ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

What Is the Role of Questioning in Young Children's Fluency?

Julianne Garbarino[a] and **Nan Bernstein Ratner**[a]

[a] Department of Hearing and Speech Sciences, University of Maryland, College Park

ARTICLE INFO

Article History:
Received July 6, 2021
Revision received January 20, 2022
Accepted May 25, 2022

Editor-in-Chief: Erinn H. Finke
Editor: Nancy E. Hall

https://doi.org/10.1044/2022_AJSLP-21-00209

ABSTRACT

Purpose: Most therapy programs for young children who stutter (CWS) involve caregiver counseling and adjustment of caregiver behavior to maximize opportunities for the child to be more fluent. One component sometimes included as a recommended adjustment is a reduction in caregiver question asking, as question asking is hypothesized to increase language formulation demands on the child. However, there is limited research to guide clinician advisement to caregivers that has controlled for numerous potential confounding factors, including utterance length and grammaticality, that may impact potential stressors on children. Our aim was to assess whether there was an empirical basis for such recommendations by comparing disfluency profiles of answers to questions and nonanswer utterances produced by children during spontaneous play with parents and examiners.
Method: We analyzed fluency and structural properties as well as pragmatic function of 15,782 utterances from language samples produced by 32 CWS and 32 children who do not stutter (CWNS) who were between 28 and 50 months of age. CWS and CWNS were matched on gender and age within 4 months and were matched as closely as possible on maternal education.
Results: For utterances produced by CWS, answers to adult questions were significantly *less* likely to contain stuttering-like disfluencies than other utterance types, and this was still true after controlling for utterance length and grammaticality. In contrast, for utterances produced by CWNS, answers to questions were significantly *more* likely to be disfluent than other utterance types after controlling for length and grammaticality.
Conclusion: Given the current findings, some prior research, and the documented potential benefits in language development for adult question asking of children, we do not believe that clinicians need to recommend changes to typical question-asking behavior by caregivers of CWS.

A component of most published treatments for stuttering in young children is adjustment of caregiver behavior. This practice of advising caregivers how to talk to young children who stutter (CWS) dates back at least to Johnson's (1949) open letter to a mother of a child who stutters, originally published in *You and Your Child* magazine in April 1941. Johnson advises a mother to be cautious not to respond in any way to her child's disfluency in order to avoid causing or worsening stuttering (an advice we now know not to be well founded; Bloodstein et al., 2021).

Over the past 3 decades, some of the advice offered to caregivers has been amalgamated into the Demands and Capacities Model (DCM; Adams, 1990; Starkweather, 1987). Under the DCM, limitations in a child's ability to quickly respond to stimuli, coordinate movements of the speech mechanism, and/or plan and execute sequences of movements (i.e., capacity limitations) interact with internal and external demands to make fluency disruption more likely. Recommendations following the DCM are meant to increase capacities for fluent speech and/or decrease the demands on a child's speech production system. Examples of DCM-based adjustments include improvements in turn-taking behaviors within the family context and reductions in caregiver rate of speech, syntactic complexity of utterances (see the review by Bernstein Ratner & Guitar, 2006), and question-asking behaviors (Wilkenfeld & Curlee, 1997).

Surveys of clinicians over time (e.g., Cooper & Cooper, 1996; Crichton-Smith et al., 2003) suggest that

Correspondence to Julianne Garbarino: jgarbari@umd.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

the DCM has been immune to criticism that it is based on untested assumptions (e.g., Bernstein Ratner, 2000; Siegel, 2000). Advice that is consistent with its tenets continues to appear on self-help and informational websites, in all likelihood due to the limited literature available to clinicians in making therapy recommendations. While the DCM does not encompass all treatment programs for young CWS (Andrews et al., 2020; Onslow et al., 2021), the DCM continues to guide other treatment approaches such as RESTART-DCM and the Palin Parent–Child Interaction therapy (Franken & Laroes, 2021; Millard et al., 2008). We note that even the Lidcombe program suggests limiting the types of questions that parents ask during daily sessions to closed and binary choice questions in order to minimize formulation pressure on the child and maximize stutter-free speech (Onslow et al., 2021). Thus, the DCM has made inroads even into therapies not widely viewed as DCM motivated. In a review of studies published in the 1970s, 1980s, and early 1990s, Nippold and Rudzinski (1995) cautioned that there was limited evidence connecting caregiver speech behaviors, such as question asking, to increased disfluency rates in their children. Given the impact of the DCM in clinical preparation and practice, we believe that continued critical discussion of the therapy practices rooted in the DCM remains relevant.

The focus of this article is one recommendation based on the DCM: that caregivers of CWS reduce the frequency of their question asking as a way to reduce communicative stressors that may be associated with a child's stuttering rate. While there are no known negative influences on child language development when caregivers speak more slowly or increase turn-taking latencies, the recommendation that caregivers adjust the form and function of their input to the child has the potential to adversely impact child language development. Long-standing research in typical child language development specifically shows a major formative role of caregiver questions, as a form of grammatically rich child-directed speech, in language acquisition. For example, children's production of specific *wh*-question types has been linked to their frequency in their mothers' productions (Rowland et al., 2003). Furthermore, fathers' use of *wh*-questions predicts children's later vocabulary and reasoning skills (Rowe et al., 2017). Caregiver use of questions is not only linked to children learning to produce questions themselves; children who hear more maternal models of yes/no questions have shown greater growth in their use of auxiliaries (Newport et al., 1977). Last, adult questioning is also part of many caregiver–child routines such as shared book reading, which plays a critical role in supporting children's language, literacy, and knowledge development (Birbili & Karagiorgou, 2009; Zibulsky et al., 2019). Thus, assessing the impact of caregiver questions on the speech of CWS is particularly needed in order for clinicians to have access to evidence-based guidance for families of CWS.

Relatedly, recent research has found that children with higher rates of language growth are more likely to recover from stuttering (Hollister et al., 2017; Leech et al., 2017, 2019). Concurrently, research is growing steadily to show the critical role of caregiver verbal input in children's language growth (Hoff, 2003, 2013; Huttenlocher et al., 2002). Thus, it is appropriate to evaluate whether these traditional caregiver recommendations noticeably change the child's rate of stuttering and whether such changes provide a benefit beyond any negative impact that reduced use of caregiver questioning may have on the language development of CWS.

The recommendation of a reduction in caregiver questioning, based on the DCM, is widely available in both clinician manuals and websites designed to provide advice or "tips" to caregivers of young CWS. Our examples below are merely illustrative rather than exhaustive. For instance, the "7 Tips for Talking With Your Child" web page on the Stuttering Foundation of America's website suggests, "Asking questions is a normal part of life – but *try to resist asking one after the other* [emphasis added]. Sometimes it is more helpful to comment on what your child has said and wait" (Stuttering Foundation of America, n.d.). WebMD's web page about "How to Help a Stuttering Child" says, "*Minimize questions* [emphasis added] and interruptions when your child is speaking" (WebMD, n.d.). The American Speech-Language-Hearing Association (ASHA) web page with information for the public about stuttering says, "Indirect strategies are ways to help make it easier for your child to talk. These strategies can include slowing down your own speech and *asking fewer questions* [emphasis added]" (ASHA, n.d.). In addition to websites intended for the public, textbooks and treatment manuals for clinicians suggest reductions in questioning for caregivers of CWS. Yairi and Ambrose (2005) wrote, "The clinician encourages parents to identify a small change that may facilitate the child's fluency…. The fluency disrupting factors that are most prevalent during parent–child interaction are rapid speech rate, poor listening, inadequate turn-taking, interruptions, *repeated questioning* [emphasis added], adult response time latency, and syntactic/semantic complexity of parents' speech" (p. 437). With the RESTART-DCM method, an indirect approach described in depth in its manual, clinicians work with families to reduce demands (linguistic and otherwise) in a specific way tailored to the child's profile and the family's dynamic. Per the RESTART-DCM manual, a reduction of linguistic demands may be indicated if the parent asks many questions and makes few statements, and one possible change that the parent may make is "*mainly using statements (not questions)* [emphasis added] when speaking and using words which are appropriate, in terms of length and complexity, for the child's developmental age and language profile (according to the assessment)" (Franken & Laroes, 2021, p. 17). In providing

these examples, we wish to emphasize that many documented, effective treatments for children include numerous components such as the modification of parent–child interactions. Some of these sources (e.g., Franken & Laroes, 2021; Yairi & Ambrose, 2005) emphasize assessing the role of communicative factors within a child's profile and individualizing recommendations. We simply observe that these types of recommendations are considered common practice and flow naturally from a general orientation that views the reduction of communicative stress on a child who stutters as beneficial for maintaining the child's fluency.

While there is limited research directly assessing the role of question answering in altering early profiles of stuttering frequency, several studies have provided information that aids us in hypothesizing about how stuttering events might be influenced in the context of children's answers to questions. In an early account questioning the relationship between caregiver questions and fluency, Weiss and Zebrowski (1992) enrolled eight CWS aged 48–127 months (4;0–10;7 [years;months]; M = 83 months [6;11]) and reported that they stuttered more often in assertive utterances (including requests for information, actions, clarifications, or attention; utterances that label or describe; and performatives such as jokes) than in responses to requests (including answers to questions). Assertive utterances were longer and more complex than responsive utterances (many of which, in their samples, were one-word responses to questions), so length or complexity factors could have accounted for this finding. With a sample of 20 CWS aged 34–69 months (2;10–5;9; M = 52 months [4;4]), Ryan (2000) found that his participants also stuttered more on statements, commands, corrections, and questions than on answers. This may have been due to answers tending to be short (often one word long in his samples). Similarly, Yaruss (1999) found, for 12 CWS aged 40–66 months (3;4–5;6; M = 55 months [4;7]), that children's questions, a type of assertive utterance, were more likely to contain stuttering-like disfluencies (SLDs) than declarative utterances but did not assess whether answers to questions had particular fluency properties. Wilkenfeld and Curlee (1997) conducted a study with a within-subject, alternating-treatments design, replicated with a total of three participants aged 29 months (2;5), 44 months (3;8), and 56 months (4;8). Their study was specifically designed to assess rates of stuttering of young CWS in response to questions and comments. In the A condition, the examiner, interacting with the CWS, produced only questions and requests for verbal responses. In the B condition, the examiner produced only comments. There was no association between condition and rate of stuttering, indicating that these CWS did not stutter more in response to questions.

Byrd et al. (2011) provided additional clarity about the potential relationship between an utterance's assertiveness or responsiveness and the likelihood of stuttering for 15 CWS aged 28–94 months (2;4–7;10; M = 58 months [4;10]). They found that while assertive utterances produced by CWS were more likely to contain SLDs than responsive utterances, this difference was no longer significant after accounting for utterance length and complexity. They broke utterances down into a greater number of categories than simply assertive or responsive, but their coding system did not specify whether each responsive utterance was an answer to a question or a response to a statement. Distinguishing between responses to statements and answers to questions while controlling for syntactic factors would allow us to better understand whether adult question asking is associated with the rate of stuttering and whether there is an evidence base for reducing these questions in order to assist with lowering a child's rate of stuttering. We believe that both clinicians and caregivers deserve that greater attention be paid to increasing the evidence base for treatment recommendations in early stuttering.

Readers may be wondering why it is necessary in 2022 to examine whether children's answers to adult questions are more or less disfluent than other utterances, given that evidentiary support for the DCM has long been discussed, the evidence that responsiveness is not associated with increased rates of stuttering, and Wilkenfeld and Curlee's (1997) findings specifically assessing child stuttering in response to adult questions and comments for three CWS. The answer is that, despite all of these, the recommendation of reduced use of adult questioning when talking with CWS has persisted. Our hope is that if the results of a well-controlled study of a large number of utterances from a larger sample of CWS are consistent with previous results, caregiver advisement can be streamlined to include those elements most likely to maximize children's fluency and remove those that are likely to be ineffective and run counter to advice to caregivers of other children in the early stages of language development.

We must make one final note about our aims in studying whether question answering is or is not associated with a reduction in SLD production. We are not suggesting that reducing a child's rate of SLD production is the only (or necessarily, best) measure of whether a treatment approach for young CWS is effective and evidence based. Still, it is important, for the reasons we have outlined above, that clinicians have access to data that enable a better understanding of whether adult questioning is or is not empirically associated with childhood disfluency rates.

Research Questions

Our research questions were as follows:

1. Do CWS produce SLDs more often when they are producing assertive or responsive utterances?

2. Do CWS produce SLDs more often when they are answering questions than in utterances that are not answers to questions?
3. Within responsive utterances produced by CWS, are answers to questions more likely to contain SLDs than responses to comments?

For each of these questions, we were interested in (a) the answer without controlling for syntactic factors known to impact fluency and (b) the answer when controlling for these factors. Addressing these questions without controlling for other factors would permit us to make clearer comparisons with previous research. Then, controlling for syntactic factors would allow us to better isolate and understand the impact of the utterance's function.

Method

This study involved a secondary analysis of previously collected language samples gathered as part of studies assessing child language and/or fluency. All original data collection was completed using institutional review board–approved procedures. The current analysis was determined exempt by the University of Maryland's Institutional Review Board.

Data Sharing

This study makes use of data sharing and secondary data coding and analysis. Among the benefits to using language samples from a data sharing platform is increased power due to larger numbers of participants than are typical of single-site studies. In order to reap the benefits of data sharing, small concessions are required; some information that researchers or readers may desire to know can be unavailable. We believe that the benefits of the opportunity to analyze large number of utterances (i.e., 7,914 fully intelligible utterances from CWS and 8,414 fully intelligible utterances from children who do not stutter [CWNS] or 16,328 fully intelligible utterances total) far outweigh the drawbacks.

Participants

Participants were 64 children, 32 CWS and 32 CWNS, who enrolled in one of the three original studies from which we drew our participants. Seventeen of the CWS came from the Ratner corpus, available at Fluency.Talkbank.org (Bernstein Ratner & Silverman, 2000; Miles & Bernstein Ratner, 2001; Silverman & Bernstein Ratner, 2002; Wagovich & Bernstein Ratner, 2007), which was collected in the mid-to-late 1990s. The other 15 CWS came from the first year of the longitudinal University of Maryland–Carnegie Mellon University (UMD-CMU) corpus (Bernstein Ratner & MacWhinney, 2018). The first-year data from these children were collected between 2016 and 2020 at the University of Maryland.

Each child who stutters was matched to a child who does not stutter whose language samples were collected in similar play-based contexts. CWS and CWNS were matched on age (within 4 months) and gender and as closely as possible on maternal education. Of the 17 CWS from the Ratner corpus, 10 were matched to CWNS from the same corpus, and seven were matched to CWNS from the Weismer corpus available at Talkbank.org (Ellis Weismer et al., 2013). UMD-CMU CWS were matched to UMD-CMU CWNS.

Of the CWS, 26 participated in language sampling with both an examiner and a caregiver; five, with a parent only; and one, with an examiner only. Of the CWNS, 25 participated with both an examiner and a caregiver; six, with a parent only; and one, with an examiner only. Previous research with young CWS has found similar levels of stuttering occurring in play-based samples collected with examiners and parents (Meyers, 1986; Yaruss, 1997). Given this previous research and that approximately 80% of the CWS and CWNS in this study interacted with both examiners and parents, familiarity was balanced, and differences in familiarity were unlikely to have impacted results.

Participants were between 28 and 50 months of age. Twelve were girls, and 52 were boys. All were monolingual English speakers. Maternal education levels were available for the 15 CWS and 15 CWNS from the UMD-CMU corpus. All of these CWS and CWNS had mothers with a college or graduate degree, with the exception of one child who stutters with a mother who had taken some college courses. Please see Table 1.

Diagnosis of Stuttering

Children were diagnosed as stuttering on the basis of parent report plus a clinical finding of 3% syllables stuttered. Five participants whose recordings included 1.2%–2.7% syllables stuttered were also included as CWS on the basis of the parent's report that the child was usually less fluent at home and the parent's description of blocks or prolongations of consonants in the child's typical speech. One of these five participants presented with a weighted SLD score above 4.5 on the FluCalc computations of their language samples, with weighted SLD scores above 4 being highly suggestive of stuttering (Ambrose & Yairi, 1999). CWS from the Ratner corpus were tested and recorded within 4 months of stuttering onset, and CWS from the UMD-CMU corpus were tested and recorded within 1 year of stuttering onset.

Testing for Typical Language Development

All participants were determined to have language skills within typical developmental limits. Available

Table 1. Description of participants.

Pair	Corpora	Gender	CWS		CWNS	
			Age	Intell. utts.	Age	Intell. utts.
1	UC-UC	Male	28	99	30	242
2	R-W	Male	28	183	29	254
3	R-W	Male	29	274	29	177
4	R-R	Male	29	276	29	167
5	UC-UC	Male	31	64	32	258
6	R-W	Male	31	248	31	147
7	R-W	Male	31	255	31	297
8	R-W	Male	32	48	32	196
9	R-R	Male	32	120	32	267
10	R-R	Female	32	139	31	183
11	UC-UC	Female	32	237	33	357
12	UC-UC	Male	32	442	31	483
13	R-W	Male	33	230	33	107
14	R-R	Male	33	252	33	288
15	R-W	Female	33	348	32	238
16	UC-UC	Male	34	215	34	305
17	UC-UC	Female	34	272	30	277
18	UC-UC	Male	34	274	33	480
19	UC-UC	Female	34	470	35	451
20	R-R	Male	35	71	35	174
21	UC-UC	Male	35	338	35	323
22	R-R	Male	36	460	39	41
23	R-R	Male	37	284	37	147
24	UC-UC	Male	39	373	36	350
25	R-R	Male	40	137	40	279
26	UC-UC	Male	40	238	38	415
27	R-R	Male	41	65	41	240
28	R-R	Male	45	166	46	147
29	UC-UC	Male	46	352	48	319
30	UC-UC	Male	47	172	47	200
31	UC-UC	Female	47	298	45	285
32	UC-UC	Male	47	514	50	320

Note. Age is in months. Intell. utts. is the number of fully intelligible utterances produced by the child. R-W indicates that a child who stutters from the Ratner corpus was matched to a child who does not stutter from the Weismer corpus. R-R indicates that both participants were from the Ratner corpus. UC-UC indicates that both participants were from the University of Maryland–Carnegie Mellon University corpus. CWS = children who stutter; CWNS = children who do not stutter.

standardized test scores plus mean length of utterance in morphemes (MLU_m) were used to construct a language profile for each child. Each participant was required to have a majority of their scores within 1.5 *SD*s of the mean or higher. These guidelines were generally based on Tomblin et al.'s (1996) recommended criteria for diagnosis of language disorders.[1] While complete information about any previous diagnoses of these participants was not available due to the secondary analysis nature of the design, we can be reasonably certain that no participant presented with a language disorder based upon these exclusion criteria.

The three corpora used different sets of measures to construct this profile of participants' language skills. In the UMD-CMU corpus, available tests were the Peabody Picture Vocabulary Test–Fourth Edition (Dunn & D. M. Dunn, 2007) and the Clinical Evaluation of Language Fundamentals–Preschool, Second Edition Sentence Structure and Concepts and Following Directions subtests (Semel et al., 2004). In the Ratner corpus, available test scores for each child were some combination of the Peabody Picture Vocabulary Test–Revised (Dunn & L. M. Dunn, 1981), the Expressive One-Word Picture Vocabulary Test–Revised (Gardner, 1990), and the Clinical Evaluation of Language Fundamentals–Preschool Linguistic Concepts and Word Structure subtests (Wiig et al., 1992). In the Weismer corpus, the MacArthur Communicative Development Inventories Words and Sentences form (Fenson et al., 1993) and the Preschool Language Scale–Third Edition Auditory Comprehension and Expressive Language subtests (Zimmerman et al., 1992) were administered. MLU_m was based on all fully intelligible utterances so that it could be compared with dynamic norms available through the CHILDES (Child Language Data Exchange System) database (Bernstein Ratner et al., 2020; MacWhinney & Bernstein Ratner, 2022).

[1]According to Tomblin et al.'s (1996) criteria for diagnosing language impairment in kindergartners, children must score at least 1.25 *SD*s below the mean on two of five composite scores.

Language Sample Collection and Preparation

All CWS and CWNS participated in play-based language sampling tasks with examiners and/or caregivers, and transcripts documented these interactions. Transcripts were in CHAT (Codes for the Human Analysis of Transcripts) format, for analysis with CLAN (Computerized Language Analysis) language sample analysis software (MacWhinney, 2000). All transcripts were linked to either audio or video recordings of the interactions. All available adult and child utterances in the recordings were transcribed. As this study involved secondary data analysis, we cannot describe when during testing sessions recordings from the Weismer and Ratner corpora were taken. Recording in the UMD-CMU corpus began at the beginning of the testing sessions. The recordings ranged in length from 10.1 to 59.6 min (M = 10.1, SD = 13.0). As shown in Table 1, individual transcripts ranged from 41 to 514 fully intelligible child utterances (M = 255.1, SD = 113.9).

Disfluency Coding

SLDs and typical disfluencies (TDs) were coded using CHAT's fluency coding conventions (Bernstein Ratner et al., 2020). SLDs coded were part-word repetitions, prolongations, blocks, broken words, and monosyllabic word repetitions.[2] As noted by Yairi and Ambrose (2005), such events are present, albeit at lower frequency and without obvious tension or struggle, in the speech of all children learning to produce spoken language. TDs coded were fillers "um" and "uh," word fragments, multisyllabic word repetitions, multisyllabic word revisions, phrase repetitions, and phrase revisions. SLDs and TDs could be coded on the same word (or phrase) if both were observed.

Some transcripts included repetitions of (a) a person's or an animal's name (e.g., "Pig pig where are you?"); (b) "look," "hey," or "oh" (e.g., "Hey hey that's the dragon wing"); (c) "yes," "no," or another word marking agreement or disagreement (e.g., "No no you're the doctor"); or (d) an adjective repeated for emphasis (e.g., "He's a big big cat"). When there was a repetition of one of these types and intonation that made the intentionality of the repetition for emphasis markedly evident, the utterance was classified as fluent unless it contained another disfluency. This decision was made to avoid overmarking disfluencies in cases of intentional repetition and impacted 99 out of the 16,328 fully intelligible utterances.

Segmentation Procedure

Samples were segmented into utterances based on the "two out of three criteria" rule described by Bernstein Ratner et al. (2020). It is based on Stockman's (2010) finding that the use of at least two criteria in boundary placement results in more reliable segmentation decisions. Under this rule, an utterance boundary is placed when two of the following three criteria are met: (a) a perceptible pause, (b) a terminal intonation contour, and (c) a complete grammatical structure. Also following the work of Bernstein Ratner et al., preposed elements such as "yes" were joined to the main utterance unless they met two out of the three criteria. In addition, similar to Rispoli (2003), a maximum of two independent clauses connected by a coordinating conjunction could occur in one utterance. After this, another utterance began.

Consensus Reliability Procedure

For morpheme-by-morpheme transcription, segmentation, and disfluency coding, a consensus reliability procedure was developed, based on those used in previous language sample analysis studies assessing disfluency (Hollister et al., 2017; Rispoli, 2003). There were two passes through the transcript and then a consensus meeting with a third coder to discuss any locations where the second-pass coder disagreed with the first-pass transcription.

First pass. Because of the high number of people involved in initial transcript creation across the three original studies (over 25, as documented in transcript details), initial transcripts from TalkBank were considered a starting point. To increase consistency, the first author reviewed and corrected all transcripts to complete the first pass. Beginning with the completion of this first pass, all further checking and coding was done by our lab team.

Second pass. Second passes were completed by undergraduate research assistants (RAs) who had not worked on the original transcription projects. This pass involved checking morpheme-by-morpheme transcription (i.e., whether the words and morphemes transcribed matched what was heard), utterance segmentation, and

[2]Some readers might disagree with our decision to classify some disfluencies as SLDs even in the absence of perceptible tension (with the exception of utterances such as "Pig pig where are you?" where context and intonation made the intentionality of the repetition quite clear). We would agree with them (and Wingate, 2001, whose letter to the editor is titled "SLD Is Not Stuttering") that not all SLDs are moments of stuttering. We were faced with the need to balance the imprecision of listener perception of tension (Tichenor et al., 2017), with the fact that some nontense monosyllabic whole-word repetitions and part-word repetitions produced by CWS will probably not be true moments of stuttering. Given the discrepancies between speaker and expert observer ratings of tension (Tichenor et al., 2017), we chose to draw the line at classifying all part-word and monosyllabic whole-word repetitions as stuttering, in a decision we felt would increase reliability. However, requiring perceptible tension to classify a part-word or monosyllabic whole-word repetition as an SLD would be another reasonable approach in a study addressing the questions assessed in this study. We invite continued discussion about disfluency classification and any future research addressing questions similar to ours but requiring tension for SLD identification.

disfluency transcription. Before initiating second passes, RAs completed approximately 20 hr of training and passed two tests each for morpheme-by-morpheme transcription, segmentation, and disfluency transcription.

During the second pass, RAs marked any locations where they disagreed with the first-pass transcription on morpheme-by-morpheme transcription, segmentation, or disfluency coding. RAs were simply told that the first pass was created by a different transcriber to avoid their deferral to the first author's transcription decisions.

Consensus procedure. During consensus discussions, the second-pass and consensus raters discussed any sections in question to attempt to reach consensus. The mean numbers of marked utterances per participant brought to these discussions were 0.83 (SD = 1.34) for segmentation, 4.77 (SD = 4.24) for fluency coding, and 6.64 (SD = 6.30) for morpheme-by-morpheme transcription. Consensus was reached for 100% of the utterances.

Available Data Summary and Variable Coding

After consensus was achieved, utterances containing only a filler or word fragment and utterances in which the child trailed off were excluded. This left 15,782 utterances for further analysis. The following utterance-level factors were hand-coded or automatically computed.

Length

Length in morphemes was obtained using CLAN's automatic parser, namely, MOR. This parser has been found to be 94% accurate and is thought to result in utterance lengths that are more accurate and consistent than mean length of utterance computation by hand (Bernstein Ratner & MacWhinney, 2016).

Grammaticality

Utterances were coded as grammatical or ungrammatical based on whether the intended utterance (disregarding disfluencies) was acceptable in the adult grammar of the dialect(s) spoken by the child. Most participants spoke only mainstream American English (MAE) and were evaluated using protocols based on the work of Eisenberg et al. (2012). For participants who spoke both MAE and African American English (AAE; n = 5), utterances were considered grammatical if they were adultlike in either dialect (Oetting & MacDonald, 2001; Oetting & Pruitt, 2005).

Function

Function coding was based on the system used by Weiss and Zebrowski (1992) and Byrd et al. (2011), originally described by Fey (1986). Questions and nonquestion ("other") assertive utterances were both considered assertive utterances. Answers to questions and nonanswer ("other") responsive utterances were both considered types of responsive utterances. Utterances were assigned to one of the five categories described below.

Questions. This included questions with subject–auxiliary inversion (e.g., "Can you get this on?") as well as utterances with a declarative sentence structure in which intonation indicated that they were questions (e.g., "This one's full?"). It also included tag questions (e.g., "It's mine, right?").

Other assertive utterances. This code was given to any nonquestion utterance in which the child (a) brought up a new topic; (b) solicited information, an action, or attention; (c) labeled; (d) described facts or events; or (e) stated rules or explanations. For example, when not in response to an adult's utterance, "Sheep go on here," "Your name is 'cowboy,'" "I can open it," and "I need something else" would all have been coded as other assertive utterances.

Answer to a question. Any utterance in which a child was answering an adult's question was given this code. In line with Fey (1986), the child's utterance must have been an attempt to answer the question to receive this code, but the child's answer was not required to be correct. For instance, if the child was holding a toy car and an adult asked, "Is that a car?" and the child said, "Yes," "No," or "Yeah but it's not fast," all of these would have been coded as answers. However, if the child said, "Cars go fast," this would have been coded as an assertive utterance.

Other responsive utterances. This code was used for any utterance, other than an answer to a question, in which a child was responding to an adult. It was also used for utterances in which the child repeated the previous adult utterance or their own immediately preceding utterance. This included any utterances that were statements of agreement with the adult's preceding declarative utterance (e.g., if the adult said, "Your Lego tower is big," and the child said, "Yeah it is") and responses to requests for information (e.g., if the adult said, "Tell me which truck you want," and the child said, "the green one"). If the adult's utterance had been "Which truck do you want?" and the child had responded in the same way, this would have been coded as an answer to a question.

Unclear function. This code was used sparingly and only in cases in which it was unclear what preceded the child's utterance, and this may have impacted function coding. For example, if the child was the first person to speak after recording began and this made it difficult to determine the function of that first utterance, "unclear" was coded. These utterances were excluded from the analysis.

Coding Reliability

Because length was computed automatically through CLAN, the reliability of this variable depended primarily on accurate morpheme-by-morpheme transcription and utterance boundary placement. These were checked through the consensus transcription reliability process.

Grammaticality was hand-coded by RAs and underwent the consensus reliability procedure. Before participating in grammaticality coding for MAE-only speakers, RAs passed two tests of grammaticality coding for MAE speakers. RAs who worked on grammaticality coding in transcripts where the child spoke AAE and MAE received additional training in AAE features. This additional training involved synchronous instruction with the first author, reading and studying an internal lab manual describing AAE features, and further discussion and opportunities to ask questions. Any structure described as a feature of AAE when used by an African American child was considered grammatical if there was any doubt. These coders then passed two tests of grammaticality coding for bidialectal speakers. The first author completed the second pass and marked any utterances on which she disagreed with the first-pass decision. Finally, a third coder reviewed utterances on which there was disagreement between the two coders (and to avoid their deferral to the first author's coding, they were not told which coder chose which code). This process resolved 100% of the disagreements.

Function was also hand-coded and used similar reliability methods to those used for grammaticality. RAs who had passed two function coding tests completed the first function coding pass. The first author served as the second-pass coder and marked any utterances on which she disagreed with first-pass coding. A third coder, not told which coder had chosen which function, reviewed these utterances and resolved 100% of the disagreements.

Data Reduction

Utterances were excluded from the analyses if it would have been difficult to determine their intent. Excluded utterances were those that were (a) fully or partially unintelligible, (b) incomplete (i.e., interrupted or trailed off), or (c) composed only of fillers or fragments. There were 7,689 eligible utterances produced by CWS. Of these, three were coded as having unclear functions and so were excluded from the analyses, leaving 7,686 utterances remaining.

Planned Analyses

All analyses were run in R (Version 3.6.1; R Core Team, 2019) with the lme4 package (Version 1.1-21; Bates et al., 2015). The models used in this study were designed to assess how utterances produced by CWS of varied functions were associated with the odds of stuttering. To address each research question, we first ran a model assessing utterance function only, without controlling for linguistic predictors of stuttering. Next, models controlling for other linguistic predictors of stuttering were run. These controlled linguistic factors were utterance length and grammaticality. Length was controlled because of the well-established association between utterance length and disfluency (e.g., Buhr & Zebrowski, 2009; Logan & Conture, 1995; Zackheim & Conture, 2003). Grammaticality was controlled because it was a significant predictor of disfluency in the data set used here (Garbarino, 2021); the role of grammaticality will be more fully addressed in another paper currently in preparation. While syntactic complexity has also been associated with elevated disfluency (Buhr & Zebrowski, 2009; Logan & Conture, 1995; Zackheim & Conture, 2003), it was not a significant predictor of disfluency in this data set after controlling for length and grammaticality (Garbarino, 2021); hence, complexity was not included in these models. Because of the six models tested, .05 was divided by 6, and α was set to .008.

Question 1

The first question addressed was whether assertive and responsive utterances produced by CWS differed in the odds that they would contain at least one SLD. The first model run to address this question included a binary predictor representing whether an utterance was assertive (including questions and other assertive utterances) or responsive (including answers to questions and other responsive utterances) and a random intercept for participant. The second model added a binary grammaticality predictor and cluster-mean centered utterance length in morphemes as fixed factors. These models used the 7,686 fully intelligible utterances produced by CWS that could be coded as questions, other assertive utterances, answers to questions, or other responsive utterances. The mixed-effects models used to address this question and all other research questions do not assume equal numbers of data points within clusters (i.e., they do not assume equal numbers of utterances for each child; Maas & Hox, 2005).

Question 2

The next question was whether answers in the speech of CWS were more likely to contain at least one SLD than all other utterance types combined. The first model looking at this question included a binary function predictor (answer vs. any utterance that was not an answer) and a random intercept for participant. The second model addressing this question added a binary grammaticality predictor and cluster-mean centered utterance length in morphemes as fixed factors. These models used the same 7,686 utterances included in the models addressing Research Question 1.

Question 3

The third question concerned the two types of responsive utterances in the speech of CWS. Answers were compared with nonanswer responsive utterances. The first model assessing this question included a binary answer

Table 2. Characteristics of assertive and responsive utterances produced by children who stutter.

Utterance type	*n*	Length *M* (*SD*)	% ungrammatical	% SLD-containing
Assertive utterances				
Questions	898	3.4 (2.0)	22.4	27.3
Other assertive	2,894	3.9 (2.4)	22.3	23.6
Assertive total	3,792	3.8 (2.3)	22.3	24.5
Responsive utterances				
Answers	3,008	2.6 (2.2)	10.0	13.4
Other responsive	886	2.1 (1.7)	6.4	7.0
Responsive total	3,894	2.5 (2.1)	9.2	12.0
Total	7,686	3.1 (2.3)	15.7	18.1

Note. Length is in morphemes. SLD = stuttering-like disfluency.

versus other responsive utterance predictor and a random effect of participant. The second model looking at Research Question 3 added the fixed factors of a binary grammaticality predictor and cluster-mean centered utterance length in morphemes. These models used the 3,984 fully intelligible utterances produced by CWS coded as answers to questions or other responsive utterances.

Results

Descriptive Statistics

In this sample, 18.1% of all utterances produced by CWS contained at least one SLD, and the breakdown by utterance type is available in Table 2. With regard to length in morphemes, the mean length of all utterances produced by CWS was 3.1 (SD = 2.3), with mean lengths by utterance type provided in Table 2. Regarding grammaticality, 15.7% of all utterances were ungrammatical, and grammaticality by utterance type is reported in Table 2. In order to describe the sample without the impact of length, Table 3 shows similar descriptive statistics for the 1,030 utterances produced by CWS that were exactly three morphemes long.

Question 1

Without controlling for other factors, assertive utterances produced by CWS had 2.76 times higher odds of containing SLDs than responsive utterances did (z = 15.33, $p < .001$). Controlling for grammaticality and length, assertive utterances produced by CWS still had 1.74 times higher odds of having at least one SLD than responsive utterances did (z = 7.56, $p < .001$).

Question 2

Without controlling for grammaticality or length, utterances produced by CWS that were not answers had 1.98 times higher odds of containing at least one SLD than answers did (z = 9.97, $p < .001$). Controlling for grammaticality and length, utterances produced by CWS that were not answers still had 1.43 times higher odds of containing at least one SLD than answers did (z = 4.72, $p < .001$).

Question 3

Finally, the last two models compared answers with other responsive utterances. Without controlling for other

Table 3. Characteristics of three-morpheme assertive and responsive utterances produced by children who stutter.

Utterance type	*n*	% ungrammatical	% SLD-containing
Assertive utterances			
Questions	202	22.8	31.2
Other assertive	475	26.5	20.0
Assertive total	677	25.4	23.3
Responsive utterances			
Answers	285	15.4	19.6
Other responsive	68	11.8	13.2
Responsive total	353	14.7	18.4
Total	1,030	21.7	21.7

Note. SLD = stuttering-like disfluency.

factors, answers had 1.92 times greater odds of having at least one SLD than other responsive utterances did ($z = -4.50$, $p < .001$). Holding grammaticality and length constant, there was a trend toward answers from CWS having greater odds of having at least one SLD than other responsive utterances ($z = -0.42$ $p = .009$). Controlling for grammaticality and length, answers from CWS had 1.52 times greater odds of containing at least one SLD compared with other responsive utterances. See Table 4 for model output.

Additional Post Hoc Analyses

CWNS

One possibility is that the origin of the recommendation for caregivers to avoid asking questions was early clinicians' and researchers' observations of how disfluency related to question answering in children with typical levels and types of disfluencies in their speech. We wondered whether it might be the case that CWNS would be more disfluent when answering questions. Beyond our curiosity, however, there are theoretical motivations for conducting this type of analysis. In most cases, TD and stuttering relate to language in similar ways, as they both appear to relate to language planning. Increases in both TDs and SLDs are more likely to occur in longer and more complex utterances (e.g., Buhr & Zebrowski, 2009; Zackheim & Conture, 2003). Both types of disfluencies are produced more often at the beginnings of utterances (Bernstein, 1981; Buhr & Zebrowski, 2009; Choi et al., 2020; Gaines et al., 1991). Because of these similarities, it would be useful to compare SLDs produced by CWS and TDs produced by CWNS to gain additional information about the role of question answering in disfluency production more broadly.

Therefore, we also assessed whether disfluencies in the speech of CWNS increased in answers to questions. The questions addressed were as follows.

1. Are CWNS more disfluent when they are producing assertive or responsive utterances?
2. Are CWNS more disfluent when they are answering questions than in utterances that are not answers to questions?
3. Within responsive utterances produced by CWMS, are answers to questions more disfluent than responses to comments?

Six models were run with the same inputs as those run to assess SLD production in the speech of CWS. However, in these models, only utterances from CWNS were used, and the outcome variable was whether the utterance was

Table 4. Results from mixed-effects models of stuttering-like disfluency (SLD) production by utterance function for children who stutter.

Model and effect	*B*	*OR*	*z*	*p*	*B*	*OR*	*z*	*p*
Responsive vs. assertive		*Function alone*				*Function, grammaticality, length*		
Fixed effects								
Intercept	–2.31	0.10	–15.59	< .001	–2.45	0.09	–14.59	< .001
Function	1.01	2.76	15.33	< .001	0.55	1.74	7.56	< .001
Grammaticality					0.74	2.10	8.99	< .001
Length					0.38	1.46	23.42	< .001
Random effect	*Var.*				*Var.*			
Participant	0.59				0.76			
Answers vs. all other utts.		*Function alone*				*Function, grammaticality, length*		
Fixed effects								
Intercept	–2.18	0.11	–14.77	< .001	–2.38	0.09	–14.05	< .001
Function	0.68	1.98	9.97	< .001	0.36	1.43	4.72	< .001
Grammaticality					0.77	2.16	9.33	< .001
Length					0.39	1.48	24.40	< .001
Random effect	*Var.*				*Var.*			
Participant	0.56				0.75			
Answers vs. other resp.		*Function alone*				*Function, grammaticality, length*		
Fixed effects								
Intercept	–2.09	0.12	–13.87	< .001	–2.64	0.07	–14.70	< .001
Function	–0.66	0.52	–4.50	< .001	–0.42	0.66	–2.61	.009
Grammaticality					0.90	2.45	5.91	< .001
Length					0.44	1.56	17.12	< .001
Random effect	*Var.*				*Var.*			
Participant	0.55				0.75			

Note. Function coded as follows: 0 = all responsive, 1 = all assertive; 0 = answers, 1 = all other utterances; 0 = answers, 1 = other responsive utterances. Outcomes coded as follows: 0 = not SLD-containing, 1 = SLD-containing. The number of participants is 32. For the responsive versus assertive and answers versus all other utterances models, the number of utterances is 7,686. For the answers versus other responsive utterances model, the number of utterances is 3,984. *OR* = odds ratio; Var. = variance; utts. = utterances; resp. = responsive utterances.

fluent or disfluent. Simple fluency or disfluency was used because it is unclear whether there are qualitative distinctions between TDs and the SLDs produced by CWNS that would make a meaningful difference in the outcomes of these models. In other words, hallmarks of stuttering are tension and a feeling of being "stuck" (Ambrose & Yairi, 1994; Tichenor et al., 2017; Tichenor & Yaruss, 2019). To our knowledge, there is no evidence that CWNS experience tension or a feeling of being stuck when producing behaviors coded as SLDs, and therefore, we chose not to separate SLDs and TDs in these analyses. Furthermore, by definition, we did not observe audible tension in the disfluencies produced by CWNS; if we had observed tension, they would not have been in the CWNS group. For CWNS, there were 8,093 eligible utterances initially, six of which had unclear functions, leaving 8,087 coded utterances. These 8,087 utterances were used in the models run to address Questions 4 and 5. The 3,866 fully intelligible utterances from CWNS coded as answers or other responsive utterances were used in the models addressing Question 6. Again, α was set to .008 (.05 divided by 6) due to the six models tested.

Question 4

Without controlling for other factors, there were 1.24 times higher odds of an utterance being disfluent if it was assertive rather than responsive ($z = 3.21$, $p = .001$). In contrast, when controlling for grammaticality and length, there was a 1.27 times increase in the odds of an utterance being disfluent if it was responsive rather than assertive ($z = -3.25$, $p < .001$).

Question 5

Without controlling for other factors, there was a nonsignificant trend toward an increase in the odds that an utterance from a child who does not stutter would be disfluent if it was an answer compared with an utterance with another function ($z = -1.96$, $p = .050$). When controlling for grammaticality and length, answers from CWNS had 1.61 times greater odds of being disfluent compared with utterances that were not answers ($z = -6.52$, $p < .001$).

Question 6

Without controlling for other predictors, answers produced by CWNS had 4.17 times greater odds of being disfluent than other responsive utterances did ($z = -7.74$, $p < .001$). Finally, controlling for grammaticality and length, answers produced by CWNS still had 3.45 times greater odds of being disfluent than other responsive utterances did ($z = -6.67$, $p < .001$). See Table 5 for descriptive statistics and Table 6 for model output.

Types of Other Responsive Utterances

In the coding system discussed so far, all other responsive utterances were grouped together. The clinical question that this study aims to address is whether advising caregivers of CWS to make comments rather than ask questions is in fact effective at reducing the likelihood that a child's utterance will contain an SLD. There is one step toward better understanding this clinical issue that can be taken with the data available here: The nature of the utterances that children are responding to with their other responsive utterances could be examined. Are children responding to comments that caregivers may use in place of questions? Or are children responding to utterances that are not specifically intended to prompt the child to respond? To address this question, other responsive utterances were classified further based on the communicative interaction that they were a part of, as outlined below. The child's other responsive utterance itself was coded, but the preceding adult utterance and conversational

Table 5. Characteristics of assertive and responsive utterances produced by children who do not stutter.

Utterance type	*n*	Length *M* (*SD*)	% ungrammatical	% disfluent
Assertive utterances				
Questions	1,041	3.7 (2.1)	18.9	12.6
Other assertive	3,180	4.4 (2.5)	20.3	14.9
Assertive total	4,221	4.2 (2.4)	20.0	14.4
Responsive utterances				
Answers	3,045	2.7 (2.4)	7.9	14.6
Other responsive	821	2.1 (1.8)	4.0	4.3
Responsive total	3,866	2.6 (2.3)	7.1	12.4
Total	8,087	3.4 (2.5)	13.8	13.4

Note. Length is in morphemes.

Table 6. Results from mixed-effects models of disfluency production by utterance function for children who do not stutter.

Model and effect	*B*	*OR*	*z*	*p*	*B*	*OR*	*z*	*p*
Responsive vs. assertive		*Function alone*				*Function, grammaticality, length*		
Fixed effects								
Intercept	–2.15	0.12	–20.69	< .001	–2.11	0.12	–19.46	< .001
Function	0.22	1.24	3.21	.001	–0.24	0.79	–3.25	.001
Grammaticality					0.43	1.54	4.74	< .001
Length					0.25	1.28	18.02	< .001
Random effect	*Var.*				*Var.*			
Participant	0.24				0.26			
Answers vs. all other utts.		*Function alone*				*Function, grammaticality, length*		
Fixed effects								
Intercept	–1.95	0.14	–18.52	< .001	–1.94	0.14	–17.60	< .001
Function	–0.14	0.87	–1.96	.050	–0.48	0.62	–6.52	< .001
Grammaticality					0.44	1.56	4.83	< .001
Length					0.26	1.29	18.61	< .001
Random effect	*Var.*				*Var.*			
Participant	0.24				0.26			
Answers vs. other resp.		*Function alone*				*Function, grammaticality, length*		
Fixed effects								
Intercept	–2.01	0.10	–14.64	< .001	–3.18	0.04	–21.22	< .001
Function	–1.41	0.24	–7.74	< .001	–1.25	0.29	–6.67	< .001
Grammaticality					0.40	1.50	2.40	.016
Length					0.27	1.31	12.96	< .001
Random effect	*Var.*				*Var.*			
Participant	0.44				0.37			

Note. Function coded as follows: 0 = all responsive, 1 = all assertive; 0 = answers, 1 = all other utterances; 0 = answers, 1 = other responsive utterances. Outcomes coded as follows: 0 = fluent, 1 = disfluent. The number of participants is 32. For the responsive versus assertive and answers versus all other utterances models, the number of utterances is 8,087. For the answers versus other responsive utterances model, the number of utterances is 3,866. *OR* = odds ratio; Var. = variance; utts. = utterances; resp. = responsive utterances.

interaction as a whole were considered during the coding process.

Prompted. This code was used when the child responded to an utterance in which the adult was clearly attempting to get the child to respond. For instance, it was used when an adult said, "I wonder what this is," and the child's next utterance was, "That'll help us get on there."

Repetition. This was coded when the child's utterance was a partial or complete repetition of their own preceding utterance or the preceding adult utterance. For example, it was used when an adult said, "I think that's actually supposed to be pickles," and the child followed with "pickles."

Spontaneous. This was used when the child appeared to have decided on their own to say something in response to an adult, including spontaneous agreements, disagreements, and protests. It was used, for instance, when an adult said, "I think we should play with the castle," and the child replied, "No."

When the 886 other responsive utterances produced by CWS were coded using these definitions, 20 (2.3%) were prompted, 241 (27.2%) were repetitions, and 625 (70.5%) were spontaneous. Prompted other responsive utterances are the subtype of greatest interest because they are the ones likely to occur in response to an adult substituting a comment for a question. Given the small number of prompted other responsive utterances, no additional statistical analyses were performed.

Discussion

Utterance Function and SLD Production by CWS

The primary aim of this study was to evaluate the recommendation that caregivers should consider avoiding asking questions to CWS, so as to help CWS stutter less often. This recommendation is made both in materials intended to prepare clinicians to work with CWS (Franken & Laroes, 2021; Onslow et al., 2021; Yairi & Ambrose, 2005) and on websites aimed at caregivers and other family members of CWS (ASHA, n.d.; Stuttering Foundation of America, n.d.; WebMD, n.d.). Models were run to gain more information about the relationship between stuttering and the function of utterances produced by CWS. In general, conventional advisement regarding caregiver question asking was not supported by the results of our analyses.

First, we found that assertive utterances from CWS had 2.76 times greater odds of having at least one SLD compared with responsive utterances. This difference was

still significant after controlling for grammaticality and length.

In order to better assess whether answers were associated with higher levels of stuttering, answers were compared with all utterances that were not answers (assertive utterances plus other responsive utterances). For CWS, utterances that were not answers had 1.98 times (and significantly) greater odds of having at least one SLD compared with answers, and this continued to be significant when controlling for grammaticality and length. That answers have only about half the odds of containing an SLD compared with other utterances, without controlling for other factors, is a critical result in terms of indicating why a general advisement for caregivers to avoid asking questions to CWS may *not* result in decreased rates of stuttering.

Finally, answers were compared with other responsive utterances. When caregivers are instructed to avoid questioning, the proportion of the child's utterances that are responses to nonquestion comments is expected to increase. Therefore, comparing answers to other responsive utterances addressed whether child utterances that may replace answers are likely to be more fluent than answers. Answers produced by CWS had 1.92 times (and significantly) higher odds than other responsive utterances of containing SLDs, and this difference approached significance when controlling for grammaticality and length. These results might appear at first to suggest that if caregivers are trying to prompt children to talk, prompting children with statements that will result in nonanswer responsive utterances may be helpful. However, only 2.3% of other responsive utterances were prompted by adult statements of the type that might be used instead of questions. Given this small percentage, the decreased likelihood of stuttering in other responsive utterances compared with answers does not indicate to us that caregivers should avoid questions and instead make comments if they aim to reduce the frequency of stuttering.

This study's results are consistent with Byrd et al.'s (2011) finding that responsive utterances were less likely to be stuttered than assertive utterances, without controlling for other linguistic factors. However, current results are inconsistent with those in the work of Byrd et al. when linguistic factors were controlled. This difference may be due to power differences or differences in the age of children studied (with this study's children being younger on average). An examination of how age might impact the association between question answering and stuttering is beyond the scope of this study and may be of interest in future research.

Wilkenfeld and Curlee (1997) found, in a within-subject, alternating-treatments study replicated with three participants, that stuttering was not associated with whether the examiner was asking questions versus making comments; current results, based on a much larger set of data, are consistent with this finding. Our results are also consistent with those in studies reporting that responsive utterances or answers were less likely to contain stuttering than assertive utterances (Ryan, 2000; Weiss & Zebrowski, 1992). While it is common that caregivers are advised to reduce questioning when talking to CWS, we lack empirical reports indicating that CWS as a group tend to stutter more often when answering questions than when producing other utterance types.

When caregivers are advised to reduce their questioning of young children, an implicit assumption is made that the benefit to CWS from the supposed decrease in SLD production outweighs any benefit they would gain from being asked questions. However, hearing questions has positive effects on child language development; thus, we need to consider whether these benefits of caregiver advisement are sufficiently measured to ensure they achieve their intended goals. Being asked questions is associated with the acquisition of specific *wh*-question structures (Rowland et al., 2003), increased growth in vocabulary and reasoning skills (Rowe et al., 2017), and the acquisition of auxiliary verbs (Newport et al., 1977).

We wish to make one final note about the clinical recommendation of reductions in question asking for caregivers of young CWS. The first question clinicians may consider before asking parents to modify the communicative environment is whether less stuttering in any particular interaction is the primary aim of the intervention (rather than, for instance, managing frustration or establishing/maintaining positive beliefs about the child's communication skills). If reductions in stuttering are a therapeutic aim, then a blanket recommendation that caregivers should avoid questions is, according to our analysis, not likely to facilitate a child's fluency. As noted, we are not aware of any empirical studies showing *increased* stuttering for CWS when they are answering questions. Furthermore, growth in expressive language skills is associated with recovery from stuttering among CWS (Hollister et al., 2017; Leech et al., 2017, 2019), and limiting caregiver questioning may have unintended negative consequences for language development (Birbili & Karagiorgou, 2009; Newport et al., 1977; Rowe et al., 2017; Rowland et al., 2003; Zibulsky et al., 2019). With all of this said, clinicians may observe that for a particular CWS–caregiver pair, there may be certain question types or questioning styles worth adjusting for some period of time if they appear to provoke disproportionate amounts of stuttering behaviors and if reduction of such moments is a primary aim of the therapy plan.

Utterance Function and Disfluency Production by CWNS

What is the source of the clinical recommendation to reduce questions asked of CWS? Asking caregivers to

reduce questioning has been done by experts since at least the 1970s (Van Riper, 1973). One possibility is that this concern was derived from observing elevated disfluency rates for CWNS answering questions, and thus, we tested whether this was the case. Our results suggest that questioning interactions may be somewhat different in CWNS than in CWS and that it may have appeared logical to extend the reasoning to CWS. For CWNS, responsive utterances had 1.24 times (and significantly) higher odds of being disfluent compared with assertive utterances, after controlling for length and complexity. Answers to questions produced by CWNS had 1.61 times (and significantly) higher odds of being disfluent compared with other utterances, controlling for grammaticality and length. Finally, answers produced by CWNS, when compared with other responsive utterances, had 3.45 times higher odds of being disfluent, controlling for grammaticality and length.

Conflicting findings have driven continued debate about whether CWS and CWNS have differing temperamental profiles (see Bloodstein et al., 2021, for a review), but if these differences do in fact turn out to be present, they may partially explain the differing fluency patterns in the ways that the two groups answer questions. For instance, CWS have been reported to have higher behavioral inhibition than CWNS (Choi et al., 2013; Ntourou et al., 2020), and higher behavioral inhibition has been related to the production of fewer words in a language sampling context (Tumanova et al., 2020). CWS with higher behavioral inhibition may be more resistant to answering questions and, therefore, may choose to answer different types of questions than CWNS. This, in turn, may have impacted the profiles of fluency we observed here and suggests that future investigations that probe question-asking and question-answering patterns between CWS and CWNS and their parents on structural, functional, and temperament-focused levels of analysis would be informative.

Limitations

The generalizability of these findings may be limited to other children who share characteristics with these participants, specifically to *monolingual, English-speaking children without obvious language delays or disorders*. The maternal education level for this sample, when available, was also more advanced than the general population of American parents, and this may further limit generalizability.

That we used secondary data analysis in this study allowed us to assess a larger set of participants than previous studies looking at stuttering in answers to questions or responsive utterances, each of which enrolled three to 20 CWS (Byrd et al., 2011; Ryan, 2000; Weiss & Zebrowski, 1992; Wilkenfeld & Curlee, 1997; Yaruss, 1999). In order to reap these benefits, we made small concessions in that some information about participants and the original data collection that we or our readers may wish to know is not available. The tests used to ensure that participants did not have language disorders varied across studies. In addition, the retrospective nature of this study meant that researchers were unable to use screening measures, such as the Diagnostic Evaluation of Language Variation–Screening Test (Seymour et al., 2003) to definitively determine dialect use. We believe, however, that the benefits of the design far outweigh the limitations it introduced and that any impact on the results would have been minimal given that most p values were less than .001.

Last, while there are many factors that can influence a child's fluency, this study was focused on a particular set of predictors based on findings of previous studies assessing child fluency. This study is part of a larger project that will analyze additional predictors of child fluency.

Summary

In summary, these analyses evaluated disfluencies in children's answers to questions, as compared with other types of utterances, in a corpus of over 15,000 utterances. For CWS, our results showed significantly fewer SLDs in utterances that were answers to questions than in other types of utterances. Given our findings, previous similar findings in smaller studies, and the lack of any previous studies finding increases in stuttering produced by CWS in response to questions, we believe that clinicians do not need to recommend that caregivers of CWS need to restrict the language forms they use in conversation with these children.

Acknowledgments

This material is based upon work supported by National Science Foundation (NSF) Grant 1449815, awarded to Colin Phillips (PI); Julianne Garbarino was supported as a trainee on this grant. This work was also supported by National Institute on Deafness and Other Communication Disorders Grant R01DC015494, awarded to Nan Bernstein Ratner and Brian MacWhinney (Co-PIs), and NSF BCS-1626300/1626294, awarded to Nan Bernstein Ratner (Co-I) and Brian MacWhinney (Co-I). This study was completed by the first author in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Maryland. The authors would like to thank Allison Vance, Elizabeth Kim, Gabrielle Woska, Madison Burns, Kaitlin McCloskey, Rachel Raqueno, and Valerie Hsieh for their assistance

with this project and Natasha Cabrera, Rochelle Newman, Jared Novick, and Stacy Wagovich for their feedback on the study as members of the first author's dissertation committee.

References

Adams, M. R. (1990). The demands and capacities model I: Theoretical elaborations. *Journal of Fluency Disorders, 15*(3), 135–141. https://doi.org/10.1016/0094-730X(90)90014-J

Ambrose, N. G., & Yairi, E. (1994). The development of awareness of stuttering in preschool children. *Journal of Fluency Disorders, 19*(4), 229–245. https://doi.org/10.1016/0094-730X(94)90002-7

Ambrose, N. G., & Yairi, E. (1999). Normative disfluency data for early childhood stuttering. *Journal of Speech, Language, and Hearing Research, 42*(4), 895–909. https://doi.org/10.1044/jslhr.4204.895

American Speech-Language-Hearing Association. (n.d.). *Stuttering*. https://www.asha.org/public/speech/disorders/stuttering/

Andrews, C., Trajkovski, N., O'Brian, S., & Onslow, M. (2020). *The Westmead Program treatment guide* (Version 1.1). https://www.uts.edu.au/sites/default/files/2020-09/Westmead%20Program%20Treatment%20Guide_v1.0_Aug2020.pdf

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bernstein, N. (1981). Are there constraints on childhood disfluency? *Journal of Fluency Disorders, 6*(4), 341–350. https://doi.org/10.1016/0094-730X(81)90021-8

Bernstein Ratner, N. (2000). Performance or capacity, the model still requires definitions and boundaries it doesn't have. *Journal of Fluency Disorders, 25*(4), 337–346. https://doi.org/10.1016/S0094-730X(00)00083-8

Bernstein Ratner, N., Brundage, S. B., & Fromm, D. (2020). *A clinician's complete guide to CLAN and PRAAT*. https://talkbank.org/manuals/Clin-CLAN.pdf

Bernstein Ratner, N., & Guitar, B. (2006). Treatment of very early stuttering and parent-administered therapy: The state of the art. In N. Bernstein Ratner & J. Tetnowski (Eds.), *Current issues in stuttering research and practice* (pp. 99–124). Lawrence Erlbaum Associates.

Bernstein Ratner, N., & MacWhinney, B. (2016). Your laptop to the rescue: Using the Child Language Data Exchange System archive and CLAN utilities to improve child language sample analysis. *Seminars in Speech and Language, 37*(2), 74–84. https://doi.org/10.1055/s-0036-1580742

Bernstein Ratner, N., & MacWhinney, B. (2018). Fluency Bank: A new resource for fluency research and practice. *Journal of Fluency Disorders, 56,* 69–80. https://doi.org/10.1016/j.jfludis.2018.03.002

Bernstein Ratner, N., & Silverman, S. (2000). Parental perceptions of children's communicative development at stuttering onset. *Journal of Speech, Language, and Hearing Research, 43*(5), 1252–1263. https://doi.org/10.1044/jslhr.4305.1252

Birbili, M., & Karagiorgou, I. (2009). Helping children and their parents ask better questions: An intervention study. *Journal of Research in Childhood Education, 24*(1), 18–31. https://doi.org/10.1080/02568540903439359

Bloodstein, O., Bernstein Ratner, N., & Brundage, S. B. (2021). *A handbook on stuttering* (7th ed.). Plural.

Buhr, A., & Zebrowski, P. (2009). Sentence position and syntactic complexity of stuttering in early childhood: A longitudinal study. *Journal of Fluency Disorders, 34*(3), 155–172. https://doi.org/10.1016/j.jfludis.2009.08.001

Byrd, C. T., Coalson, G., & Bush, C. (2011). The communicative intent of stuttered utterances. *Journal of Interactional Research in Communication Disorders, 1*(2), 253–275. https://doi.org/10.1558/jircd.v1i2.253

Choi, D., Conture, E. G., Walden, T. A., Lambert, W. E., & Tumanova, V. (2013). Behavioral inhibition and childhood stuttering. *Journal of Fluency Disorders, 38*(2), 171–183. https://doi.org/10.1016/j.jfludis.2013.03.001

Choi, D., Sim, H., Park, H., Clark, C. E., & Kim, H. (2020). Loci of stuttering of English- and Korean-speaking children who stutter: Preliminary findings. *Journal of Fluency Disorders, 64,* 105762. https://doi.org/10.1016/j.jfludis.2020.105762

Cooper, E. B., & Cooper, C. S. (1996). Clinician attitudes towards stuttering: Two decades of change. *Journal of Fluency Disorders, 21*(2), 119–135. https://doi.org/10.1016/0094-730X(96)00018-6

Crichton-Smith, I., Wright, J., & Stackhouse, J. (2003). Attitudes of speech and language therapists towards stammering: 1985 and 2000. *International Journal of Language & Communication Disorders, 38*(3), 213–234. https://doi.org/10.1080/1368282031000086282

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test–Revised*. AGS.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Pearson Assessments.

Eisenberg, S., Guo, L.-Y., & Germezia, M. (2012). How grammatical are 3-year-olds? *Language, Speech, and Hearing Services in Schools, 43*(1), 36–52. https://doi.org/10.1044/0161-1461(2011/10-0093)

Ellis Weismer, S., Venker, C. E., Evans, J. L., & Moyle, M. J. (2013). Fast mapping in late-talking toddlers. *Applied Psycholinguistics, 34*(1), 69–89. https://doi.org/10.1017/S0142716411000610

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D. J., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *The MacArthur Communicative Development Inventory*. Singular.

Fey, M. (1986). *Language intervention with young children*. Allyn & Bacon.

Franken, M.-C., & Laroes, E. (2021). *RESTART-DCM method: Revised edition*. https://restartdcm.nl

Gaines, N. D., Runyan, C. M., & Meyers, S. C. (1991). A comparison of young stutterers' fluent versus stuttered utterances on measures of length and complexity. *Journal of Speech and Hearing Research, 34*(1), 37–42. https://doi.org/10.1044/jshr.3401.37

Garbarino, J. (2021). *Utterance-level predictors of stuttering-like, stall, and revision disfluencies in the speech of young children who do and do not stutter* [Doctoral dissertation, University of Maryland]. ProQuest Dissertations Publishing.

Gardner, M. F. (1990). *Expressive One-Word Picture Vocabulary Test–Revised*. Academic Therapy Publications.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368–1378. https://doi.org/10.1111/1467-8624.00612

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology, 49*(1), 4–14. https://doi.org/10.1037/a0027238

Hollister, J., Van Horne, A. O., & Zebrowski, P. (2017). The relationship between grammatical development and disfluencies in preschool children who stutter and those who recover. *American Journal of Speech-Language Pathology, 26*(1), 44–56. https://doi.org/10.1044/2016_AJSLP-15-0022

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*(3), 337–374. https://doi.org/10.1016/S0010-0285(02)00500-5

Johnson, W. (1949). An open letter to the mother of a stuttering child. *Journal of Speech and Hearing Disorders, 14*(1), 3–8. https://doi.org/10.1044/jshd.1401.03

Leech, K. A., Bernstein Ratner, N., Brown, B., & Weber, C. M. (2017). Preliminary evidence that growth in productive language differentiates childhood stuttering persistence and recovery. *Journal of Speech, Language, and Hearing Research, 60*(11), 3097–3109. https://doi.org/10.1044/2017_JSLHR-S-16-0371

Leech, K. A., Bernstein Ratner, N., Brown, B., & Weber, C. M. (2019). Language growth predicts stuttering persistence over and above family history and treatment experience: Response to Marcotte. *Journal of Speech, Language, and Hearing Research, 62*(5), 1371–1372. https://doi.org/10.1044/2019_JSLHR-S-18-0318

Logan, K. J., & Conture, E. G. (1995). Length, grammatical complexity, and rate differences in stuttered and fluent conversational utterances of children who stutter. *Journal of Fluency Disorders, 20*(1), 35–61. https://doi.org/10.1016/0094-730X(94)00008-H

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum Associates.

MacWhinney, B., & Bernstein Ratner, N. (2022). Dynamic norming and open science. *Journal of Speech, Language, and Hearing Research, 65*(3), 1183–1185. https://doi.org/10.1044/2022_JSLHR-22-00019

Meyers, S. C. (1986). Qualitative and quantitative differences and patterns of variability in disfluencies emitted by preschool stutterers and nonstutterers during dyadic conversations. *Journal of Fluency Disorders, 11*(4), 293–306. https://doi.org/10.1016/0094-730X(86)90017-3

Miles, S., & Bernstein Ratner, N. (2001). Parental language input to children at stuttering onset. *Journal of Speech, Language, and Hearing Research, 44*(5), 1116–1130. https://doi.org/10.1044/1092-4388(2001/088)

Millard, S. K., Nicholas, A., & Cook, F. M. (2008). Is parent–child interaction therapy effective in reducing stuttering? *Journal of Speech, Language, and Hearing Research, 51*(3), 636–650. https://doi.org/10.1044/1092-4388(2008/046)

Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 109–149). Cambridge University Press.

Nippold, M. A., & Rudzinski, M. (1995). Parents' speech and children's stuttering: A critique of the literature. *Journal of Speech and Hearing Research, 38*(5), 978–989. https://doi.org/10.1044/jshr.3805.978

Ntourou, K., DeFranco, E. O., Conture, E. G., Walden, T. A., & Mushtaq, N. (2020). A parent-report scale of behavioral inhibition: Validation and application to preschool-age children who do and do not stutter. *Journal of Fluency Disorders, 63,* 105748. https://doi.org/10.1016/j.jfludis.2020.105748

Oetting, J. B., & McDonald, J. L. (2001). Nonmainstream dialect use and specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(1), 207–223. https://doi.org/10.1044/1092-4388(2001/018)

Oetting, J. B., & Pruitt, S. (2005). Southern African-American English use across groups. *Journal of Multilingual Communication Disorders, 3*(2), 136–144. https://doi.org/10.1080/14769670400027324

Onslow, M., Webber, M., Harrison, E., Arnott, S., Bridgman, K., Carey, B., Sheedy, S., O'Brian, S., MacMillan, V., Lloyd, W., & Hearne, A. (2021). *The Lidcombe Program treatment guide*. https://www.lidcombeprogram.org/helpful-resources/helpful-downloads/

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rispoli, M. (2003). Changes in the nature of sentence production during the period of grammatical development. *Journal of Speech, Language, and Hearing Research, 46*(4), 818–830. https://doi.org/10.1044/1092-4388(2003/064)

Rowe, M. L., Leech, K. A., & Cabrera, N. (2017). Going beyond input quantity: *Wh*-questions matter for toddlers' language and cognitive development. *Cognitive Science, 41*(S1), 162–179. https://doi.org/10.1111/cogs.12349

Rowland, C. F., Pine, J. M., Lieven, E. V. M., & Theakston, A. L. (2003). Determinants of acquisition order in *wh*-xquestions: Re-evaluating the role of caregiver speech. *Journal of Child Language, 30*(3), 609–635. https://doi.org/10.1017/S0305000903005695

Ryan, B. P. (2000). Speaking rate, conversational speech acts, interruption, and linguistic complexity of 20 pre-school stuttering and non-stuttering children and their mothers. *Clinical Linguistics & Phonetics, 14*(1), 25–51. https://doi.org/10.1080/026992000298931

Semel, E., Wiig, E. H., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals Preschool–Second Edition*. Pearson Assessments.

Seymour, H. N., Roeper, T. W., & de Villiers, J. (2003). *Diagnostic Evaluation of Language Variation–Screening Test*. Ventris Learning.

Siegel, G. M. (2000). Demands and capacities or demands and performance? *Journal of Fluency Disorders, 25*(4), 321–327. https://doi.org/10.1016/S0094-730X(00)00081-4

Silverman, S., & Bernstein Ratner, N. (2002). Measuring lexical diversity in children who stutter: Application of *vocd*. *Journal of Fluency Disorders, 27*(4), 289–304. https://doi.org/10.1016/S0094-730X(02)00162-6

Starkweather, C. W. (1987). *Fluency and stuttering*. Prentice-Hall.

Stockman, I. J. (2010). Listener reliability in assigning utterance boundaries in children's spontaneous speech. *Applied Psycholinguistics, 31*(3), 363–395. https://doi.org/10.1017/S0142716410000032

Stuttering Foundation of America. (n.d.). *7 tips for talking with your child*. Retrieved June 22, 2021, from https://www.stutteringhelp.org/7-tips-talking-your-child

Tichenor, S., Leslie, P., Shaiman, S., & Yaruss, J. S. (2017). Speaker and observer perceptions of physical tension during stuttering. *Folia Phoniatrica et Logopaedica, 69*(4), 180–189. https://doi.org/10.1159/000486032

Tichenor, S., & Yaruss, J. S. (2019). Stuttering as defined by adults who stutter. *Journal of Speech, Language, and Hearing Research, 62*(12), 4356–4369. https://doi.org/10.1044/2019_JSLHR-19-00137

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39*(6), 1284–1294. https://doi.org/10.1044/jshr.3906.1284

Tumanova, V., Woods, C., & Razza, R. (2020). The role of behavioral inhibition for conversational speech and language characteristics of preschool-age children who stutter. *American*

Journal of Speech-Language Pathology, 29(2), 638–651. https://doi.org/10.1044/2019_AJSLP-19-00026

Van Riper, C. (1973). *The treatment of stuttering*. Prentice-Hall.

Wagovich, S. A., & Bernstein Ratner, N. (2007). Frequency of verb use in young children who stutter. *Journal of Fluency Disorders, 32*(2), 79–94. https://doi.org/10.1016/j.jfludis.2007.02.003

WebMD. (n.d.). *How to help a stuttering child*. Retrieved June 22, 2021, from https://www.webmd.com/children/guide/how-to-help-stammering-child

Weiss, A. L., & Zebrowski, P. M. (1992). Disfluencies in the conversations of young children who stutter: Some answers about questions. *Journal of Speech and Hearing Research, 35*(6), 1230–1238. https://doi.org/10.1044/jshr.3506.1230

Wiig, E. H., Secord, W. A., & Semel, E. (1992). *Clinical Evaluation of Language Fundamentals–Preschool*. Psychological Corporation.

Wilkenfeld, J. R., & Curlee, R. F. (1997). The relative effects of questions and comments on children's stuttering. *American Journal of Speech-Language Pathology, 6*(3), 79–89. https://doi.org/10.1044/1058-0360.0603.79

Wingate, M. E. (2001). SLD is not stuttering. *Journal of Speech, Language, and Hearing Research, 44*(2), 381–383. https://doi.org/10.1044/1092-4388(2001/031)

Yairi, E., & Ambrose, N. G. (2005). *Early childhood stuttering: For clinicians by clinicians*. Pro-Ed.

Yaruss, J. S. (1997). Clinical implications of situational variability in preschool children who stutter. *Journal of Fluency Disorders, 22*(3), 187–203. https://doi.org/10.1016/S0094-730X(97)00009-0

Yaruss, J. S. (1999). Utterance length, syntactic complexity, and childhood stuttering. *Journal of Speech, Language, and Hearing Research, 42*(2), 329–344. https://doi.org/10.1044/jslhr.4202.329

Zackheim, C. T., & Conture, E. G. (2003). Childhood stuttering and speech disfluencies in relation to children's mean length of utterance: A preliminary study. *Journal of Fluency Disorders, 28*(2), 115–142. https://doi.org/10.1016/S0094-730X(03)00007-X

Zibulsky, J., Casbar, C., Blanchard, T., & Morgan, C. (2019). Parent question use during shared reading time: How does training affect question type and frequency? *Psychology in the Schools, 56*(2), 206–219. https://doi.org/10.1002/pits.22219

Zimmerman, I., Steiner, V., & Pond, R. (1992). *Preschool Language Scale* (3rd ed.). Psychological Corporation.