
Occupation				
	Education-Primary/Secondary (K-12)	8	13	21
	Education-College, university, adult	1	2	3
	Retired	1	4	5
	Finance and Insurance	5	0	5
	Software or Technical Service	5	0	5
	Homemaker	0	1	1
	Retail	1	1	2
	Software	0	2	2
	Government and Public Administration	1	0	1
	Health Care and Social Assistance	2	2	4
	Manufacturing-Other	1	0	1
	Hotel and Food Services	0	1	1
	Legal Services	0	1	1
	Transportation and Warehousing	1	0	1
	Agriculture, Forestry, Fishing, or Hunting	0	1	1
	Other	7	6	13
	Multiple	7	6	13
Age				
	Mean	47.05	42.10	44.51
	Missing	0	2	2
Marital Status				
	Married	29	26	55
	Single	11	7	18
	Divorced	0	7	7
Number of Children				
	Total Children	52	64	50
	Mean number of children/person	1.3	1.6	1.45
	Missing	0	1	1
Ethnicity				
	Hispanic	4	2	6
	Not Hispanic	36	38	74
Race				
	White	36	37	73
	Asian	0	1	1
	Black	0	1	1
	American Indian	1	0	1
	Mixed	2	0	2
	Missing	1	1	2
Education Level				
	High school	2	2	4
	1 or more years of college, no degree	3	4	7
	Associate Degree	3	3	6
	Bachelor's Degree	17	17	34
	Master's Degree	13	10	23
	Professional Degree	1	3	4
	Doctorate	1	1	2
Occupation in Education				
	No job in primary/secondary education	30	26	56
	Job in primary/secondary education	10	14	24
Political Affiliation				
	Mean on the VAS[a]	47.5	39.65	43.58

[a] Higher mean on VAS is considered more conservative; Lower mean is considered more liberal.

in participating in the study.

2.2. Procedures

2.2.1. Narrative stimuli

Individual narratives used as stimuli in this study were the same as those used in Ludlow (2013). These were originally collected over the course of Redmond et al. (2011) investigation into the psycholinguistic characteristics of children with SLI and ADHD. In Ludlow (2013), these narratives were presented to raters as audio sample files. In the current study, these three samples were transcribed orthographically and presented to raters. Orthographic and phonemic transcripts of each child can be found in Table 2.

Table 2
Orthographic and Phonetic Transcript of Language Samples.

	Orthographic	Phonemic
SLI Child	They went back.	ðe wɛnt bæk
	They went to, they went to home after school.	θe wɛnt tu, θe wɛnt tu hom æftɚ skul
	Mom said, “What, where you, where you like to go?”	mɑm sɛd wət wɛɚ ju wɛɚ ju lɑɪk tu go
	They shout, “McDonald’s!”	ðe ʃɑʊt mɪkdɑnlds
	They went in the car went to McDonald’s.	ðe wɛnt ɪn θə kɑr wɛnt tu mɪkdɑnlds
	The kid said, “I want a bi a hamburger.”	ðə kɪd sɛd aɪ want ə bɪ ə hæmbəɡɚ
	The mom said, “I want a salad.”	ðə mɑm sɛd aɪ want ə sælɪd
	The uh their kid wanted a happy meal and a… Coke.”	ðə ə ðɚ kɪd wantɪd ə hæpi mil ænd ə kok

	Orthographic	Phonemic
TN Child	When Lisa got home their mother said “We’re going out to dinner”	wɛn lisə gat hom ðɛɚ məðɚ səd wɪɚ goiŋ aʊt tu dɪnɚ
	“Where do you want to go?”	wɛɚ du ju want tu go
	Lisa shouted, “Mickanddonald’s!”	lisə ʃaʊtɪd mɪkænd danlds
	They both hoppeded in the car.	ðe boθ haptɪd ɪn ðə kar
	And their mother derove them to the nearest Mickanddonald’s	ænd ðɛɚ məðɚ dəov ðɛm tu ðə nɪəɪst mɪkænd danlds
	Lisa couldn’t, didn’t know what to order.	lisə kʊdnt dɪdnt no wət tu ordɚ
	And… and her brother went to order.	ænd ænd hɚ brəðɚ wɛnt tu ordɚ
	When they got up to the counter, her and her mother ordered a salad	wɛn ðe gat əp tu ðə kʊuntɚ hɚ ænd hɚ məθɚ ordəd ə sælɪd
	And Lisa finally decided she wanted a cheeseburger and a Coke with a vanilla shake.	ænd lisə faɪnli disaɪdɪd ʃi wantɪd ə tʃisbəɡɚ ænd ə kok wɪθ ə vənɛlə eʃk
	When her “that’s twelve dollars and fifty cents,” said the clerk.	wɛn hɚ ðæts twɛlv dalə̆s ænd fɪfti sɛnts sɛd ðə klə̆k
	And their mother reached for the purse and then she realized she had left it at home on the counter.	ænd ðɛɚ məðɚ riʃt for ðə pə̆s ænd ðɛn ʃi rilə̆laɪzd ʃi hæd left ɪt hom an ðə kaʊntɚ

	Orthographic	Phonemic
ADHD Child	The problem was the m the purse wasn’t there.	ðə prabləm wə̆s ðə m ðə pə̆s wə̆snt ðɛɚ
	And they wanted-d uh a chocolate ice cream cone and um… I think a drink.	ænd ðe wantɪd d ʌ ə tʃaklɛt aɪs krɪm kon and ʌm aɪ θɪnk ə drink
	They really we need a drink	ðe rili wi nidɪd ə drink
	They was thirsty	ðe wə̆s θə̆sti
	And um… and um… and um… the, the mother said, “Tonight we were going to McDonald’s.”	ænd ʌm ænd ʌm ænd ʌm ðə ðe məðɚ sɛd tunaɪt wi wɚ goiŋ to mɪkdanlds
	And at night, they, they jumped in the car	ænd æt naɪt ðe ðe dʒə̆mpt ɪn ðə kar
	And the mother forgot her purse	ænd ðə məðɚ fə̆gat hɚ pə̆s
	And then they drove to McDonald’s, the nearest one.	ænd ðɛn ðe drov tu mɪkdanlds ðə nɪəɪst wən
	And they went in	ænd ðe wɛnt ɪn
	And they didn’t know what the do to th, d, decide.	ænd ðe dɪdnt no wət ðə du tu θ d dɪsaɪd

When they occurred, pauses and fillers (e.g. “um” and “uh”) were incorporated into the transcriptions. The children who provided the narratives were each seven years old, white, male, and monolingual speakers of Standard American English. These boys were matched on maternal educational level and on nonverbal intelligence (within one standard error of measurement) based on the Naglieri (2003). The assigned gender of the child was changed in the questionnaires to test for gender effects of the speaker and interaction between the rater and the speaker. Each sample was one minute in length and did not contain any details that could have

Table 3
Speaker Characteristics.

	ADHD	SLI	TN
Age (years; months)	7; 1	7; 2	7; 9
Grade	1st	1st	2nd
Nonverbal IQ[a]	99	94	100
PCC [b]	92.50	93	100
% Grammatical T-Units [c]	90	42.86	83.34
Behavioral [d]	62	50	50
Verbal [e]	17	11	23

Note: Speaker characteristics’ test results are based on tests and criteria utilized by a certified speech-language pathologist and a clinical psychologist to determine each individual diagnosis.

[a] Naglieri Nonverbal Ability Test, standard score (M = 100, SD = 15).
[b] Percent Consonant Correct (total correct consonants/total consonants).
[c] Percent Grammatical T-Units.
[d] Child Behavior Checklist, DSM-ADHD syndrome scale.
[e] Clinical Evaluation of Language Fundamentals-Fourth Edition.

potentially revealed personal or demographic characteristics of the speakers. High levels of phonological accuracy were present in the narratives (range percent consonant correct scores: 92.5–100). Characteristics of each child speaker can be found in Table 3. A certified speech-language pathologist and a clinical psychologist provided independent diagnoses of SLI and ADHD. The child with typical neurodevelopment was not receiving any special services.

Children retold the "McDonald's Story" from the *Test of Narrative Language* (Gillam & Pearson, 2004), a standardized clinical measure used to collect age-referenced information on children's ability to understand, reproduce, and generate stories. The "McDonald's Story" describes a family's trip to McDonald's wherein after placing their orders, the mother discovers she has forgotten her wallet. After being told the story and asked comprehensions questions, children are asked to retell the narrative. The narrative retellings produced were representative of the children's respective designation (SLI, ADHD, TN). In other words, similar to the conversational outcomes reported by Redmond (2004), the narrative produced by the boy with SLI contained several grammatical errors (e.g., *"Mom said where you like to go?"*) and the narrative from the boy with ADHD was temporally disorganized and included tangential details (e.g., *"um I think a drink, they really needed a drink"*). In contrast, the boy with TN retold the story with high levels of accuracy and only minor errors (e.g., *"they hoppeded in the car"*) (See Table 2).

2.2.2. Rater questionnaire

The questionnaire required approximately 15 min to complete and was finished in one testing session. The protocol consisted of an introductory training section orienting participants to the use of 100 mm long visual analog scale (VAS) (e.g. "*Use this line to indicate how warm or cold you think this room is.*"). Following the training, participants were presented with three narratives and 15 questions after each narrative. Raters were given a short description of the task required by the children (i.e. *"In each case, an adult has just read a brief story to each child about a family trip to McDonald's. The transcripts denote three different children's various retellings of the exact same story.*"). Raters were informed that the transcripts they were reading were in the child's exact words.

After reading each narrative, raters provided evaluations using a VAS in response to 15 question prompts about either the narrative itself, the speaker, or the speaker's family. Participants also completed a set of questions providing demographic information about themselves. The narrative questionnaire can be found in Table 4. Twelve of the 15 questions used in the current study were used by Ludlow (2013), who had extrapolated her questions from previous inventories. We added 3 items asking raters to evaluate the speaker's mathematical abilities (*how good at math do you think this child is?)* and characterize their classroom experiences (*how often do you think this child misbehaves in class?; how often do you think this child is teased by their classmates?*). Visual analog scale values assigned to individual questions were summed and averaged to generate five composite measures: narrative attributes, academic attributes, social attributes, behavioral attributes, and family attributes (For the composite summary, see Table 5).

The raters were unaware of the clinical status of each speaker (SLI, ADHD, TN). Instead, they were provided with brief generic

Table 4
Narrative Questionnaire.

Question	Source
1. How well does the child tell their story?	Rice et al. (1993)
Not well-Well	Ludlow (2013)
2. How relevant were the details included?	Newman and McGregor (2006)
Not Relevant-Relevant	Ludlow (2013)
3. How correct was their grammar?	Newman and McGregor (2006)
Very Incorrect-Correct	Ludlow (2013)
4. How smart do you think this child is?	Rice et al. (1993)
Below Average-Well Above Average	Ludlow (2013)
5. How good do you think this child is at their school work?	Overby et al. (2007)
Very Poor-Very Good	Ludlow (2013)
6. How good at math do you think this child is?	Kiefer & Sekaquaptewa (2007)
Very Poor-Very Good	
7. How likely is it that this child could try harder?	Bebout and Arthur (1992)
Not Likely-Very Likely	Ludlow (2013)
8. How often do you think this child is distracted in class?	Ludlow (2013)
Not often-Very often	
9. How often do you think this child misbehaves in class?	Overby et al. (2007)
Not often-Very often	
10. How often do you think this child is teased by their classmates?	Overby et al. (2007)
Not often-Very often	
11. How easy do you think this child makes new friends?	Overby et al. (2007)
Not easy-Very easy	Ludlow (2013)
12. How often do you think this child is a classroom leader?	Rice et al. (1993)
Not often-Very often	Ludlow (2013)
13. How often do you think this child's parents read to them?	Ludlow (2013)
Never-Daily	
14. How educated do you this child's parents are?	Rice et al. (1993)
Less than high school-Advanced degree	Ludlow (2013)
15. How much money do you think this child's family makes each year?	DeThorne and Watkins (2001)
Well Below Average-Well Above Average	Ludlow (2013)

Table 5
Composite Summary.

Composite	Questions
Narrative Attributes	1) How well does the child tell their story?
	2) How relevant were the details included?
	3) How correct was their grammar?
Academic Attributes	4) How smart do you think this child is?
	5) How good do you think this child is at their school work?
	6) How good at math do you think this child is?
Behavioral Attributes	7) How likely is it that this child could try harder?
	8) How often do you think this child is distracted in class?
	9) How often do you think this child misbehaves in class?
Social Attributes	10) How often do you think this child is teased by their classmates?
	11) How easy do you think this child makes new friends?
	12) How often do you think this child is a classroom leader?
Family Attributes	13) How often do you think this child's parents read to them?
	14) How educated do you think this child's parents are?
	15) How much money do you think this child's family makes each year?

details about each speaker highlighting their gender. For example, a boy's description read, "*Ethan is a 7-year old boy who likes playing with Legos and watching professional wrestling on TV*," while a girl's description read, "*Olivia is a 7-year old girl who likes playing with her Barbie's in the 'Barbie Dream House' and making bracelets for her friends*." Names were selected based on the most popular baby names of 2009, which would align with their age of 7 years at the time of test administration ("Popular Names in 2009, 2016Popular Names in 2009, 2016," https://www.ssa.gov/cgi-bin/namesbystate.cgi). Gender loaded actions and toys were chosen based on ratings provided by various websites: (Bradley, 2016a, http://www.beyondtalk.net/gift-boys/; Bradley, 2016b, http://www.beyondtalk.net/gift-girls/; Gumbinner, 2013; "Top Toys for 7 Year Old Girls, 2016," https://www.fatbraintoys.com/toys/toys_by_ages/girls/top_picks_7.cfm; "The coolest birthday gifts for 7 year olds, 2016," http://coolmompicks.com/ultimate-birthday-party-gift-guide/coolest-birthday-gifts-for-7-year-olds/).

2.2.3. Design

A Latin square design and linear mixed model ANOVA analysis were used to control for potential order effects. Eight versions of the narrative questionnaire were created, wherein the serial order (1st, 2nd, 3rd) of the narrative, the speaker's group, (SLI, ADHD, TN) and the speaker's assigned gender (M, F) was counterbalanced (see Table 6 for narrative questionnaire versions). Each combination of serial position, clinical group, and child speaker's assigned gender was rated by five male and five female participants. Prior to analyses, responses to items 7, 8, 9, and 10 were reverse coded to ease interpretation. Descriptive statistics, including means, standard deviations, and ranges were calculated for each questionnaire item, the speaker's assigned gender (TM, TF), the rater's gender (RM, RF), and the speaker's group (SLI, ADHD, TN). These values are presented in Table 7.

3. Results

Complete data were available for 66/80 participants (82.5%). Speaker ratings were complete for 72/80 (90%) and demographic information was complete for 75/80 (93.75%). Individual items were missing from 13 questionnaires and in one questionnaire, seven items were missing. Fourteen instances of missing narrative/speaker questions occurred (14/3,600 = 0.4%). Five demographic questions (5/800 = 0.6%) were unanswered. VAS values provided by the raters ranged from 0 mm to 100 mm, indicating that as a group, participants had used the whole scale. ANOVA analyses examining main effects for speaker group, speaker assigned gender, rater gender, and their interactions were used to address our research questions. Effect sizes were calculated using partial eta squared.

Table 6
: Latin-Square Design Questionnaire Versions.

	Order of Appearance in the Questionnaire		
Questionnaire	First	Second	Third
1	SLI[a], GIRL[b]	TN, BOY	ADHD, BOY
2	ADHD, GIRL	SLI, GIRL	TN, GIRL
3	TN, GIRL	ADHD, BOY	SLI, GIRL
4	ADHD, GIRL	SLI, GIRL	TN, BOY
5	SLI, BOY	TN, GIRL	ADHD, GIRL
6	ADHD, BOY	SLI, BOY	TN, GIRL
7	TN, BOY	ADHD, GIRL	SLI, BOY
8	SLI, BOY	TN, BOY	ADHD, BOY

[a] Clinical status of the child.
[b] Assigned gender of the child.

Table 7
Descriptive Statistics.

	ADHD Speaker Sample		SLI Speaker Sample		TN Speaker Sample	
	Male	Female	Male	Female	Male	Female
Raters						
Male						
Narrative	25.38[a]	19.28	28.78	31.75	66.73	57.65
Quality	(16.63)[b]	(10.79)	(16.29)	(16.86)	(21.19)	(22.46)
	2.67-53.00[c]	*8.00-40.33*	*3.67-59.67*	*12.00-67.33*	*9.67-89.33*	*12.33-89.67*
Academic	33.85	30.12	35.70	40.95	64.07	56.52
Attributes	(16.39)	(17.15)	(14.65)	(19.45)	(17.60)	(21.23)
	5.67-55.67	*5.00-58.33*	*7.00-59.67*	*6.33-90.33*	*9.67-92.00*	*14.67-94.33*
Behavioral	34.28	43.15	44.00	46.85	49.11	57.02
Attributes	(12.40)	(9.46)	(8.30)	(11.60)	(14.03)	(10.68)
	11.33-47.33	*22.3-60.33*	*25.3-57.67*	*9.33-64.00*	*8.33-66.00*	*27.67-69.00*
Social	39.58	40.93	46.15	47.05	62.85	58.20
Attributes	(16.02)	(11.34)	(13.97)	(16.67)	(18.18)	(17.69)
	6.00-58.33	*18.00-56.67*	*24.00-77.00*	*6.00-81.67*	*8.00-94.00*	*15.67-95.33*
Family	36.33	31.50	36.30	39.23	63.03	53.08
Attributes	(15.33)	(15.70)	(15.21)	(20.83)	(17.22)	(21.64)
	7.33-58.67	*5.33-53.33*	*9.00-55.00*	*7.67-82.33*	*8.67-91.33*	*15.67-94.67*
Female						
Narrative	23.17	21.32	30.62	32.75	60.33	64.65
Quality	(16.40)	(14.41)	(17.68)	(15.31)	(23.86)	(19.25)
	2.33-55.00	*4.67-53.33*	*5.33-77.00*	*9.00-78.67*	*14.00-95.00*	*23.00-96.33*
Academic	31.27	35.28	34.45	41.47	55.43	60.75
Attributes	(17.29)	(15.04)	(16.91)	(13.19)	(19.77)	(19.47)
	3.00-54.33	*13.00-57.67*	*4.00-67.67*	*16.00-77.00*	*20.67-92.67*	*24.00-94.33*
Behavioral	45.41	43.50	44.75	46.85	54.25	56.38
Attributes	(16.79)	(9.72)	(15.83)	(11.43)	(14.26)	(9.23)
	23.67-88.33	*20.00-57.33*	*26.00-86.00*	*23.33-69.33*	*35.00-92.33*	*35.67-67.00*
Social	38.03	39.26	45.40	46.36	57.06	63.51
Attributes	(12.45)	(12.15)	(14.74)	49(14.26)	(16.76)	(15.04)
	8.00-53.00	*13.67-57.33*	*22.67-62.33*	*22.33-90.67*	*28.00-96.33*	42.00-90.00
Family	35.72	36.72	49.00	44.80	56.42	57.38
Attributes	(18.10)	(12.62)	(18.69)	(14.10)	(20.92)	(21.94)
	2.00-57.00	*12.00-59.67*	*5.67-75.33*	*11.67-76.33*	*6.67-93.00*	*17.33-96.00*

[a] Mean.
[b] Standard Deviation.
[c] Range.

Cohen (1988) benchmarks were used to interpret observed effect sizes (small = 0.01, medium = 0.06, large = 0.14). ANOVA statistics are presented in Table 8.

1 *Are the ratings collected from our participants using a text vignette of children's narratives similar to those obtained by* Ludlow (2013) *using the same narratives but presented in that study to participants as audio samples?*

Ludlow (2013) studied mothers who provided ratings for audio samples, whereas the current study examines male and female participants ratings through orthographic transcriptions of the same samples presented as either male or female. Twelve questions were directly comparable between the current study and from the questionnaire used by Ludlow (2013). Fig. 1 compares the means and standard deviations obtained from the mothers in Ludlow (2013) study and the female participants in this study for the twelve corresponding questions. By restricting our evaluation to the female ratings of the male speakers in the current study we ensured parity with the stimuli and participants in Ludlow (2013). As shown in Fig. 1, there was a tendency for current participants to assign lower ratings to the narratives that were presented as orthographic transcriptions relative to Ludlow's participants who were presented with audio files. Two series of twelve one-way ANOVAs, one for audio and one for orthographic transcripts, were conducted to compare patterns of rater differences across the 12 questions as a function of stimuli type (audio file and orthographic transcription). Follow-up pairwise analyses (Sidak, $p < .05$) were used to test significant differences between the speaker groups (ADHD, SLI, TN). Results from the univariate ANOVAs are presented in Table 9. Nine of the twelve questions provided the same patterns across the audio and orthographic group ratings (TN > SLI = ADHD). Of the three questions that had dissimilar rating patterns, the TN speaker maintained significant advantages over the speakers with SLI and ADHD in two of the three analyses (TN > ADHD), while there was no difference between the SLI and ADHD groups. The most discrepant ratings between judgments based on the audio sample and the text version of the narratives were found for the question, "*how likely is it that this child could try harder?*" This was the only question where ratings of the speakers from the clinical groups were not significantly different from the speaker with TN group in the orthographic ratings. Overall, for the female raters in Ludlow (2013) and the female raters in the current study sample, ratings

Table 8
ANOVA results.

		Narrative			Academic			Behavioral			Social			Family		
ANOVAs	*df*	*F*	*p*	Partial η^2	*F*	*p*	Partial η^2	*F*	*p*	Partial η^2	*F*	*p*	Partial η^2	*F*	*p*	Partial η^2
Speaker group	(2, 237)	110.50	< **.001**	.492	51.39	< **.001**	.311	22.05	< **.001**	.162	40.14	< **.001**	.260	35.19	< **.001**	.236
Rater gender	(1, 238)	0.06	.815	< .001	0.04	.851	< .001	3.10	.072	.013	0.19	.662	.001	0.37	.543	.002
Speaker assigned gender	(1, 238)	0.30	.584	.001	0.58	.447	.003	5.33	**.022**	.023	0.29	.593	.001	0.02	.886	< .001
Speaker group X Rater gender	(2, 235)	0.04	.963	< .001	0.20	.819	.002	1.03	.355	.009	0.04	.958	< .001	0.32	.729	.003
Speaker group X Speaker assigned gender	(2, 235)	0.72	.488	.006	0.98	.376	.009	0.23	.795	.002	0.01	.996	< .001	1.63	.197	.014
Rater gender X Speaker assigned gender	(1, 236)	1.46	.227	.006	2.73	.100	.012	3.31	.070	.014	0.89	.346	.004	2.43	.120	.011
Speaker group X Rater gender X Speaker assigned gender	(2, 233)	0.81	.447	.007	0.51	.604	.004	0.84	.434	.007	0.91	.406	.008	0.16	.851	.001

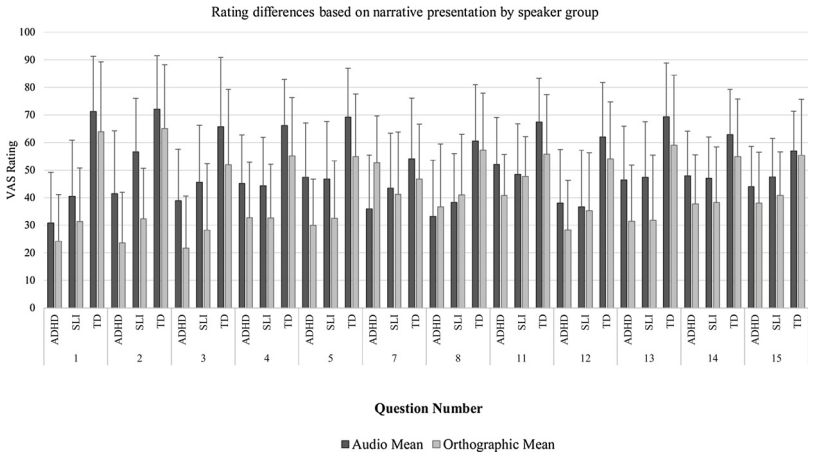

Fig. 1. Speaker group rating differences.
Note. Date displayed in this figure reflects means and standard deviations from Ludlow (2013) and from female raters in the current study. Error bars indicate standard deviations. Refer to Table 4 for the text of each question number.

Table 9
Comparisons of Speaker Group by audio and orthographic presentations.

Question	Audio F (2, 178)	η2	Post hoc	Orthographic F (2, 58)	η2	Post hoc
How well does the child tell their story?	69.69***	.44	TN > SLI > ADHD	20.45***	.42	TN > SLI = ADHD
How relevant were the details included?	30.38***	.26	TN > SLI > ADHD	23.66***	.45	TN > SLI = ADHD
How correct was their grammar?	24.83***	.22	TN > SLI = ADHD	9.02***	.24	TN > SLI = ADHD
How smart do you think this child is?	30.29***	.26	TN > SLI = ADHD	8.16***	.22	TN > SLI = ADHD
How good do you think this child is at their school work?	25.87***	.23	TN > SLI = ADHD	9.28***	.27	TN > SLI = ADHD
How likely is it that this child could try harder?	11.89***	.12	TN > SLI = ADHD	1.69	.06	TN = SLI = ADHD
How often do you think this child is distracted in class?	33.01***	.27	TN > SLI = ADHD	4.93*	.15	TN = SLI SLI = ADHD TN > ADHD
How easily do you think this child makes new friends?	21.04***	.19	TN > SLI = ADHD	3.78*	.12	TN = SLI SLI = ADHD TN > ADHD
How often do you think this child is a classroom leader?	30.58***	.26	TN > SLI = ADHD	8.73***	.24	TN > SLI = ADHD
How often do you think this child's parents read to them?	25.89***	.23	TN > SLI = ADHD	9.21***	.25	TN > SLI = ADHD
How educated do you think this child's parents are?	18.52***	.17	TN > SLI = ADHD	4.92*	.15	TN > SLI = ADHD
How much money do you think this child's family makes each year?	12.90***	.13	TN > SLI = ADHD	5.20**	.15	TN > SLI = ADHD

* $p < .05$.
** $p < .01$.
*** $p < .001$.

provided in response to either audio samples or orthographic transcriptions yielded similar patterns of speaker differences. This suggests that, for the most part, orthographic transcriptions can be used to examine attributional biases associated with children's communication differences. We addressed research questions 2–4 using ratings provided by the current study sample of males and females using orthographic transcriptions

2 *Do raters perceive narratives by children with SLI and ADHD more negatively than a child with typical neurodevelopment?*

Significant medium to large main effects were found for speaker group (SLI, ADHD, TN) across 11 of the 12 questions (all $ps < .05$, partial η^2 range: 11 - 0.37) as shown in Table 9. Follow-up pairwise analyses (Sidak, $p < .05$) were used to test significant differences between the speaker classifications (ADHD, SLI, TN). Higher ratings were consistently assigned to the TN child as compared to children from the speaker groups. The pattern was confirmed across nine questions, such that ADHD = SLI < TN (all $ps < .01$). The speaker with ADHD was rated significantly lower than the speaker with SLI (ADHD < SLI < TN) in two questions related to narrative quality ("*how well does the child tell their story*?" and "*how relevant were the details included?*").

3 *Are there differences between ratings provided by male and female participants across different traits and characteristics?*

The ANOVA revealed no significant rater effects on ratings for any of the composites: narrative, academic, behavioral, social, or

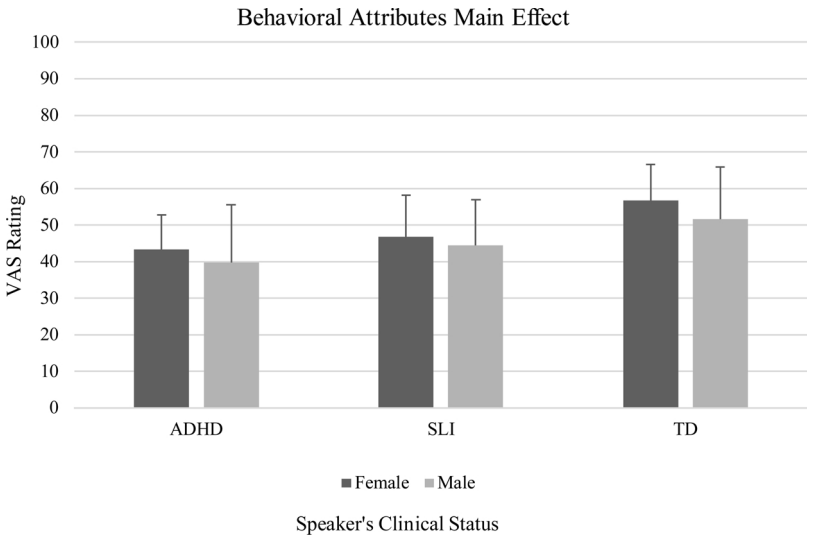

Fig. 2. Behavior Attributes Main Effect for Speaker's Gender.
Note. Sidak results of pairwise analyses indicated that the male speakers were rated significantly lower than female speakers for behavioral attributes (p = .022). Error bars indicate standard deviations.

family (*p* range: .072–.851). This finding was unexpected and contradicted previous research examining biases attached to physical and intellectual disabilities that have suggested male raters would provide more pejorative ratings than female raters.

4 *Are there differences across different traits and characteristics between the ratings assigned to speakers identified as male and speakers identified as female?*

The ANOVA revealed a small, but significant main effect for speaker's assigned gender on the ratings participants assigned to the behavioral attributes composite ($p = .022$, partial $\eta^2 = 0.023$). Based on previous reports of speaker gender effects in rater evaluations, we predicted that when the speaker was identified to our participants as female they would be rated more pejoratively than when they were presented as male. We found no support for this prediction. Follow-up pairwise analyses revealed that when speakers were presented as male ($M = 45.30$, $SD = 14.87$) they were in fact rated significantly more pejoratively on the behavioral composite than when they were presented as female ($M = 48.96$, $SD = 11.66$: M < F, see Fig. 2). The other composites (narrative, academic, social, or family) did not reveal significant rating differences between the narratives presented as being produced by a boy or a girl.

5 *Are there significant interactions among speaker's clinical status, speaker's gender, and speaker's gender?*

None of the potential interaction terms were statistically significant (Table 8). Contrary to our predictions, male raters did not provide more favorable ratings to male speakers. In fact, a tendency emerged in the opposite direction of our expectations on the behavior composite. Group means indicated male raters provided more favorable ratings to female speakers, but this interaction failed to reach statistical significance ($p = .07$).

4. Discussion

Specific Language Impairment (SLI) and Attention-Deficit/Hyperactivity Disorder (ADHD) represent two of the most common neurodevelopmental disorders and their management consumes considerable educational and health resources. Elevated risks for a variety of long-standing social difficulties have been implicated in each condition, including peer neglect and victimization (Redmond et al., 2011). How peers, parents, teachers, and other adults interpret children's atypical communication is a likely contributor to these social difficulties. These interpretations start with listener's snap judgments of speaker attributes. Based on very limited exposures, people readily make a variety of assumptions about children's intellectual capacities, their levels of motivation, their social skills, and their family backgrounds.

Two decades of research on the issue consistently document the presence of robust, multidimensional, negative biases, within listeners' initial impressions of speakers who display the linguistic differences associated with SLI (e.g. elevated rates of grammatical error). Although ADHD has not yet enjoyed the same level of consideration as SLI, preliminary indications suggest that communicative differences associated with ADHD, such as false starts and revisions, trigger similar attributional biases and negative stereotypes. Individual rater variation in these judgments has been documented. However, the source of this variation has remained elusive. For example, educational background, professional experiences, personal experiences with disability, or age differences across listeners contribute very variance.

In this study, we considered the potential influence gender might have on listener's reactions to atypical communication along with stimuli presentation (audio versus orthographic). We used children's narrative retellings collected from a standardized test, which kept the story content uniform across speakers. Three male speakers (SLI, ADHD, and TN) who had provided narrative retellings over the course of a previous study were matched on the basis of maternal educational level, nonverbal ability, and narrative length. In order to examine potential gender effects, we used orthographic versions of these narratives, essentially stripping away any phonological or paralinguistic features that could have provided raters with reliable cues to the speaker's gender. These transcriptions were presented to 80 raters (40 males and 40 females) as being produced by either a male or female speaker. Raters used visual analog scales to evaluate each speaker across a range of traits and characteristics.

Our results aligned closely with previous studies and documented further that the language differences associated with SLI and ADHD triggered strong negative reactions in adults, especially in comparison with children with typical neurodevelopment. These negative judgments between SLI children and TN children and ADHD children and TN children were seen in all five composites: narrative, academic, behavioral, social and family attributes. Most notably are the negative attributes associated with ratings that cannot be directly inferred from the orthographic transcriptions, indicating that adults tend to make negative judgments about ADHD and SLI children's academic, behavioral, social and family attributes based on their ability to retell a narrative.

Although the results from this study were not directly comparable to those of Ludlow (2013), the overall pattern of ratings our participants applied to the orthographic transcriptions was highly similar to the pattern of rating applied to the audio versions of the same narratives. This suggests that either method can be used to capture potential attributional biases, but there are probably trade-offs to consider. Orthographic stimuli might overestimate effects relative to how adults would probably respond during their initial encounters with children who have SLI or ADHD under real-world circumstances. On the other hand, orthographic stimuli allow investigators to directly manipulate type, severity, and frequency of individual clinical symptoms (e.g. dysfluencies, word choice errors, morphological omissions) and examine their relative effects. In addition to gender, text-based vignettes can be used to examine the relative contributions of other demographic variables (e.g. race, ethnicity, income level).

With regard to anticipated perceived gender differences, we found little support for our predictions. There were no significant differences between the ratings collected from our male and female participants across speakers or across composites. A perceived speaker gender effect was found but it applied only to items representing the questions associated with our behavior composite. When the speaker was presented to raters as a male, they were rated more likely to be perceived as having more behavior problems than when the same speakers were presented to raters as a female. Therefore, there is a small, but significant perceived gender effect through assigned genders as it relates to behavioral attributes. This result was not consistent with previous reports of more pejorative ratings being associated with female targets with disabilities. Additional research could resolve this discrepancy. Differences across studies may have been a function of the particular disabilities presented to raters.

4.1. Limitations and future directions

There were several limitations associated with this study. The chain-referral recruitment strategy may have introduced bias into our sample in that people refer personal contacts who share common demographic characteristics with them. Our sample was for the most part, white, non-Hispanic, and highly educated. More diverse study samples of adults might yield different results. This concern was mitigated somewhat by consistency in outcomes between our study and previous studies, however, future studies should examine more deliberately the possibility that racial, ethnic, and economic differences contribute to variation in adult's attributional biases. As with most studies in this area, our questionnaire did not include an "uncertain" response option. It is possible that the inclusion of this option would have altered our outcomes. However, results from DeThorne and Watkins (2001) suggest that when the "uncertain" option is offered to raters they rarely use it. Future research could address the level of certainty with which individuals provide their judgments to child speakers with atypical communication behaviors. We also relied on a single measurement method (orthographic transcriptions) to measure adult reactions to atypical communicative behaviors. A more complete picture requires a multi-method approach. Visual cues, for example, represent an examined element of adult judgments of children with atypical communication. Gilliam, Maupin, Reyes, Accavitti, and Shic (2016)) used eye-tracking measurements to examine potential implicit teacher biases in interpreting behaviors of preschool age children. Results indicated that when presented with silent video samples, teachers looked longer at boys than girls and longer at African American children than White children. Teacher participants were also more likely to interpret children's deliberately staged neutral behaviors as potentially problematic when the target children were either male or African American. These findings suggest implicit predispositions among teachers to seek confirming evidence of problematic male behavior. The observed bias based on visual information appears to align well with the presence of more pejorative ratings of children's behavior when our adult participants were told the narrative had come from a boy. In future studies, eye-tracking measurements could be combined with participant's attributional ratings to capture more fully the scope of implicit biases toward atypical communication.

4.2. Clinical implications

Although preliminary, our results on the nature of adult biases toward children's atypical communication behaviors have potential clinical implications. In most cases, the identification of language impairments in children requires a teacher or other referral source to recognize atypical communication behaviors as potential symptoms of an underlying language impairment and then to pass on their concerns to either a speech-language pathologist or an assessment team American Speech-Language-Hearing Association (2010). The results of this study and other studies on the topic indicate that symptoms of language impairments trigger a host of

potential negative social, intellectual, behavioral, and moral misattributions in adults. These common biases encourage mischaracterizations of language symptoms as the manifestation of some other deficit or disparity leading to potential decreases in appropriate referrals and access to services. Efforts directed at raising public awareness of children's language impairments, such as those offered by Boys Town National Research Hospital (www.dldandme.org) and the Raising Awareness of Language Disorder (RADLD) (www.youtube.com/user/RALLIcampaign), represent movement in the right direction towards offsetting these biases. However, on their own, these are unlikely to be sufficient. As the results of Rice et al. (1993) suggest, even speech-language pathologists, who receive extensive training on the nature and treatment of language impairments, share these biases. Public awareness campaigns combined with sensitivity training programs directed at teachers, speech-language pathologists, psychologists, and other health care professionals are likely to yield more substantive improvements.

Author statement

Manuscript title: "Exploring gender as a potential source of bias in adult judgments of children with specific language impairment and attention-deficit/hyperactivity disorder"

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript.

CRediT authorship contribution statement

Alison Elizabeth Shimko: Conceptualization, Methodology, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration, Funding acquisition. **Sean Redmond:** Conceptualization, Methodology, Software, Resources, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Amy Ludlow:** Resources, Writing - review & editing. **Andrea Ash:** Software, Validation, Formal analysis, Resources, Data curation, Writing - review & editing, Visualization.

References

Allard, E. R., & Williams, D. F. (2008). Listeners' perceptions of speech and language disorders. *Journal of Communication Disorders, 41*(2), 108–123. https://doi.org/10.1016/j.jcomdis.2007.05.002.

American Speech-Language-Hearing Association (2010). *Guidelines for the roles and responsibilities of the school-based speech-language pathologist [Guidelines]*. Available fromwww.Asha.org/policy.

Bebout, L., & Arthur, B. (1992). Cross-cultural attitudes toward speech disorders. *Journal of Speech Language and Hearing Research, 35*(1), 45. https://doi.org/10.1044/jshr.3501.45.

Bogdan, R., & Biklen, D. (1993). Handicapism. In M. Negler (Ed.). *Perspectives of disabilities* (pp. 69–76). (2nd ed.). Palo Alto, CA: Health Markets Research.

Bradley, D. (2016a). *Gift ideas for boys by age & category* (n.d.a) Retrieved February 10, 2016, fromhttp://www.beyondtalk.net/gift-boys/.

Bradley, D. (2016b). *Gift ideas for girls by age & category* (n.d.b) Retrieved February 11, fromhttp://www.beyondtalk.net/gift-girls/.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences, 2nd Edition. Hillsdale, N. J. *Lawrence Erlbaum*.

DeThorne, L. S., & Watkins, R. V. (2001). Listeners' perceptions of language use in children. *Language, Speech, and Hearing Services in Schools, 32*(3), 142–148. https://doi.org/10.1044/0161-1461(2001/012).

Fine, M., & Asch, A. (1981). Disabled women: Sexism without the pedestal. *Journal of Sociology and Social Welfare, 8*, 233.

Gartner, A., Lipsky, D. K., & Turnbull, A. P. (1991). *Supporting families with a child with a disability: An international outlook.* Baltimore: Paul H. Brookes.

Gillam, R. B., & Pearson, N. A. (2004). *Test of narrative language.* Austin, TX: Pro-Ed.

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). *Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions?* Yale Child Study Center991–1013 September.

Gumbinner, L. (2013). *What to buy a 7 year old girl: 13 gift ideas.* Retrieved February 17, 2016, fromhttp://coolmompicks.com/blog/2013/01/28/best-gifts-for-seven-year-old/.

Hasson, H., & Arnetz, J. E. (2009). The impact of an educational intervention on nursing staff ratings of quality of older people care: A prospective, controlled intervention study. *International Journal of Nursing Studies, 46*(4), 470–478. https://doi.org/10.1016/j.ijnurstu.2008.04.007.

Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., Brownlie, E. B., ... Wang, M. (1999). Fourteen-year follow-up of children with and without Speech/Language impairments. *Journal of Speech and Hearing Research, 42*(3), 744. https://doi.org/10.1044/jslhr.4203.744.

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes, gender identification, and math-related outcomes. *Psychological Sciences, 18*(1), 13–18. https://doi.org/10.1111/j.1467-9280.2007.01841.x.

Ludlow, A. (2013). *Narrative and language use: How mothers regard school-aged children* (Master's Thesis). Retrieved fromhttp://content.lib.utah.edu/utils/getfile/collection/etd3/id/2532/filename/2538.pdf.

Naglieri, J. A. (2003). *Naglieri nonverbal ability test—Individual administration manual.* San Antonio, TX: The Psychological Corporation.

National Institutes of Health [NIH] (2009). Consensus Development Conference Statement: Diagnosis and treatment of attention-deficit/hyperactivity disorder (ADHD). *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(2), 182–193. https://doi.org/10.1097/00004583-200002000-00018.

Nelson, J. A. (1994). *The Disabled, the media, and the information age.* Westport, CT: Greenwood Press.

Newman, R. M., & McGregor, K. K. (2006). Teachers and laypersons discern quality differences between narratives produced by children with or without SLI. *Journal of Speech Language and Hearing Research, 49*(5), 1022–1036. https://doi.org/10.1044/1092-4388(2006/073).

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. https://doi.org/10.1111/jcpp.12573.

Overby, M., Carrell, T., & Bernthal, J. (2007). Teachers' perceptions of students with speech sound disorders: A quantitative and qualitative analysis. *Language, Speech, and Hearing Services in Schools, 38*(4), 327–341. https://doi.org/10.1044/0161-1461(2007/035).

Panek, P. E., & Jungers, M. K. (2008). Effects of age, gender, and causality on perceptions of persons with mental retardation. *Research in Developmental Disabilities, 29*(2), 125–132. https://doi.org/10.1016/j.ridd.2007.01.002.

Popular Names in 2009 (2016). *Popular names in 2009* (n.d.). Retrieved February 19, 2016, fromhttps://www.ssa.gov/cgi-bin/namesbystate.cgi.

Rice, C. J. (2009). Attitudes of undergraduate students toward people with intellectual disabilities: Considerations for future policy makers. *College Student Journal, 43*(1), 207–215.

Rice, M. L., Hadley, P. A., & Alexander, A. L. (1993). Social biases toward children with speech and language impairments: A correlative causal model of language

limitations. *Applied Psycholinguistics, 14*, 445–471. https://doi.org/10.1017/s0142716400010699.
Redmond, S. M. (2016). Language impairment in the attention-deficit/hyperactivity disorder context. *Journal of Speech Language and Hearing Research, 59*(1), 133–142. https://doi.org/10.1044/2015_JSLHR-L-15-0038 62.
Redmond, S. M. (2004). Conversational profiles of children with ADHD, SLI and typical development. *Clinical Linguistics & Phonetics, 18*(2), 107–125. https://doi.org/10.1080/02699200310001611612.
Redmond, S. M., Thompson, H. L., & Goldstein, S. (2011). Psycholinguistic profiling differentiates specific language impairment from typical development and from attention-deficit/hyperactivity disorder. *Journal of Speech Language and Hearing Research, 54*(1), 99–117. https://doi.org/10.1044/1092-4388(2010/10-0010).
Scahill, L., & Schwab-Stone, M. (2000). Epidemiology of adhd in school-age children. *Child and Adolescent Psychiatric Clinics of North America, 9*(3), 541–555. https://doi.org/10.1016/s1056-4993(18)30106-8.
Tang, C. S., Davis, C., Wu, A., & Oliver, C. (2000). Chinese children's attitudes toward mental retardation. *Journal of Developmental and Physical Disabilities, 12*, 73–87.
The coolest birthday gifts for 7 year olds (2016). *The coolest birthday gifts for 7 year olds* (n.d.). Retrieved May 15, fromhttp://coolmompicks.com/ultimate-birthday-party-gift-guide/coolest-birthday-gifts-for-7-year-olds/.
Tomblin, J. B., & Nippold, M. A. (2014). *Understanding individual differences in language development across the school years.* New York, NY: Psychology Press.
Tomblin, J. B., 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. https://doi.org/10.1044/jslhr.4006.1245.
Top Toys for 7 Year Old Girls (2016). *Top toys for 7 year old girls* (n.d.). Retrieved May 13, 2016, fromhttps://www.fatbraintoys.com/toys/toys_by_ages/girls/top_picks_7.cfm.
Vilchinsky, N., Werner, S., & Findler, L. (2010). Gender and attitudes toward people using wheelchairs: A multidimensional perspective. *Rehabilitation Counseling Bulletin, 53*(3), 163–174. https://doi.org/10.1177/0034355209361207.
Weisel, A., & Florian, V. (1990). Same-and cross-gender attitudes toward persons with physical disabilities. *Rehabilitation Psychology, 35*(4), 229. https://doi.org/10.1037//0090-5550.35.4.229.
Wenker, R. B., Wegener, J. G., & Hart, K. J. (1996). The impact of presentation mode and disfluency on judgments about speakers. *Journal of Fluency Disorders, 21*(2), 147–159. https://doi.org/10.1016/0094-730x(95)00110-s.