JSLHR

Research Article

Explicit and Implicit Verbal Response Inhibition in Preschool-Age Children Who Stutter

Julie D. Anderson[a] and Stacy A. Wagovich[b]

Purpose: The purpose of this study was to examine (a) explicit and implicit verbal response inhibition in preschool children who do stutter (CWS) and do not stutter (CWNS) and (b) the relationship between response inhibition and language skills.

Method: Participants were 41 CWS and 41 CWNS between the ages of 3;1 and 6;1 (years;months). Explicit verbal response inhibition was measured using a computerized version of the grass–snow task (Carlson & Moses, 2001), and implicit verbal response inhibition was measured using the baa–meow task. Main dependent variables were reaction time and accuracy.

Results: The CWS were significantly less accurate than the CWNS on the implicit task, but not the explicit task. The CWS also exhibited slower reaction times than the CWNS on both tasks. Between-group differences in performance could not be attributed to working memory demands. Overall, children's performance on the inhibition tasks corresponded with parents' perceptions of their children's inhibition skills in daily life.

Conclusions: CWS are less effective and efficient than CWNS in suppressing a dominant response while executing a conflicting response in the verbal domain.

Executive function (EF) refers to the conscious control we exert over our behaviors, emotions, and thoughts (Zelazo & Müller, 2010). It is a broad term that includes a variety of domain-general cognitive abilities such as working memory, cognitive flexibility, and inhibition. These functions, for example, allow us to make mental shifts quickly, adapt to new situations or information, and control our behavior. Importantly, components of EF have different developmental trajectories. The rudiments of EF emerge during the first few years of life and continue to develop and become more integrated throughout the preschool years, with more complex forms of EF continuing to develop well into adolescence (Best & Miller, 2010). Response inhibition, the focus of this study, is among the first to develop (Jurado & Rosselli, 2007). The bulk of development in this area appears to take place during the preschool years, with growth continuing into the early school-age years (e.g., Gerstadt, Hong, & Diamond, 1994; for discussion, see Montgomery & Koeltzow, 2010).

[a]Department of Speech and Hearing Sciences, Indiana University, Bloomington
[b]Department of Communication Science and Disorders, University of Missouri, Columbia

Correspondence to Julie D. Anderson: judander@indiana.edu

Editor: Julie Liss
Associate Editor: Hans-Georg Bosshardt

Received April 5, 2016
Revision received August 10, 2016
Accepted October 8, 2016
https://doi.org/10.1044/2016_JSLHR-S-16-0135

The preschool years represent a period of rapid language growth as well. The relationship between spoken language development and EF is coupled—language use supports the development of EF skills that, in turn, promote further advances in language development (Müller, Jacques, Brocki, & Zelazo, 2009; Pisoni, Conway, Kronenberger, Henning, & Anaya, 2010). With the development and internalization of language processes, children are also able to use language to regulate behavior, for example, through the use of self-talk (Müller et al., 2009). Fuhs and Day (2011) examined EF and verbal skills in preschool-age children enrolled in a Head Start program. The main finding was that growth in EF skills over several months, from fall semester to spring semester, could be predicted by children's verbal skills, which had been measured in the fall.

The primary purpose of this study was to examine whether children who stutter (CWS) exhibit weaknesses in one particular component of EF, namely, response inhibition. To put this study into context, we first review relevant background information concerning the nature of response inhibition and what is currently known about the cognitive processing skills of CWS, focusing in particular on inhibition.

Response Inhibition

Inhibition is not presently considered to be a unitary construct; rather, it appears to consist of different types (Müller & Kerns, 2015). A variety of taxonomies have been

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

developed in an attempt to differentiate among different types of inhibition. For example, Friedman and Miyake (2004) classify inhibition into three separable types. The first is prepotent response inhibition, which involves simply suppressing a dominant (prepotent) response. More complex forms of prepotent response inhibition (referred to hereafter as *response inhibition*), which are the focus of this study, involve not only suppressing a dominant response but also executing a conflicting, subdominant response. The additional processing requirement in complex response inhibition—having to hold an arbitrary rule in mind—is thought to require greater working memory demand (Anderson & Reidy, 2012; Best & Miller, 2010; Carlson & Moses, 2001; Garon, Bryson, & Smith, 2008). The second type is resistance to distractor interference, which involves resisting irrelevant information to complete a task. Finally, the third type is resistance to proactive interference, which involves suppressing competing irrelevant information from previous task exposures to complete the current task. Of these three types, the first two are similar, with the distinction being that response inhibition focuses on a dominant response to a stimulus and the extent to which that response can be suppressed.

A classic complex response inhibition task is the color-word Stroop task (Stroop, 1935), in which participants view color words printed in a competing color and are asked to name the color of the printed word rather than reading the word itself (e.g., for the word *green* written in blue print, the participant would respond with *blue*). The task involves suppressing the dominant response, which is simply to read the word as printed, and instead responding with the color of the print, the subdominant response. Unfortunately, the color-word Stroop task is not appropriate for use with preschool children because it requires literacy (MacLeod, 1991). A variety of alternative, non–reading-based, Stroop-like measures have, therefore, been developed specifically for children in this age group. The most commonly used and extensively studied Stroop-like measure for preschoolers is the day–night task (Gerstadt et al., 1994), in which children are instructed to say the word *day* when viewing a card depicting nighttime and *night* when shown a picture of daytime.

Gerstadt and colleagues (1994) used the day–night task with preschool and early school-age children. Overall, they found that children's response accuracy improved from 3½ to 7 years of age. However, from ages 3;5 to 4;5 (years;months), response accuracy actually decreased while processing speed increased dramatically, suggesting potential speed–accuracy trade-offs during this period. In addition, children younger than 5 years had greater difficulty suppressing the dominant response as the task progressed. The authors used a control condition in a different group of younger children (aged 3½ to 5 years) to examine whether the working memory demands of the task (i.e., having to remember two rules) were responsible for the difficulty these children were having with the task. The control task tapped memory by presenting children with two cards containing an abstract design (either a squiggle or checkerboard card) and asking them to respond to one design with the word *day* and the other design with the word *night*. Because there was no dominant response to suppress in this task, it removed inhibition demands. The children performed with high accuracy (about 90% across age groups) on this control task, suggesting that the day–night task primarily challenges response inhibition, as opposed to memory.

Carlson and Moses (2001) developed the grass–snow task, a manual variant of the day–night task. Used with children as young as 3 years, participants are asked to point to a white card when they hear "grass" and point to a green card when they hear "snow." Carlson (2005) found that, even among "young 3-year-olds" (ages 36–41 months), 40% received a passing score of 12 of 16 items correct. Performance improved with age, such that of the "older 4 year olds" (ages 54–59 months), 84% received a passing score. In addition, Simpson and Riggs (2009) reported that 3-year-old children are able to retain the rules of the task throughout administration, suggesting that working memory demands are minimal. Thus, the grass–snow task is an appropriate measure of verbal inhibition for preschool-age children. In addition, because it is a manual response task, it lends itself well to the measurement of processing speed (by modifying it in the form of a button-press task) as well as response accuracy.

Most of what we know about response inhibition in preschool children comes from studies that have used the day–night task and its highly correlated variant, the grass–snow task (Guy, Rogers, & Cornish, 2012). These tasks and their variants (e.g., mommy–me, yes–no, black–white, happy–sad) are similar to the adult-appropriate Stroop in that they require children to remember the instructions over a series of trials, suppress the dominant (prepotent) response associated with a stimulus, and activate a conflicting subdominant response (Montgomery & Koeltzow, 2010). In a button-press version of the grass–snow task, for example, the prepotent responses of pressing the grass button when hearing the word *grass* and the snow button when hearing the word *snow* must be suppressed; the child is asked to press "grass" in response to the word *snow* and vice versa. Prepotency in these tasks is magnified because both the stimulus and response set are identical (Müller & Kerns, 2015; Simpson & Riggs, 2005, 2009). As a result, a correct response on one trial (e.g., pressing "grass" when hearing "snow") is also the incorrect response on a subsequent trial (e.g., pressing "grass" when hearing "grass").

Several forms of evidence exist that child-appropriate Stroop-like tasks such as these directly measure response inhibition. First, when the stimulus items and response set do not overlap (e.g., the child hears the word *grass* and then has to press the button for *pants* or some other categorically unrelated object), children's performance dramatically improves because interference between the stimulus items and response set is removed. This suggests that the primary demand in these tasks is response inhibition (Montgomery & Koeltzow, 2010; Simpson & Riggs, 2005). Second, when a delay is interposed between the presentation of the stimulus and when the child can respond on these tasks, children's performance markedly improves (e.g., Diamond, Kirkham, & Amso, 2002; Jones, Rothbart, & Posner, 2003; Ling, Wong, & Diamond, 2016; Montgomery & Fosco, 2012; Simpson et al., 2012). As noted by Ling et al. (2016), preschool children are often so eager to respond that they tend to respond with the first thing that comes to mind, making inhibition particularly challenging. By requiring the child to wait, the delay allows them to respond more thoughtfully and, at the same time, it enables the prepotent response to decay, making inhibition easier. Finally, children's performance on the day–night task and its variants have been shown to be highly congruent with their performance on other behavioral inhibition tasks (e.g., Simpson & Riggs, 2005; Tillman, Thorell, Brocki, & Bohlin, 2008; Wolfe & Bell, 2007).

In sum, conflict tasks such as the grass–snow task are strongly substantiated by research as valid measures of response inhibition in preschool children. Nevertheless, similar to most behavioral tasks, these tasks also require some degree of working memory because children must remember the rules throughout the task. Research has revealed that the memory demands of these tasks, however, tend to be minimal (e.g., Diamond et al., 2002; Gerstadt et al., 1994; Ling et al., 2016; Miller, Giesbrecht, Müller, McInerney, & Kerns, 2012).

Response Inhibition in CWS

Within the broader area of self-regulation, several studies focusing on the temperament characteristics of CWS have suggested that these children differ from children who do not stutter (CWNS; e.g., Anderson, Pellowski, Conture, & Kelly, 2003; Arnold, Conture, Key, & Walden, 2011; Embrechts, Ebben, Franke, & van de Poel, 2000).[1] In particular, there has been recent focus on the area of attentional control in CWS (Anderson & Wagovich, 2010; Chou, 2014; Eggers, De Nil, & Van den Bergh, 2010, 2012; Felsenfeld, van Beijsterveldt, & Boomsma, 2010; Johnson, Conture, & Walden, 2012; Kaganovich, Wray, & Weber-Fox, 2010; Karrass et al., 2006; Schwenk, Conture, & Walden, 2007). Some of these studies have used parent or teacher questionnaires, most of them finding that CWS are rated less well than CWNS on aspects of attention (Eggers et al., 2010; Felsenfeld et al., 2010; Karrass et al., 2006; cf. Anderson & Wagovich, 2010). Of those studies that have used direct measures of attention, either behavioral or electrophysiological, most have yielded similar findings (Chou, 2014; Eggers et al., 2012; Kaganovich et al., 2010; Schwenk et al., 2007; cf. Johnson et al., 2012), with CWS demonstrating weaker skills in aspects of attention allocation. For example, Schwenk et al. (2007) found that CWS demonstrated more shifts in attention away from conversation with their parents and toward the video camera in the room. Eggers et al. (2012) examined attention in CWS and CWNS, ages 4 to 9 years, using a children's version of the Attention Network Test (Fan, McCandliss, Sommer, Raz, & Posner, 2002; Rueda et al., 2004). This test is intended to measure the following three networks or aspects of attention: alerting (maintaining alertness), orienting (selecting relevant information from input), and executive attention (inhibiting a dominant response). The CWS group demonstrated significantly less efficiency in orienting and reduced efficiency in executive attention, a difference that approached significance. This latter trend, as discussed later, is particularly relevant because executive attention involves inhibitory control.

Event-related potential evidence presents a similar picture in terms of attention allocation in CWS (Chou, 2014; Kaganovich et al., 2010). For example, Kaganovich and colleagues (2010) presented CWS and CWNS, 4 to 5 years of age, a set of two types of tones: a 1-kHz tone, which occurred frequently within the set, and a 2-kHz tone (the "deviant" tone), which occurred infrequently. Event-related potential was used to examine nonlinguistic auditory processing between groups as well as attention allocation and working memory updating. Findings revealed that although there were no differences between groups in the P1 and N1 components, which are related to the basic auditory processing of stimuli, the P3 component was elicited by the deviant tones only for the CWNS group. The fact that this component was not elicited for the CWS group in response to the deviant tones was interpreted as an indication of reduced efficiency in attention allocation and working memory updating.

The related area of inhibition (or inhibitory control) in CWS has received less focus. Studies using data from parent questionnaires have produced mixed findings, with some studies reporting evidence of reduced inhibition among CWS (e.g., Eggers et al., 2010; Embrechts et al., 2000) and others reporting no differences between CWS and CWNS (e.g., Anderson & Wagovich, 2010).

Even fewer studies have directly measured response inhibition in CWS through behavioral tasks. As discussed earlier, Eggers and colleagues (2012) explored this construct using the Attention Network Test (Fan et al., 2002; Rueda et al., 2004). One aspect of attention, executive attention, involves actively inhibiting a dominant response. On executive attention, the CWS showed a trend toward reduced efficiency, relative to the CWNS. In a subsequent study, Eggers, De Nil, and Van den Bergh (2013) examined inhibitory control using the go/no-go task of the Amsterdam Neuropsychological Tasks (de Sonneville, 2003). The go/no-go task requires participants to press a button in response to a symbol of a man running but not in response to a symbol of a man standing. On this task, CWS between the ages of 4;10 and 10;0 produced significantly more false alarms (pressing the button in response to the man standing) and premature responses than CWNS. However, the groups did not differ on the number of misses (not pressing the button in response to the man running). The mean reaction time for false alarms was also significantly shorter for children in the CWS group. The authors interpreted these findings as suggestive of CWS having "a less controlled response style,"

[1] Some conceptualizations of temperament include domains, such as attentional and inhibitory control, that are typically considered to be more cognitive in nature.

(Eggers et al., 2013, p. 6) with false alarms and premature responses indicating greater impulsivity. The authors noted that, as a group, the CWS performed quickly across items, regardless of errors on prior items. The more typical response, when one begins to produce errors, is to reduce speed to improve accuracy. Taken together, these two studies suggest that CWS may have reduced inhibition relative to CWNS.

In contrast to these findings, however, a recent event-related potential study by Piispala, Kallio, Bloigu, and Jansson-Verkasalo (2016) examined 11 CWS, ages 6;3 to 9;5, and 19 age-matched CWNS during production of the same go/no-go task as described previously (de Sonneville, 2003). The CWS had longer N2 and P3 latencies on the go trials (in response to the figure of the man running) than the CWNS, but not on the no-go trials. Given that there were no group differences on trials in which children were to inhibit their response, the findings were interpreted to suggest that the groups did not differ in inhibitory control. It should be noted that the children in each of these studies were older, with most in the school-age years. In addition, because the tasks were nonverbal, it is not possible to assess the role of verbal demands on performance.

Our study differs from previous work in two important ways. First, it is an examination of prepotent response inhibition, as defined by Friedman and Miyake (2004), using a Stroop-like component. In particular, we began tasks by establishing that the dominant button press response was, in fact, the dominant response for a particular child. Once established, we instructed the child to respond with the subdominant response. Thus, the study is an attempt to measure complex inhibition using experimental tasks similar to or identical to those used in the extant child development literature.

The second difference relates to the overarching rationale for this study, which is to explore whether a specific domain-general process, complex response inhibition, is linked to verbal demands in CWS. The study was undertaken to investigate not only whether CWS might differ from CWNS in response inhibition but also the extent to which performance might be linked to the verbal demands of the tasks. Therefore, we used both explicit and implicit verbal response inhibition tasks, which enables us to examine the effect of variations in linguistic demand on inhibition processes. It is the first study of CWS, to our knowledge, to explore response inhibition while taking into account linguistic processing demands.

Purpose of the Study

The purpose of this study was to examine the accuracy and processing speed of response inhibition in preschool-age children who do and do not stutter using both explicit and implicit verbal tasks. We also examined the effect of working memory on children's performance in the two tasks and parents' perceptions of children's response inhibition in everyday settings. The main hypothesis was that CWS would perform less accurately and respond more slowly than CWNS on both the explicit and implicit verbal response inhibition tasks. In addition, we hypothesized that children's performance on the two tasks would be associated with children's inhibition abilities in everyday life situations, as reported by their parents.

Method

Participants

Participants were two groups of 41 children (N = 82) between the ages of 3;1 and 6;1 who do (CWS) and do not (CWNS) stutter. All children spoke American English as their primary language. None of the children had a history of neurological, hearing, intellectual, or speech-language (other than stuttering) problems per parent report and examiner observation and testing. Children were identified for participation by their parents who had heard about the study through newspaper or magazine advertisements, posted flyers, and referrals from other parents, speech-language pathologists, or preschool and daycare centers.

Group Classification Criteria

Children were classified as CWS or CWNS based on the frequency of stuttered disfluencies (part-word repetitions, single-syllable word repetitions, sound prolongations, or blocks; Yairi & Ambrose, 2005; cf. Pellowski & Conture, 2002) produced during a parent–child conversational interaction. Stuttered disfluencies occur with much greater frequency among individuals who stutter compared with those who do not stutter, and thus they are the primary means by which the two groups of individuals are differentiated (Yairi & Ambrose, 2005).

During the parent–child conversational interaction, children and their parent(s) conversed with one another for 20 to 30 minutes while seated at a small table with age-appropriate toys. A 300-word sample was obtained for each child and analyzed for stuttered disfluencies. Stuttering severity was also estimated for children in the CWS group using the Stuttering Severity Instrument-4 (SSI-4; Riley, 2009).

To be classified as a CWS, the children had to exhibit three or more stuttered disfluencies, on average, per 100 words of conversational speech and receive a total score of 12 or above on the SSI-4 (25 CWS were classified as mild and 16 as moderate). The mean stuttered disfluencies for the CWS was 6.31 (SD = 2.94) and the mean parent-reported time since stuttering onset (Yairi & Ambrose, 1992) was 18.18 months (SD = 11.40).

Children in the CWNS group were required to exhibit fewer than three stuttered disfluencies, on average, per 100 words of conversational speech. The mean stuttered disfluencies for CWNS was 0.57 (SD = 0.58). A Mann–Whitney test revealed that the CWS exhibited significantly more stuttered disfluencies than the CWNS, $z = -7.82$, $p < .001$, suggesting that the two groups of children had been appropriately classified by virtue of their stuttering behavior.

Group Matching Criteria

Children in the CWS group were matched to children in the CWNS group by age (± 4 months) and gender

(12 girls, 29 boys per group) at each data collection location. The CWS had a mean age of 52.32 months (SD = 11.29), and the CWNS had a mean age of 52.85 months (SD = 11.45), a nonsignificant difference, $t(80) = -0.21$, p = .83. The CWS and CWNS were also equated by family socioeconomic status using Hollingshead's Four-Factor Index of Social Position (Hollingshead, 1975), which takes both parental education and occupational status into account, yielding a family social position score from 8 (lower class) to 66 (upper class). The CWS had a mean social position score of 51.29 (SD = 13.64), and the CWNS had a mean of 48.83 (SD = 10.99), a nonsignificant difference based on the Mann–Whitney test, $z = -1.52$, p = .13.

Procedures

Testing was conducted at Indiana University and the University of Missouri. At each data collection location, children spent approximately 2 to 3 hr, during the course of two sessions, engaged in the following procedures: (a) a conversational interaction with their parents (described previously), (b) standardized speech-language testing and hearing screening, (c) a simple auditory detection task, and (d) explicit and implicit verbal response inhibition tasks. The children also participated in several other tasks unrelated to the present investigation. All experimental tasks were presented in random order across participants. In addition, parents completed the Children's Behavior Questionnaire–Short Form (Putnam & Rothbart, 2006) to obtain parent-report data on children's inhibition abilities in everyday life.

Speech-Language Tests and Hearing Screening

Children were administered four standardized, norm-referenced speech and language tests to ensure that their speech and language skills were typically developing. The following speech-language tests were administered: (a) Peabody Picture Vocabulary Test-4 (Dunn & Dunn, 2007), a receptive vocabulary measure; (b) Expressive Vocabulary Test-2 (Williams, 2007), an expressive vocabulary measure; (c) Test of Early Language Development-3 (Hresko, Reid, & Hammill, 1999), a receptive and expressive language measure; and (d) the Sounds-in-Words subtest of the Goldman-Fristoe Test of Articulation–2 (Goldman & Fristoe, 2000), a speech sound articulation measure. All participants received a standard score of 85 or higher (no lower than 1 SD below the mean) on all four speech and language tests, with no significant between-group differences observed based on a multivariate analysis of variance (p values = .43 to .95).

Children's hearing was also screened using bilateral pure tone testing at 20 dB HL for 1000, 2000, and 4000 Hz (American Speech-Language-Hearing Association, 1997). Of the 82 participants, 81 passed the hearing screening, suggesting that their hearing was within normal limits. One child, a female CWS, refused to complete the hearing screening, and thus the objective status of her hearing is unknown. However, the parents of this child had no concerns about her hearing, and the child did not exhibit any difficulty hearing during testing. For these reasons, the child was retained in the study.

Simple Auditory Detection Task

Children completed a simple auditory detection (SAD) task in which they pressed a button whenever they heard a 2000-Hz test tone. The SAD task was administered to ensure that the two groups of children were comparable in their basic auditory and motor processing abilities.

Children were seated in front of a computer and given the following instructions: "We are going to play a listening game. You will hear a 'beep' and your job will be to press the button as fast as you can when you hear the 'beep'. We are going to practice first so you can see how the game is played." After these instructions, children placed their hands on a mark in front of a response box and were presented with three live-voice trials in which the experimenter vocally simulated a beep and the child responded by pressing a single button. This was followed by three computer practice trials in which children pressed the button whenever they heard a 2000-Hz test tone. The practice trials were repeated until it was clear that the children understood the task. After practice, children completed the experimental task, which was identical to the computer practice, except that it contained 13 trials. Unlike the practice trials, however, children received no visual or verbal feedback on their performance during the experimental task. Each 2000-Hz test tone was presented for 1,000 ms, with an intertrial interval that varied randomly between 1,500 to 3,000 ms to reduce the predictability of the test tone.

The SAD task was developed using E-Prime v. 2.0 software by Psychology Software Tools, Inc. (PST; Sharpsburg, PA). A PST Serial Response Box, which features five button keys (all but one of the button keys were covered), was directly connected to the computer via the serial port. The latency of the child's response (i.e., reaction time [RT]) was measured in milliseconds from the onset of the auditory stimulus to the onset of the child's manual response and recorded onto the computer with the E-Prime software program.

Children's responses were recorded by the examiner as correct (child correctly responded to the test tone), commission errors (child responded before the test tone), and omission errors (child failed to respond). Responses that were delayed (child responded, but only after a delay) or contained a button press error (child responded, but failed to fully depress the button key) were also noted by the examiner. Main dependent variables were RT for correct responses and accuracy (number of correct responses).

Both groups of children performed well on the SAD task. Of 1,066 total trials (82 participants × 13 trials), there were only 16 omission errors (1.50%) and three commission errors (0.28%). The CWS (M = 12.73; SD = 0.92) produced, on average, slightly fewer correct responses than the CWNS (M = 12.80; SD = 0.84), but this difference failed to reach statistical significance, z = 0.71, p = .47. RT data for the SAD task were analyzed only for correct responses. RT values that were more than 2 SDs above or below the mean for each group were treated as outliers and, thus, eliminated from the final data corpus (see Ratcliff, 1993). This resulted in the loss of 4.88% of the RT data for the CWS and 6.03% for the CWNS. Mean RT on the SAD task was 1,256.89 ms (SD = 367.15 ms) for the CWS and 1,252.82 ms (SD = 440.84 ms) for the CWNS. An independent samples t-test revealed no significant differences in RT between the CWS and CWNS groups on the SAD task, $t(80) = .05$, $p = .96$. In general, these findings indicate that the CWS and CWNS were comparable in their performance on the SAD task, and thus if the two groups of children should differ on the response inhibition tasks, this difference cannot be easily attributed to differences in their basic auditory and motor processing abilities.

Explicit and Implicit Verbal Response Inhibition Tasks

A new computerized adaptation of the grass–snow task (Carlson & Moses, 2001), created using E-Prime v. 2.0 software, was used to examine explicit verbal response inhibition. Children were seated in front of a computer and instructed to place their hands on a mark in front of the PST Serial Response Box. The first button key on the response box was framed in green, and underneath it was a picture of grass, and the fifth button key was framed in white with silver snowflakes underneath it (with the three button keys in between covered). Children were then given the following instructions: "We are going to play a listening game. You will hear two words: 'grass' and 'snow.' When you hear the word 'grass,' you need to press the 'snow' button as fast as you can. When you hear the word 'snow,' you need to press the 'grass' button as fast as you can. We are going to practice first so you can see how the game is played." After the instructions, children completed three practice phases. Children first named the colors of grass and snow (color identification phase), and then they identified the grass and snow buttons (button identification phase). All of the children correctly identified both the colors and buttons. After the identification phase, children completed a training phase in which they heard each response stimulus (*grass, snow*) once and pressed the opposite button of the word they heard. The training phase was repeated until the child answered both correctly, with the examiner recording the number of practice trials needed. After the training phase, the children were presented with 16 experimental trials consisting of eight presentations of the word *grass* and eight of the word *snow*. Children pressed the grass button whenever they heard the word *snow* and the snow button whenever they heard the word *grass*.

Auditory stimuli (*grass, snow*) were digitally recorded by a male speaker and edited to 32-bit resolution at a sampling rate of 44 kHz. The intensity of the auditory stimuli was also equated for root-mean-square amplitude using Adobe Audition CS6 (Adobe Systems, Inc., San Jose, CA). Auditory stimuli were presented in a fixed random order, with the first four trials being identical to the last four trials. Each word was presented for 500 ms, with an intertrial interval of 2,500 ms. At the conclusion of the experiment, which lasted approximately 5 minutes, the children were asked to recall the instructions and their responses were coded as correct, incorrect, or no response.

Children's responses to the experimental trials were scored as correct (child produced the subdominant response), incorrect (child produced the dominant response), commission errors (child responded before the stimulus was presented), and omission errors (child failed to respond). Main dependent variables included RT for correct responses, which was measured from the onset of the auditory stimulus to the onset of the button press response; response accuracy (number of correct responses); and omission error rates.[2] Other variables included the number of practice trials needed, responses to the recall question, and response accuracy on the first and last four trials.

Implicit verbal response inhibition was measured using the newly developed baa–meow task. This task is similar to the grass–snow task, except that meaningful nonverbal stimuli (*meow, baa*) were used instead of explicit verbal stimuli, and pictures of an orange cat and white sheep were placed under the two button keys, which were framed in orange and white, respectively.

Nonverbal auditory stimuli (*meow, baa*) were obtained from the sound database of Marcell, Borella, Greene, Kerr, and Rogers (2000), each with a duration of approximately 1,000 ms, and equated for root-mean-square amplitude. For practice, children first identified the sounds that sheep and cats make (sound identification phase), and then they identified the cat and sheep buttons (button identification phase). All of the children correctly identified both sounds and buttons. After the identification phase, they completed a training phase in which they heard each auditory stimulus (*meow, baa*) once and were asked to respond by pressing the opposite button of the sound they heard. The training phase was repeated until the child answered both correctly. Children then completed 16 experimental trials in which they pressed the sheep button when they heard the *meow* (n = 8) and the cat button when they heard the *baa* (n = 8). The methods, procedures, coding, and data analysis for this task were otherwise identical to the grass–snow task described previously.

Parent-Reported Inhibition

The Children's Behavior Questionnaire-Short Form (CBQ-SF; Putnam & Rothbart, 2006), a caregiver report of children's temperament, was used to assess children's inhibition in a natural setting. Parents rated their children on 94 items using a 7-point Likert scale, with 1 being *extremely untrue of your child* and 7 being *extremely true of your child*. Scores for each item were averaged to form 15 temperament scales, but only the Inhibitory Control scale was used in the present study. *Inhibitory control* is defined by the CBQ-SF as the capacity to plan and refrain from inappropriate approach responses upon request or in

[2] Although commission errors were recorded, their rates were not statistically analyzed as they occurred with insufficient frequency.

novel or uncertain situations (Rothbart, Ahadi, Hershey, & Fisher, 2001). Higher scores on this scale indicate a stronger presence of the characteristic. Thus, higher inhibitory control scores reflect an increased ability to inhibit responses to irrelevant stimuli, whereas lower scores represent a decreased ability.

Statistical Analyses

Statistical analyses were conducted using IBM SPSS Statistics for Windows, Version 23 (IBM Corp., Armonk, NY), and SAS software, Version 9.4 of the SAS System for Windows (SAS Institute Inc., Cary, NC). Prior to analysis, data screening procedures (e.g., outliers, normality) were conducted for each descriptive and main dependent variable to determine whether they met the assumptions required for parametric tests. The normality assumption was assessed using the following procedures: (a) visual inspection of the distribution (e.g., histograms, P–P plots), (b) descriptive assessment of the data (e.g., skewness, kurtosis, dispersion), and (c) univariate normality testing (e.g., Shapiro–Wilk test).

For the descriptive measures, results of the normality assessment indicated that six of the nine continuous variables (chronological age, all four standardized speech-language test scores, and SAD RT) were normal; the remaining three variables (stuttered disfluencies, family socioeconomic status, and SAD accuracy) were non-normal. Independent-samples *t* tests and multivariate analysis of variance were, therefore, used to compare the mean values of the variables that approximated a normal distribution, whereas the Mann–Whitney *U* test was used to assess group differences in the non-normally distributed descriptive variables. The descriptive variables were also examined as potential covariates using Pearson's product-moment or Spearman's rank correlation coefficients. The correlational analyses revealed that chronological age was moderately to highly correlated with many of the dependent variables and, thus, was added as a covariate to most models to control for differences in chronological age. SAD RT scores were also moderately to highly correlated with RT in the two response inhibition tasks. Therefore, SAD RT scores served as a covariate for the RT analyses to ensure that the results were not driven by differences in basic motor RT.

For the main dependent measures, the normality assessment revealed that two of the four continuous variables (response accuracy and omission error rates) in each response inhibition task were not normally distributed, whereas the remaining two variables (RT data and mean response accuracy and RT difference scores) were normally distributed. Attempts to power transform the non-normal variables failed to correct the underlying distribution, and thus distribution free methods were used to analyze these variables. In particular, response accuracy for both response inhibition tasks was fitted using a generalized linear model with a gamma probability distribution, as this model is appropriate for positive, continuous, non-normal data and can accommodate covariates. The omission error rates, however, were skewed toward zero in each task and, thus, these data were analyzed using a zero-inflated Poisson regression model. The normally distributed variables—namely, RT and mean accuracy and RT difference scores—were evaluated using mixed-model analyses of covariance (ANCOVA) and ordinary least square (OLS) regression, respectively. Assumptions for each analysis were tested by, for example, examining the residual distributions and P–P plots and found to be adequate.

For the memory measures, Fisher's exact test was used, as appropriate, to compare percentages in each response inhibition task. Response accuracy on the first and last four trials of each task was analyzed using the Wilcoxon signed ranks test, as the data were not normally distributed or associated with any covariates. The number of practice trials in each task, which also failed to achieve normality, was analyzed using a generalized linear model with a gamma probability distribution. Finally, correlational analyses involving one or more non-normally distributed variables were analyzed using Spearman's rank (nonparametric) correlational coefficient, whereas Pearson's product-moment (parametric) partial correlation coefficients were used to analyze variables that were normally distributed.

Results

The purpose of this study was to assess explicit and implicit verbal response inhibition in CWS versus CWNS. This was accomplished by examining the accuracy and speed of children's responses on a computerized adaptation of the grass–snow task (Carlson & Moses, 2001) and the baa–meow task. The effect of working memory on children's performance was also analyzed, along with the relationship between response inhibition performance and parent-reported inhibition.

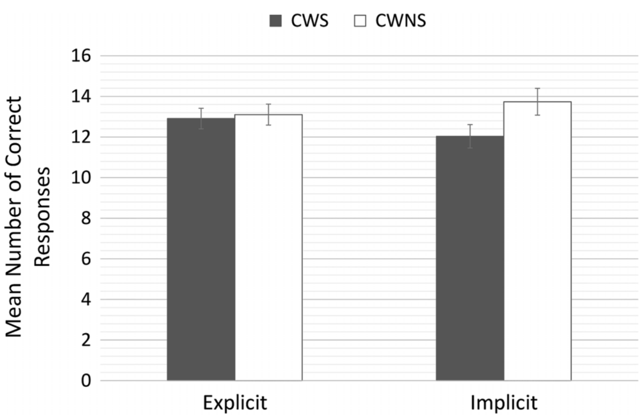
Response Accuracy

A generalized linear model was used to examine the relations between the dependent (response accuracy) and independent (group) variables. Separate analyses were performed for each response inhibition task, with chronological age serving as a covariate. Each analysis was conducted with a gamma probability distribution and log-link function.

For the grass–snow task, the omnibus test for the model fit was significant, $\chi^2(2) = 27.98$, $p < .001$, indicating that the model accounted for a significantly greater amount of the variance in response accuracy than the intercept-only model. The main effect of group was not significant, $\chi^2(1) = 0.07$, $p = .79$ (see Figure 1). Thus, contrary to expectation, the response accuracy of children in the CWS group (adjusted $M = 12.91$, $n = 41$) was comparable to that of children in the CWNS group (adjusted $M = 13.10$, $n = 41$). The main effect of the covariate was significant, $\chi^2(1) = 27.82$, $p < .001$, indicating that chronological age accounted for a significant amount of the variation in children's response accuracy.

For the baa–meow task, the omnibus analysis again indicated a significant overall effect, $\chi^2(2) = 16.04$, $p < .001$,

Figure 1. Adjusted mean (and standard error of the mean) number of correct responses for the preschool children who do stutter (CWS) and preschool children who do not stutter (CWNS) groups in the explicit response inhibition (grass–snow) and implicit response inhibition (baa–meow) tasks.



indicating that the fitted model was a better fit to the data than the intercept-only model. The main effect of group was significant, $\chi^2(1) = 3.81$, $p = .05$. As shown in Figure 1, the CWS (adjusted $M = 12.03$, $n = 41$) were significantly less accurate than the CWNS (adjusted $M = 13.74$, $n = 41$) on the task. Results further revealed a significant covariate main effect, $\chi^2(1) = 11.82$, $p = .001$, suggesting that at least some of the variation in children's response accuracy was accounted for by chronological age.

An OLS regression was used to examine whether group membership and chronological age significantly predicted the mean difference in response accuracy between tasks (difference = baa–meow accuracy − grass–snow accuracy). The results of the regression indicated that the two predictors explained 7% of the variance, $R^2 = .065$, $F(2, 81) = 2.74$, $p = .07$. Group membership significantly predicted the mean difference scores ($\beta = -1.53$, $p = .04$), but the same was not true of chronological age ($\beta = -0.04$, $p = .27$). These findings indicate that the mean difference in performance between tasks is, on average, 1.53 points lower for the CWS group, meaning that the performance of the CWS favored the grass–snow task, whereas the opposite was true for the CWNS.

Omission Error Rates

A zero-inflated Poisson regression model was used to analyze the relationship between group (CWS, CWNS) and the number of omission errors, controlling for chronological age, for each response inhibition task. For the grass–snow task, findings revealed that although the covariate, chronological age, significantly contributed to the model, $\chi^2(1) = 16.04$, $p < .001$, group did not, $\chi^2(1) = 0.35$, $p = .55$. Thus, children in the CWS group (unadjusted $M = 0.37$, $SD = .73$) were comparable to the children in the CWNS group (unadjusted $M = 0.34$, $SD = .97$) in the number of omission errors produced during the grass–snow task. Similar results were found for the baa–meow task. The covariate effect, $\chi^2(1) = 4.01$, $p = .04$, was significant, but not the main effect of group, $\chi^2(1) = 1.36$, $p = .24$ (CWS: unadjusted $M = 0.61$, $SD = 1.07$; CWNS: unadjusted $M = 0.20$, $SD = .60$).

Processing Speed

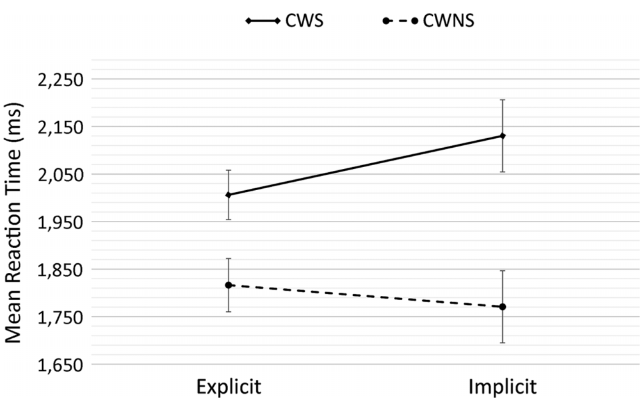
RT data were analyzed using a mixed-model ANCOVA with task (grass–snow or baa–meow) as a within-subject variable and group (CWS or CWNS) as a between-subject variable. Chronological age and SAD RT scores were included as covariates. Significant effects were followed up with post hoc analyses. As a measure of the strength of the association, the effect size indicator partial eta square (η_p^2) is reported for each statistical comparison, with .14 representing a large effect, .06 a medium effect, and .01 a small effect (Cohen, 1988).

Prior to analyzing the RT data, we removed correct responses that were 2 *SD*s above or below the mean (i.e., outliers) from the data corpus for each group of children in each task (see Ratcliff, 1993). This resulted in the removal of 6.92% of the data in the grass–snow task and 3.26% of the data in the baa–meow task. Following the removal of outliers, children who had fewer than eight (< 50%) useable RT responses were excluded from the analyses, along with their age- and gender-matched peers. This resulted in the removal of 12 pairs of children across both tasks ($N = 58$).

The mixed-model ANCOVA revealed a significant main effect of group, $F(1, 54) = 10.49$, $p = .002$, $\eta_p^2 = .16$, with the CWS performing more slowly across both tasks (adjusted M = 2,068.18 ms, $n = 29$) than the CWNS (adjusted M = 1,793.46 ms, $n = 29$). The main effect of task failed to reach significance, $F(1, 54) = 0.62$, $p = .44$, $\eta_p^2 = .01$ (grass–snow: adjusted M = 1,911.13 ms, $n = 58$; baa–meow: adjusted M = 1,950.51 ms, $n = 58$). Although the covariate main effect was significant for chronological age, $F(1, 54) = 3.93$, $p = .05$, $\eta_p^2 = .07$, and SAD RT, $F(1, 54) = 20.60$, $p < .001$, $\eta_p^2 = .28$, the interaction effects between these covariates and task were not significant: age: $F(1, 54) = 0.67$, $p = .42$, $\eta_p^2 = .01$; SAD RT: $F(1, 54) = 0.07$, $p = .80$, $\eta_p^2 = .001$.

Findings from the mixed-model ANCOVA further revealed a significant Task × Group interaction effect, $F(1, 54) = 4.24$, $p = .04$, $\eta_p^2 = .07$, indicating that RT performance on the two tasks significantly differed in the CWS and CWNS. To further examine this interaction, an ANCOVA test was used to analyze between-group differences in RT for each task, with chronological age and SAD RT scores as covariates. Findings revealed a significant between-group effect for the grass–snow task, $F(1, 54) = 5.73$, $p = .02$, $\eta_p^2 = .10$, as well as the baa–meow task, $F(1, 54) = 11.23$, $p = .001$, $\eta_p^2 = .17$. As shown in Figure 2, the CWS (grass–snow: adjusted M = 2,006.05 ms, $n = 29$; baa–meow: adjusted M = 2,130.30 ms, $n = 29$) responded significantly more slowly in both tasks than the CWNS (grass–snow: adjusted M = 1,816.21 ms, $n = 29$; baa–meow: adjusted M = 1,770.71 ms, $n = 29$). The covariate main effect of SAD RT was also significant in the grass–snow, $F(1, 54) = 24.90$, $p < .001$, $\eta_p^2 = .32$, and baa–meow, $F(1, 54) = 12.18$, $p = .001$,

Figure 2. Adjusted mean (and standard error of the mean) reaction time (ms) for the preschool children who do stutter (CWS) and preschool children who do not stutter (CWNS) groups in the explicit response inhibition (grass–snow) and implicit response inhibition (baa–meow) tasks.



$\eta_p^2 = .18$, tasks, whereas the covariate effect of chronological age approached significance: grass–snow: $F(1, 54) = 2.88$, $p = .09$, $\eta_p^2 = .05$; baa–meow: $F(1, 54) = 3.54$, $p = .06$, $\eta_p^2 = .06$.

A repeated-measures ANCOVA was also used to examine within-group differences in RT between tasks for each group of children, with chronological age and SAD RT scores serving as covariates. Findings indicated no significant difference in RT between the grass–snow and baa–meow tasks for the CWS, $F(1, 26) = 0.16$, $p = .70$, $\eta_p^2 = .006$, or CWNS, $F(1, 26) = 0.88$, $p = .36$, $\eta_p^2 = .03$. Although the covariate main effect for SAD RT was significant for both groups of children—CWS: $F(1, 26) = 7.70$, $p = .01$, $\eta_p^2 = .23$; CWNS: $F(1, 26) = 13.33$, $p < .001$, $\eta_p^2 = .34$—the covariate main effect for age failed to reach significance—CWS: $F(1, 26) = 1.45$, $p = .24$, $\eta_p^2 = .05$; CWNS: $F(1, 26) = 2.62$, $p = .12$, $\eta_p^2 = .09$. However, an analysis of the mean difference in RT between tasks (difference = grass–snow RT − baa–meow RT) using OLS regression revealed an altogether different finding. In particular, the OLS regression examined whether group membership, chronological age, and SAD RT scores significantly predicted the mean difference in RT between tasks. Results revealed that the three predictors explained 9% of the variance, $R^2 = .086$, $F(3, 57) = 1.70$, $p = .18$, and although group membership significantly predicted the mean difference scores ($\beta = -169.75$, $p = .04$), the same was not true of chronological age ($\beta = 4.01$, $p = .42$) and SAD RT ($\beta = 0.03$, $p = .80$). Thus, even though within-group differences across tasks were not statistically significant for either group of children, CWS were 169.75 ms slower in the baa–meow task when compared with the grass–snow task, whereas performance of the CWNS was in the opposite direction.

To examine the potential for speed-accuracy trade-offs, we evaluated the relationships between response accuracy and RT on the two response inhibition tasks using nonparametric (Spearman's rank) partial correlation coefficients, with chronological age and SAD RT serving as the covariates. For the CWS, analyses revealed no significant correlations between accuracy and RT on either the grass–snow ($r = .22$, $p = .27$) or baa–meow ($r = -.06$, $p = .78$) tasks. Similarly, for the CWNS, accuracy was not significantly correlated with RT on the grass–snow ($r = -.0005$, $p = .99$) or baa–meow ($r = .33$, $p = .09$) tasks. These findings suggest that the slower performance of the CWS on the two response inhibition tasks cannot be accounted for by a speed-accuracy trade-off.

Working Memory Effects

To examine the potential effect of working memory on children's performance in the baa–meow and grass–snow tasks, children's responses to the recall instruction question were analyzed along with response accuracy on the first and last four trials, and the association between number of practice trials and response accuracy (Montgomery & Koeltzow, 2010).

Recall Responses

A 2 × 2 Fisher's exact test (two-tailed) was used to examine the relationship between group (CWS and CWNS) and the response to the recall question (correct, incorrect) for each task ("no responses" were excluded from the analyses). Findings revealed that the percentage of children who correctly responded to the recall question did not differ by group for either the grass–snow, $p = .24$, or the baa–meow, $p = .99$, task. In fact, as revealed in Table 1, the majority of children (80.4% to 97.6%) in both groups were able to correctly recall the instructions after each experiment.

Performance Patterns

Children's response accuracy on the first four trials of each response inhibition task was compared with the last four trials using Wilcoxon signed ranks tests. For the grass–snow task, there was no significant difference in response accuracy between the first and last four trials for the CWS, $z = -1.58$, $p = .11$, and CWNS, $z = -0.69$, $p = .49$. Response accuracy was also comparable between the first and last four trials of the baa–meow task for the CWS, $z = -0.47$, $p = .64$ and CWNS, $z = -1.89$, $p = .06$.

Practice Trials and Response Accuracy

Spearman's rank partial correlations, with chronological age serving as a covariate, revealed no significant

Table 1. Number and (percentage) of correct, incorrect, and no responses to the recall question for the grass–snow and baa–meow tasks for the children who do stutter (CWS) and children who do not stutter (CWNS).

	Grass–snow task		Baa–meow task	
Variable	CWS	CWNS	CWS	CWNS
Correct	37 (90.2)	40 (97.6)	33 (80.4)	34 (82.9)
Incorrect	3 (7.3)	0 (0)	4 (9.8)	4 (9.8)
No response	1 (2.4)	1 (2.4)	4 (9.8)	3 (7.3)

associations between the number of practice trials and response accuracy on the grass–snow task for both the CWS ($r = -.24$, $p = .14$) and CWNS ($r = .01$, $p = .96$). Likewise, the relationship between the number of practice trials and children's response accuracy on the baa–meow task failed to reach significance for the CWS ($r = -.30$, $p = .06$) and CWNS ($r = -.01$, $p = .94$).

Although not a primary focus of this series of analyses, between-group differences in the number of practice trials were analyzed separately for each response inhibition task using a generalized linear model, with chronological age added as a covariate. Analyses were conducted using a gamma probability distribution and log-link function. The omnibus test was significant for both the grass–snow, $\chi^2(2) = 32.55$, $p < .001$, and baa–meow tasks, $\chi^2(2) = 29.01$, $p < .001$, indicating that each fitted model was a better fit to the intercept-only model.

For the grass–snow task, the main effect of group was significant, $\chi^2(1) = 5.90$, $p = .01$, with the CWS (adjusted $M = 3.36$, $n = 41$) requiring significantly more practice trials than the CWNS (adjusted $M = 2.77$, $n = 41$). The main effect of the covariate was also significant, $\chi^2(1) = 26.33$, $p < .001$, which indicates that a significant amount of the variation in the number of practice trials can be accounted for by chronological age. For the baa–meow task, both the main effect of group, $\chi^2(1) = 3.97$, $p = .04$, and the covariate, $\chi^2(1) = 25.61$, $p < .001$, were significant. As with the grass–snow task, children in the CWS group (adjusted $M = 3.07$, $n = 41$) required significantly more practice trials than CWNS (adjusted $M = 2.58$, $n = 41$) on the baa–meow task.

Correlational Analyses

The relationship between children's performance (response accuracy and RT) on the two response inhibition tasks and parent-reported inhibition (CBQ-SF) was evaluated using Spearman's rank and Pearson's product-moment partial correlation coefficients, with chronological age serving as the covariate. Bonferroni corrections were not applied to these analyses to minimize Type II error inflation (for commentary concerning the problem of applying corrections to hypothesis-driven experiments, see Curtin & Schulz, 1998; Garamszegi, 2006; Nakagawa, 2004; Perneger, 1998).

The relation between children's performance (response accuracy and RT) on the two tasks and the Inhibitory Control scale of the CBQ-SF was analyzed with both groups of children combined, as both groups of children showed a similar pattern of correlations. For the grass–snow task, analyses revealed significant correlations between the Inhibitory Control scale of the CBQ-SF and response accuracy ($r = .24$, $p = .05$) and RT ($r = -.35$, $p = .008$). Thus, children who were rated by their parents as having a stronger ability to inhibit responses to irrelevant stimuli in everyday life were not only more accurate but also faster on the explicit verbal response inhibition task. On the other hand, parents who perceived their children as having a decreased ability to inhibit responses in natural settings performed less accurately and more slowly on the explicit verbal response inhibition task.

For the baa–meow task, there was a significant correlation between the Inhibitory Control scale of the CBQ-SF and RT ($r = -.42$, $p = .002$), but not response accuracy ($r = .16$, $p = .17$). These findings indicate that, similar to the grass–snow task, parents who rated their children higher in inhibitory control performed faster on the implicit verbal response inhibition task, whereas those who were rated lower in inhibitory control responded more slowly. Unlike the grass–snow task, however, parents' perceptions of their children's inhibitory control abilities were not associated with the children's accuracy on the implicit verbal response inhibition task.

Discussion

Previous research using parent questionnaires has provided some evidence to suggest that inhibition may be reduced in CWS when compared with CWNS (Eggers et al., 2010; Embrechts et al., 2000; cf. Anderson & Wagovich, 2010). However, few studies have measured response inhibition directly in CWS, and no study to our knowledge has directly compared explicit verbal inhibition with implicit verbal inhibition to examine the role of linguistic demands on the ability to inhibit a dominant response. To explore explicit and implicit verbal inhibition in preschool-age CWS and CWNS, we administered two response inhibition tasks, both of which required suppression of a dominant response and the production of a subdominant response. The two tasks differed in that one (grass–snow task) contained explicit verbal material (words), whereas the other (baa–meow task) contained implicit verbal material (meaningful nonverbal sounds). These tasks, therefore, are more complex than the go/no-go task used in Eggers et al. (2013), in which children simply had to suppress a response depending on the stimulus (e.g., pushing a button when presented with a picture of a man running and not pushing the button when presented with a picture of a man standing; de Sonneville, 2003). Nevertheless, Eggers et al. (2013) did find differences between CWS and CWNS using the simpler go/no-go task, with CWS displaying not only more false alarms and premature responses than CWNS, but also shorter overall RTs, contrary to expectations. Thus, the present study builds on these investigations by using more complex Stroop-like inhibition tasks grounded in the developmental literature and by comparing performance on an explicit versus implicit verbal version of the task.

Response Inhibition Accuracy and Processing Speed

The results of the present study indicate that CWS are less effective than CWNS in their ability to inhibit a response tendency while engaging in a conflicting response. At first glance, this decrease in effectiveness would appear to be limited to the processing of meaningful nonverbal auditory stimuli, as response accuracy did not differ between the two groups of children on the explicit verbal task.

However, an analysis of the number of practice trials required for children to demonstrate task understanding revealed that, for both tasks, the CWS required more practice than the CWNS. Thus, even though the CWS were comparable to their peers in response accuracy on the explicit verbal task, learning the task did not come as easily for them as it did for the CWNS. Age was a significant factor in children's response accuracy on both response inhibition tasks, with older children outperforming younger children. In addition, older children in both groups required fewer practice trials to demonstrate understanding of the tasks. Thus, the tasks appear to be sensitive to children's development of inhibition skills during the preschool years. Indeed, other research has also demonstrated that inhibition and other EF skills develop during the preschool years (e.g., Best & Miller, 2010; Carlson, 2005; Gerstadt et al., 1994; Jurado & Rosselli, 2007).

RT results further indicate that CWS are less efficient (i.e., slower) than CWNS on measures of explicit and implicit verbal response inhibition, with no evidence of speed-accuracy trade-offs. In general, older children responded more quickly than younger children, and children who were faster in responding on the simple auditory detection task were also faster in responding on the inhibition tasks. These findings are partially consistent with those of Gerstadt et al. (1994); they too found age effects for both accuracy and RT for their young participants (ages 3.5 to 7 years), but with some evidence of speed–accuracy trade-offs among the youngest of their children.

As indicated at the outset, there is evidence to suggest that EF and language development are linked. For example, children with language impairments perform more poorly than their peers with typical language development on a variety of EF measures, including inhibition (for a review, see Kapa & Plante, 2015). Likewise, children with attention-deficit/hyperactivity disorder have been shown to exhibit weaknesses not only in EF, especially inhibition (Bishop & Norbury, 2005; Schoemaker et al., 2012), but also language ability (Leonard, 2014). The reciprocal connection between language and inhibition is relevant to the present study because in addition to weaknesses in inhibition, CWS have been shown to perform more poorly than CWNS on language processing measures (e.g., Anderson & Conture, 2004; Byrd, Conture, & Ohde, 2007; Hartfield & Conture, 2006; Pellowski & Conture, 2005). Thus, given the observed connection between language and response inhibition, both in children who are typically developing and among these clinical populations, it is possible that the slower or less accurate performance of the CWS on response inhibition is linked to subtle differences in language processing abilities.

The precise nature of the relationship between EF and language development is not clear. However, some investigators have speculated that inhibition is involved in lexical processing. Specifically, inhibition may be involved in suppressing competing lexical representations during selection (Mirman & Britt, 2014). For example, when a target word (e.g., *cat*) is activated by its conceptual features (e.g., tail, whiskers), other semantically related lexical representations (e.g., *rat, dog*), which share some of the same conceptual features as the target word, are also activated (Dell, Chang, & Griffin, 1999; Levelt, Roelofs, & Meyer, 1999). The outcome of the process is that the word that receives the most activation will be selected. To select the correct word, attention must be focused on the target word to enhance its activation while activation of semantically related lexical entries must be inhibited (Mirman & Britt, 2014; cf. Bishop, Nation, & Patterson, 2014). Thus, if a child has weak inhibition skills, he or she may experience more interference from lexical competitors during word production. Inhibition, however, could also be involved in the development of lexical representations because children must not only attend to and integrate relevant conceptual information from the environment when forming representations, but also inhibit nonrelevant information (Im-Bolter, Johnson, & Pascual-Leone, 2006). In other words, weaknesses in inhibition could directly affect the specificity of lexical representations, which could, in turn, negatively influence lexical access. In this way, a less than fully specified lexical representation may not receive sufficient activation during lexical access, making it more vulnerable to competition from its semantic neighbors. Inhibition could conceivably contribute to the development or retrieval of stable long-term phonological representations of words in much the same way. In the case of stuttering, the net effect of weaknesses in inhibition, regardless of whether it affects lexical or phonological representations or their retrieval, may be to make words more vulnerable to fluency disruption (Anderson, 2007; Anderson & Byrd, 2008).

Another potential way in which inhibition could affect language processing and, hence, speech fluency, is during speech monitoring (e.g., Henry, Messer, & Nash, 2015; Im-Bolter et al., 2006). Among the more current models of speech monitoring, the conflict-based account proposed by Nozari, Dell, and Schwartz (2011) suggests that errors are detected when conflict (or competition) occurs at the time of selection among activated words or phonemes. Thus, when a high degree of conflict occurs at the word or phoneme layers, the resulting conflict serves as a signal to the production system that something has gone awry. Nozari et al. (2011) further suggest that the conflict-based monitoring that occurs during language production is domain-general, meaning that conflict is the root of all monitoring, regardless of the domain (e.g., speech production, motor movements) in which the error is generated (cf. Piai, Roelofs, Acheson, & Takashima, 2013). When errors are detected by the speech monitoring system, speech is presumably interrupted as the errors are repaired (Levelt et al., 1999), which may result in disfluencies (Postma & Kolk, 1993). Inhibition may play a role in speech monitoring by prohibiting the expression of incorrect speech plans (e.g., activation of an incorrect word or phoneme). An overt expression of an error would, therefore, represent a failure of inhibition. Thus, weaknesses in inhibition could, theoretically, reduce the effectiveness (e.g., result in an increase in overt speech errors) or efficiency (e.g., incorrect speech plans take longer to suppress) of the speech monitoring system. In the latter case, weak inhibition would make it more difficult for children to suppress incorrect speech plans, but this suppression would eventually be accomplished. As a result, no increase in overt speech errors would be expected, but processing (suppressing an incorrect speech plan over a correct one) would take longer. Although most monitoring models are based on adults, children as young as 2 years of age reportedly have access to a speech monitoring system that allows them to respond to and correct errors (Levy, 1999). This system is thought to undergo further development during the preschool years (Jaeger, 2004; Rispoli, 2003) and is associated with improvements in language ability (Hanley, Cortis, Budd, & Nozari, 2016; Rispoli, 2003). Thus, it is conceivable that weaknesses in inhibition could have affected the speech monitoring systems of the young CWS in our study.

Explicit Versus Implicit Verbal Response Inhibition Performance

By comparing the explicit and implicit verbal response inhibition tasks, which were otherwise identical in inhibitory control and potential memory load, it is possible to consider the effect of linguistic demands on children's performance, as well as potential strategies children may use to aid performance. We note that, based on analyses of between-task differences in accuracy and RT, the CWS were more affected by the linguistic demands of the tasks, as they had significantly larger discrepancies in response accuracy and RT across tasks than the CWNS. In particular, the CWS tended to be slower and less accurate in the implicit verbal task (baa–meow) than the explicit verbal task (grass–snow), whereas the CWNS had the opposite tendency—slightly faster and more accurate in the implicit versus explicit task (see Figures 1 and 2).

Why were the CWS more affected by the linguistic demands of the tasks, especially when the verbal input was implicit? One potential explanation relies on the literature on word versus environmental sound processing. Although the processing of words and meaningful nonverbal sounds relies on many of the same underlying semantic, cognitive, and neural mechanisms (Saygin, Dick, & Bates, 2005), there are some important differences, which have been highlighted in a series of studies by Cummings et al. (Cummings & Čeponienė, 2010; Cummings, Čeponienė, Dick, Saygin, & Townsend, 2008; Cummings et al., 2006). Using a picture–word and picture–sound matching task, Cummings et al. found that children and adults tend to respond more accurately to meaningful nonverbal stimuli (environmental sounds) than verbal stimuli (words). Although they found no significant differences in RT between words and environmental sounds, event-related potentials indicated that the latency of the N400 effect appeared significantly earlier for environmental sounds than words. Cummings et al. (2006, 2008) interpreted these N400 findings to suggest that although semantic integration is initiated simultaneously for both environmental sounds and words, environmental sounds are either recognized more quickly than words or they directly activate semantic representations, and words must first undergo lexical processing. In addition, they argued that because the RTs were similar for both environmental sounds and words, it takes longer to transform recognized environmental sounds into behavioral responses than it does words, perhaps because meaningful sounds are encountered less frequently in everyday life and, thus, have weaker connections with response modalities than do words.

Cummings et al. (2008) also reported that the N400 effect for environmental sounds significantly decreased from childhood to adulthood, whereas the N400 effect for words did not change. Findings of a protracted course of development for the processing of environmental sounds but not words indicate, according to the authors, that the identification of the former may become more automatic over time. Specifically, they suggest that early in development, the processing of environmental sounds may require verbal labeling. Over time, however, the process of identifying and interpreting environmental sounds becomes more automatic, such that verbal labeling is no longer needed. The processing of words, on the other hand, does not undergo this developmental change because "…they are already well established by the age of 7 as the dominant, and automatic, device for semantic identification and interpretation" (Cummings et al., 2008, p. 9).

To place these findings in the appropriate developmental context for the present study, the youngest children studied by Cummings et al. (2008; Cummings & Čeponienė, 2010) were 7 years of age, whereas the youngest children in our study were 3 years of age. The studies by Cummings et al. also differed from the present study in the nature of the tasks (matching vs. inhibition) and the number of meaningful nonverbal and verbal stimuli used (multiple vs. limited). Nevertheless, similar to the older children in Cummings et al. (2008; Cummings & Čeponienė, 2010), there was a tendency for CWNS to perform more accurately on the implicit verbal task when compared with the explicit verbal task, with only a slight difference in RT between the two tasks. The CWS, on the other hand, were more affected by the different tasks and appeared to have found the implicit response inhibition task more challenging than their peers. If the processing of environmental sounds undergoes developmental change, it stands to reason that the processing abilities of the younger children in our study would be even less automatic than those of the youngest children in Cummings et al. Moreover, the greater difficulty the CWS had with the implicit verbal response inhibition task may reflect reduced automaticity relative to the CWNS. That is, it may be that the CWS were relying more on the earlier developing, less efficient strategy of verbally labeling the meaningful nonverbal auditory stimulus. Accordingly, the task would seem to be relatively easy when it involves two nouns (e.g., grass and snow), but it becomes more complex when the stimuli involve meaningful nonverbal sounds; upon hearing *baa*, one must convert that to the semantic representation of a sheep, and then press the learned opposite button, cat. In short, the mental effort needed if one is verbally labeling meaningful nonverbal stimuli is considerable, as can be seen from this example.

The Impact of Working Memory

Experimental tasks of any sort generally require participants to remember the directions of the task while completing it. Failure to remember the instructions will result in degraded performance at the point at which the person "forgets" the instructions. Thus, memory of instructions poses a potential validity threat to any EF task. For the present study, we analyzed the potential impact of memory of instructions in three ways. First, we simply asked children to recall the instructions at the end of the task. The majority of children were able to successfully recall the instructions after each experiment, and there were no significant between-group differences in the proportion of children who recalled the instructions correctly. These findings are consistent with those of Simpson and Riggs (2009), who found that 3-year-old children were able to remember the instructions for an inhibition task similar to our tasks.

Second, we compared children's response accuracy on the first four trials to the last four trials in each task. We reasoned that if children forgot the instructions during the course of the experiment, they would perform more poorly on the last four trials than on the first four trials. There was no evidence, however, to support this, as response accuracy at the beginning and end of each task was comparable for both groups of children. Similar findings have also been reported for the day–night task (Deák, & Narasimham, 2003; Simpson, Riggs, & Simon, 2004).

Finally, we examined the correlation between the number of practice trials needed and children's response accuracy on the two tasks. We reasoned, as did Montgomery and Koeltzow (2010), that the number of practice trials needed would be a direct reflection of how well the rule was held in working memory; children who need less practice ought to perform more accurately on the tasks (presumably because they had no difficulty remembering the instructions), whereas those who need more practice ought to perform more poorly on the tasks (presumably because they had difficulty remembering the instructions at the beginning of the task). Contrary to expectations, however, we found no significant correlations between the amount of practice needed and subsequent accuracy on the experimental trials for both groups of children. Thus, it would appear that the amount of practice needed had little bearing on their subsequent performance. These findings could be interpreted to suggest that memory does not play a significant role in children's performance on the task—that is, the strength with which these rules are held in memory does not affect their performance.

Taken together, findings from this series of analyses suggest that the working memory demands of the tasks were not likely to have significantly impacted children's performance or contributed to the between-group differences observed in response accuracy and processing speed.

Relationship Between Inhibition Task Performance and Parents' Perceptions of Inhibition Skills

Children's performance on the explicit and implicit verbal response inhibition tasks, in general, corresponded with parents' perceptions of their children's inhibition skills (as relates to behavior regulation). Items on this parent report measure include, for example, "Can easily stop an activity when s/he is told no." and "Can wait before entering into new activities if s/he is asked to." The similar pattern of performance across measures provides some evidence of concurrent validity for the two experimental tasks used in this study. In addition, it highlights the ecological validity of the tasks; if parent reports about their children's inhibition skills in daily life correspond to the children's task performance, the tasks themselves, it would seem, are a reasonable measure of the real-life skill of inhibition. Finally, the correspondence between parent report and the experimental task performance suggests that response inhibition, as measured in the present study, is a cognitive skill that can be reliably evaluated by parents.

Conclusion, Limitations, and Future Directions

In sum, the results of the present study suggest that response inhibition is an important aspect in describing the underlying nature of stuttering in early childhood. The CWS in this study performed less well than their peers in suppressing a prepotent response quickly and accurately in explicit and implicit verbal response inhibition tasks. The CWS also exhibited a significantly greater difference in performance between tasks, suggesting that they were more affected by the linguistic demands of the tasks. Strengths of the study include the relatively large sample of CWS, all of whom were determined to have typically developing language and articulation skills. Children were also similar in other demographic characteristics, including socioeconomic status. Although data were collected from two sites, the participants were matched from within sites so that the numbers of CWS versus CWNS contributed from each site were the same. From a methodological perspective, the explicit and implicit verbal inhibition tasks were designed to be identical other than the stimuli themselves (i.e., verbal or meaningful nonverbal stimuli) to enable comparisons across tasks. Tasks and procedures were based on those used within the literature on child cognition and development.

Limitations

There are several methodological issues that may warrant consideration. First, it should be noted that the CWS fell disproportionally in the mild range of stuttering severity. Thus, findings may not be representative of children who display a greater frequency of stuttering. Second, congruent conditions (i.e., nonconflict linguistic tasks) were not included in the experimental tasks for comparison purposes. Although it would have been ideal to include congruent conditions, doing so would have potentially compromised the integrity of the tasks. In particular, it would have added a response or task shifting component, which involves the ability to flexibly switch from one way of responding to another in response to a change in rules (Garon et al., 2008; Müller & Kerns, 2015). That is, children would have to switch from using a congruent rule—pressing the button that is associated with the word or sound they hear—to an incongruent rule—pressing the button that is opposite to the word or sound they hear—and vice versa. To successfully shift from using one set of rules to another, children must overcome proactive interference; that is, they must suppress the previously used rule in favor of the new rule (Baker, Friedman, & Leslie, 2010; Kiesel et al., 2010). Task shifting, however, has additional demands in that it requires the reconfiguration of task sets (i.e., switching to a new set of rules or response set) and a higher level of working memory demand (i.e., maintaining another set of rules or response set in working memory; Anderson & Reidy, 2012; Best & Miller, 2010; Diamond, 2016).

Numerous studies have shown that the consequence of switching from one task to another is that both RT and error rates increase (e.g., Verbruggen, Liefooghe, Szmalec, & Vandierendonck, 2005). A particularly relevant example of this comes from a study by Baker et al. (2010). In this study, 224 preschool children completed a task similar to the present study, consisting of two conditions—congruent and incongruent. In both conditions, children were shown two pictures and presented with a word. In the congruent condition, children had to point to the picture they heard, whereas in the incongruent condition, children pointed to the opposite picture. Each child completed two blocks in one of the following orders: congruent–congruent, congruent–incongruent, incongruent–incongruent, and incongruent–congruent. Thus, some children completed the congruent block first and then either did the congruent condition again or switched to the incongruent condition; the same was true for the incongruent condition. Findings revealed, not surprisingly, that the children performed significantly better in the congruent condition than the incongruent condition. Of particular interest to the present study, however, is that the children who switched rules in the second block performed significantly worse than the children who did not switch rules, regardless of whether they had to switch to the congruent or incongruent condition. This was especially true for the youngest (3-year-old) participants in the study. As one might expect, however, the children had the most difficulty when they switched to the incongruent condition.[3]

Thus, given the findings of Baker et al. (2010), we opted to err on the side of caution by not including congruent conditions in our experimental paradigm, as did many studies in the literature. Furthermore, because the tasks used in the present study have already been validated with young children as measures of inhibition, we deemed it to be an unnecessary addition. We did, however, control for any potential between-group differences in basic auditory and motor processing abilities by including the SAD task as a covariate in the statistical analyses.

[3]These findings also suggest that simply randomizing the congruent and incongruent conditions across participants may not alleviate the potential problem of proactive interference.

A third and related issue is whether the difference in performance between the CWS and CWNS was a consequence of the linguistic nature of the tasks, as opposed to inhibition. Given that CWS have been shown to perform more poorly than CWNS on a variety of language outcome and processing measures (e.g., Ntourou, Conture, & Lipsey, 2011), this is not an altogether unreasonable supposition. The addition of a linguistic congruent task (e.g., pressing the "grass" button when he or she hears the word *grass* and the "snow" button when he hears *snow*) would have best controlled for this possibility, but for reasons indicated previously, we did not include these tasks in our experimental design. However, there are several reasons why it is very unlikely that the between-group differences are a result of the language demands of the tasks: (a) there were only two stimulus items used in each task, and all four stimuli were acquired early in life and were high in frequency, concreteness, and imageability (i.e., from a linguistic standpoint; the words and sounds were not particularly challenging for preschoolers); (b) all children easily identified the stimulus items with 100% accuracy; (c) the two groups of children did not differ in their performance on four speech and language measures; and (d) there were no significant differences in RT across trials for either group of children in the grass–snow task—CWS: $F(15, 390) = 0.63$, $p = .76$, $\eta_p^2 = .02$; CWNS: $F(15, 390) = 1.68$, $p = .15$, $\eta_p^2 = .06$—and baa–meow task—CWS: $F(15, 390) = 0.88$, $p = .54$, $\eta_p^2 = .03$; CWNS: $F(15, 390) = 1.61$, $p = .12$, $\eta_p^2 = .06$. The latter point warrants further explanation. One well-documented finding from the language and memory literature is that adults and children tend to exhibit faster and more accurate performance on behavioral tasks with repeated presentations of the same or similar stimuli (e.g., Anderson, 2008). Thus, if the CWS were reacting to the language processing demands instead of inhibition, we should have seen a reliable improvement in their speed of performance across trials in the two tasks. For example, in the baa–meow task, if the child repeatedly translates a sound into an animal's name, then the connection between the sound and word would strengthen as a result of repeated activation during the duration of the task, resulting in faster RTs as the task progresses. Our data, however, suggest the opposite pattern of performance in both tasks; performance remains relatively stable. In sum, there is little evidence to suggest that the between-group differences in performance on the two tasks were a consequence of linguistic demands; rather, all evidence points toward inhibition.

Future Directions

Future investigations into response inhibition might explore nonverbal tasks that do not allow for a language work-around. This could be accomplished by using environmental sounds that represent natural opposites but for which labels are not readily available. Future work might also examine linguistic strategy use among CWS in the completion of inhibition tasks to explore the extent to which they use language in complex task completion and, when they do use it, whether it enhances performance. In addition, the examination of EF skills more generally in CWS is needed. Through careful, direct observation of these domain-general processes and potential linkages to language processes, it is possible to achieve a deeper understanding of the complexity of childhood stuttering.

Acknowledgments

This research was supported by a grant awarded to Julie D. Anderson from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health (R01DC012517).

References

American Speech-Language-Hearing Association. (1997). *Guidelines for audiologic screening*. Rockville, MD: Author.

Anderson, J. D. (2007). Phonological neighborhood and word frequency effects in the stuttered disfluencies of children who stutter. *Journal of Speech, Language, and Hearing Research, 50,* 229–247.

Anderson, J. D. (2008). Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter. *Journal of Fluency Disorders, 33,* 135–155.

Anderson, J. D., & Byrd, C. T. (2008). Phonotactic probability effects in children who stutter. *Journal of Speech, Language, and Hearing Research, 51,* 851–866.

Anderson, J. D., & Conture, E. G. (2004). Sentence structure priming in young children who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 47,* 552–571.

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.

Anderson, J. D., & Wagovich, S. A. (2010). Relationships among linguistic processing speed, phonological working memory, and attention in children who stutter. *Journal of Fluency Disorders, 35,* 216–234.

Anderson, P. J., & Reidy, N. (2012). Assessing executive function in preschoolers. *Neuropsychology Review, 22,* 345–360.

Arnold, H. S., Conture, E. G., Key, A. P., & Walden, T. (2011). Emotional reactivity, regulation and childhood stuttering: A behavioral and electrophysiological study. *Journal of Communication Disorders, 44,* 276–293.

Baker, S. T., Friedman, O., & Leslie, A. M. (2010). The opposites task: Using general rules to test cognitive flexibility in preschoolers. *Journal of Cognition and Development, 11,* 240–254.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81,* 1641–1660.

Bishop, D. V. M., Nation, K., & Patterson, K. (2014). When words fail us: Insights into language processing from developmental and acquired disorders. *Philosophical Transaction of the Royal Society B, 369,* 20120403.

Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology, 2: Response inhibition. *Autism, 9,* 29–43.

Byrd, C. T., Conture, E. G., & Ohde, R. N. (2007). Phonological priming in young children who stutter: Holistic versus incremental processing. *American Journal of Speech-Language Pathology, 16,* 43–53.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28,* 595–616.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72,* 1032–1053.

Chou, F. (2014). *Behavioral and electrophysiological observations of attentional control in children who stutter* (Unpublished doctoral dissertation). University of Iowa, Iowa City.

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

Cummings, A., & Čeponienė, R. (2010). Verbal and nonverbal semantic processing in children with developmental language impairment. *Neuropsychologia, 48,* 77–85.

Cummings, A., Čeponienė, R., Dick, F., Saygin, A. P., & Townsend, J. (2008). A developmental ERP study of verbal and non-verbal semantic processing. *Brain Research, 1208,* 137–149.

Cummings, A., Čeponienė, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115,* 92–107.

Curtin, F., & Schulz, P. (1998). Multiple correlations and Bonferroni's correction. *Biological Psychiatry, 44,* 775–777.

de Sonneville, L. M. J. (2003). *Amsterdam neuropsychological tasks*. Amstelveen, the Netherlands: SONAR.

Deák, G. O., & Narasimham, G. (2003). Is perseveration caused by inhibition failure? Evidence from preschool children's inferences about word meanings. *Journal of Experimental Child Psychology, 86,* 194–222.

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science, 23,* 517–542.

Diamond, A. (2016). Why improving and assessing executive functions early in life is critical. In J. A. Griffin, P. McCardle, & L. S. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment and translational research* (pp. 11–43). Washington, DC: American Psychological Association.

Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology, 38,* 352–362.

Dunn, L., & Dunn, L. (2007). *Peabody Picture Vocabulary Test – Fourth Edition (PPVT-4)*. Circle Pines, MN: American Guidance Service, Inc.

Eggers, K., De Nil, L. F., & Van den Bergh, B. R. (2010). Temperament dimensions in stuttering and typically developing children. *Journal of Fluency Disorders, 35,* 355–372.

Eggers, K., De Nil, L. F., & Van den Bergh, B. R. (2012). The efficiency of attentional networks in children who stutter. *Journal of Speech, Language, and Hearing Research, 55,* 946–959.

Eggers, K., De Nil, L. F., & Van den Bergh, B. R. (2013). Inhibitory control in childhood stuttering. *Journal of Fluency Disorders, 38,* 1–13.

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 H. G. Bosshardt, J. S. Yaruss, & H. F. M. Peters (Eds.), *Proceedings of the Third World Congress on Fluency Disorders: Theory, research, treatment, and self-help* (pp. 557–562). Nijmegen, the Netherlands: University of Nijmegen Press.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14,* 340–347.

Felsenfeld, S., van Beijsterveldt, C. E., & 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.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology, 133,* 101–135.

Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at Head Start. *Developmental Psychology, 47,* 404–416.

Garamszegi, L. Z. (2006). Comparing effect sizes across variables: Generalization without the need for Bonferroni correction. *Behavioral Ecology, 17,* 682–687.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134,* 31–60.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2 -7 years old on a Stroop-like day-night test. *Cognition, 53,* 129–153.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation – Second Edition (GFTA-2)* [Measurement instrument]. Circle Pines, MN: American Guidance Service, Inc.

Guy, J., Rogers, M., & Cornish, K. (2012). Developmental changes in visual and auditory inhibition in early childhood. *Infant and Child Development, 21,* 521–536.

Hanley, J. R., Cortis, C., Budd, M. J., & Nozari, N. (2016). Did I say dog or cat? A study of semantic error detection and correction in children. *Journal of Experimental Child Psychology, 142,* 36–47.

Hartfield, K. N., & Conture, E. G. (2006). Effects of perceptual and conceptual similarity in lexical priming of young children who stutter: Preliminary findings. *Journal of Fluency Disorders, 31,* 303–324.

Henry, L. A., Messer, D. J., & Nash, G. (2015). Executive functioning and verbal fluency in children with language difficulties. *Learning and Instruction, 39,* 137–147.

Hollingshead, A. (1975). *Four-Factor Index of Social Status*. New Haven, CT: Yale University.

Hresko, W., Reid, D., & Hammill, D. (1999). *Test of Early Language Development–Third Edition (TELD-3)*. Austin, TX: Pro-Ed.

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77,* 1822–1841.

Jaeger, J. J. (2004). *Kids' slips: What young children's slips of the tongue reveal about language development*. Mahwah, NJ: Erlbaum.

Johnson, K. N., Conture, E. G., & Walden, T. A. (2012). Efficacy of attention regulation in preschool-age children who stutter: A preliminary investigation. *Journal of Communication Disorders, 45,* 263–278.

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6,* 498–504.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review, 17,* 213–233.

Kaganovich, N., Wray, A. H., & Weber-Fox, C. (2010). Non-linguistic auditory processing and working memory update in pre-school children who stutter: An electrophysiological study. *Developmental Neuropsychology, 35,* 712–736.

Kapa, L., & Plante, E. (2015). Executive function in SLI: Recent advances and future directions. *Current Developmental Disorders Reports, 2,* 245–252.

Karrass, J., Walden, T. A., Conture, E. G., Graham, C. G., Arnold, H. S., Hartfield, K. N., & Schwenk, K. A. (2006). Relation of emotional reactivity and regulation to childhood stuttering. *Journal of Communication Disorders, 39,* 402–423.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin, 136,* 849–874.

Leonard, L. B. (2014). *Children with language impairment* (2nd ed.). Cambridge, MA: The MIT Press.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–38.

Levy, Y. (1999). Early metalinguistic competence: Speech monitoring and repair behavior. *Developmental Psychology, 35,* 822–834.

Ling, D. S., Wong, C.D., & Diamond, A. (2016). Do children need reminders on the day–night task, or simply some way to prevent them from responding too quickly? *Cognitive Development, 37,* 67–72.

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

Marcell, M., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology, 22,* 830–864.

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal lof Cognition and Development, 13,* 395–423.

Mirman, D., & Britt, A. E. (2014). What we talk about when we talk about access deficits. *Philosophical Transaction of the Royal Society B, 369,* 20120388.

Montgomery, D. E., & Fosco, W. (2012). The effect of delayed responding on Stroop-like task performance among preschoolers. *Journal of Genetic Psychology, 173,* 142–157.

Montgomery, D. E., & Koeltzow, T. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30,* 308–330.

Müller, U., Jacques, S., Brocki, K., & Zelazo, P. D. (2009). The executive functions of language in preschool children. In A. Winsler, C. Fernyhough, & I. Montero (Eds.), *Private speech, executive functioning, and the development of verbal self-regulation* (pp. 53–68). New York, NY: Cambridge University Press.

Müller, U., & Kerns, K. (2015). Development of executive function. In R. M. Lerner, L. S. Liben, & U. Mueller (Eds.), *Handbook of child psychology and developmental science: Vol. 7. Cognitive processes* (pp. 571–623). Hoboken, NJ: Wiley.

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology, 15,* 1044–1045.

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology, 63,* 1–33.

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.

Pellowski, M. W., & Conture, E. G. (2002). Characteristics of speech disfluency and stuttering behaviors in 3- and 4-year-old children. *Journal of Speech, Language, and Hearing Research, 45,* 20–35.

Pellowski, M. W., & Conture, E. G. (2005). Lexical priming in picture naming of young children who do and do not

stutter. *Journal of Speech, Language, and Hearing Research, 48,* 278–294.
Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal, 316,* 1236–1238.
Piai, V., Roelofs, A., Acheson, D. J., & Takashima, A. (2013). Attention for speaking: Domain-general control from the anterior cingulate cortex in spoken word production. *Frontiers in Human Neuroscience, 7,* 1–14.
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.
Pisoni, D. B., Conway, C. M., Kronenberger, W., Henning, S., & Anaya, E. (2010). Executive function, cognitive control, and sequence learning in deaf children with cochlear implants. In M. Marshark & P. E. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education* (Vol. 2, pp. 439–457). New York, NY: Oxford University Press.
Postma, A., & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research, 36,* 472–487.
Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment, 87,* 102–112.
Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510–532.
Riley, G. D. (2009). *Stuttering Severity Instrument for Children and Adults – Fourth Edition (SSI-4).* Austin, TX: Pro-Ed.
Rispoli, M. (2003). Changes in the nature of sentence production during the period of grammatical development. *Journal of Speech, Language, and Hearing Research, 46,* 818–830.
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.
Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologica, 42,* 1029–1040.
Saygin, A. P., Dick, F., & Bates, E. (2005). An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults. *Behavior Research Methods, 37,* 99–110.
Schoemaker, K., Bunte, T., Wiebe, S. A., Epsy, K. A., Dekovic, M., & Mattys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry, 53,* 111–119.
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.
Simpson, A., & Riggs, K. J. (2005). Inhibitory and working memory demands of the day–night task in children. *British Journal of Developmental Psychology, 23,* 471–486.
Simpson, A., & Riggs, K. J. (2009). What makes responses prepotent for young children? Insights from the grass-snow task. *Infant and Child Development, 18,* 21–35.
Simpson, A., Riggs, K. J., Beck, S. R., Gorniak, S. L., Wu, Y., Abbott, D., & Diamond, A. (2012). Refining the understanding of inhibitory processes: How response prepotency is created and overcome. *Developmental Science, 15,* 62–73.
Simpson, A., Riggs, K. J., & Simon, M. (2004). What makes the windows task difficult for young children: Rule inference or rule use? *Journal of Experimental Child Psychology, 87,* 155–170.
Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.
Tillman, C. M., Thorell, L. B., Brocki, K. C., & Bohlin, G. (2008). Motor response inhibition and execution in the stop-signal task: Development and relation to ADHD behaviors. *Child Neuropsychology, 14,* 42–59.
Verbruggen, F., Liefooghe, B., Szmalec, A., & Vandierendonck, A. (2005). Inhibiting responses when switching: Does it matter? *Experimental Psychology, 52,* 125–130.
Williams, K. T. (2007). *Expressive Vocabulary Test–Second Edition (EVT-2).* Circle Pines, MN: American Guidance Service, Inc.
Wolfe, C. D., & Bell, M. A. (2007). The integration of cognition and emotion during infancy and early childhood: Regulatory processes associated with the development of working memory. *Brain and Cognition, 65,* 3–13.
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. (2005). *Early childhood stuttering.* Austin, TX: Pro-Ed.
Zelazo, P. D., & Müller, U. (2010). Executive function in typical and atypical development. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 574–603). Oxford, UK: Wiley-Blackwell.