JSLHR

Research Article

Does Working Memory Enhance or Interfere With Speech Fluency in Adults Who Do and Do Not Stutter? Evidence From a Dual-Task Paradigm

Naomi Eichorn,[a] Klara Marton,[b,c] Richard G. Schwartz,[b] Robert D. Melara,[d] and Steven Pirutinsky[e]

Purpose: The present study examined whether engaging working memory in a secondary task benefits speech fluency. Effects of dual-task conditions on speech fluency, rate, and errors were examined with respect to predictions derived from three related theoretical accounts of disfluencies.
Method: Nineteen adults who stutter and twenty adults who do not stutter participated in the study. All participants completed 2 baseline tasks: a continuous-speaking task and a working-memory (WM) task involving manipulations of domain, load, and interstimulus interval. In the dual-task portion of the experiment, participants simultaneously performed the speaking task with each unique combination of WM conditions.
Results: All speakers showed similar fluency benefits and decrements in WM accuracy as a result of dual-task conditions. Fluency effects were specific to atypical forms of disfluency and were comparable across WM-task manipulations. Changes in fluency were accompanied by reductions in speaking rate but not by corresponding changes in overt errors.
Conclusions: Findings suggest that WM contributes to disfluencies regardless of stuttering status and that engaging WM resources while speaking enhances fluency. Further research is needed to verify the cognitive mechanism involved in this effect and to determine how these findings can best inform clinical intervention.

Devoting more mental effort to a task is generally assumed to benefit task performance. The present study explored whether this premise applies to speech production, particularly for speakers with varying levels of speech fluency. Within the stuttering literature, both the covert repair hypothesis (CRH; Postma & Kolk, 1993) and the vicious circle hypothesis (VCH; Vasiç & Wijnen, 2005) propose that stuttering arises from excessive speech-monitoring activity and that limiting attentional or working-memory (WM) resources available for such monitoring may improve fluency. Evidence from movement sciences and sport psychology research (e.g., Beilock, Carr, MacMahon, & Starkes, 2002; Wulf & Lewthwaite, 2010) further suggests that conscious attempts to control learned motor skills interfere with automaticity and result in performance breakdown. Following on these lines of literature, the present study examined the role of WM in fluent speech production by engaging WM resources in a systematic manner and assessing whether concurrently produced speech showed appreciable fluency gains.

Chronic stuttering affects nearly 1% of adults (Bloodstein & Bernstein Ratner, 2008) and is highly resistant to treatment once it persists beyond early childhood. Although most interventions recognize the need to address symptoms other than overt stuttering behaviors and can yield significant benefits, traditional therapies typically train speakers to focus on specific aspects of speech production, such as respiration, articulation, or prosody (for a recent review, see Blomgren, 2013). In contrast, accounts of disfluency reviewed in the following section suggest that enhanced fluency may be

[a]The University of Memphis, TN
[b]The Graduate School and University Center, The City University of New York, New York
[c]Bárczi Gusztáv College of Special Education, Eötvös Loránd University, Budapest, Hungary
[d]The City College of New York, New York
[e]Georgian Court University, Lakewood Township, NJ
Correspondence to Naomi Eichorn: neichorn@memphis.edu

Editor: Jody Kreiman
Associate Editor: Hans-Georg Bosshardt
Received July 17, 2015
Revision received October 19, 2015
Accepted November 16, 2015
DOI: 10.1044/2015_JSLHR-S-15-0249

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

achieved by minimizing rather than maximizing conscious control over these component processes.

Disfluencies as a Result of Covert Repairs

The CRH (Postma & Kolk, 1993) provides an account of stuttering that is based on Levelt's (1983) model of speech production and monitoring and is largely centered around the phenomenon of *covert repairs,* which are errors repaired in the prearticulatory speech plan. When editing occurs after articulation is initiated, production of the utterance is disrupted and disfluencies are observed. According to the CRH, excessive disfluencies in persons who stutter (PWS) arise from an underlying deficit in phonological encoding. This deficit causes frequent errors to occur during speech planning, resulting in frequent subsequent repairs and overt disfluencies as an unavoidable side effect.

Research examining key claims of the CRH has produced mixed results. Although some studies report comparable phonological skills in PWS and persons who do not stutter (PWNS; e.g., Melnick, Conture, & Ohde, 2003; Ryan, 1992; Vincent, Grela, & Gilbert, 2012), considerable evidence suggests at least subtle phonological weaknesses among individuals who stutter compared with individuals who do not, on the basis of analyses of spoken output (Anderson & Byrd, 2008), priming effects (Byrd, Conture, & Ohde, 2007), rhyme judgments (R. M. Jones, Fox, & Jacewicz, 2012; Weber-Fox, Spruill, Spencer, & Smith, 2008), phoneme monitoring (Sasisekaran, De Nil, Smyth, & Johnson, 2006), and nonword repetition (Hakim & Bernstein Ratner, 2004; Sasisekaran & Weisberg, 2014).

Dual-task studies that attempted to suppress speech-monitoring resources have been similarly inconclusive. Early studies showed reductions in disfluencies when PWS were exposed to various distractions (e.g., Maraist & Hutton, 1957); however, this effect has not been replicated consistently (e.g., Oomen & Postma, 2002), and some studies report a reverse effect (i.e., increased disfluencies under dual-task conditions; Bosshardt, 1999; Oomen & Postma, 2001). Clear patterns across these outcomes are difficult to detect, and critical methodological differences likely contribute to the lack of consistent findings. For example, studies vary significantly in the speaking tasks used, as well as the complexity, automaticity, and processing domains of secondary tasks being performed. Most studies also do not distinguish between specific types of disfluency, which may be differentially influenced by dual-task demands, and do not specify the cognitive resource targeted by secondary tasks or why it was selected (for exceptions, see Bosshardt, 1999; De Nil & Bosshardt, 2000).

The present study examined two related predictions of the CRH. Because Levelt's model assumes that error signals generated by the speech monitor are received by WM (Levelt, 1983), secondary tasks that reliably tax WM were expected to affect monitoring integrity, with a resulting increase in overt speech errors. The same conditions were also expected to cause a decrease in speech disfluencies, because disfluencies essentially represent a consequence of error repair.

Disfluencies as a Result of Hypervigilance

Like the CRH, the VCH (Vasiç & Wijnen, 2005) links stuttering to speech monitoring but attributes excessive disfluency to a hypervigilant mechanism that emits alarm signals in the absence of true encoding errors. As speculated by the authors, early stuttering may arise from a child's heightened awareness to her or his own disfluencies, causing maladaptive adjustments of monitoring parameters that ultimately result in a vicious circle of increasing discontinuity and counterproductive forms of compensation.

Empirical support for the VCH was originally based on dual-task experiments conducted by Vasiç and Wijnen (2005), which showed reductions in atypical disfluencies (particularly blocks) when PWS engaged in attention-demanding tasks while speaking. Predictions of the VCH are further supported by data suggesting an association between stuttering and perfectionistic tendencies, as measured by general attitudes and beliefs (Brocklehurst, 2008) and speech-fluency ratings (Lickley, Hartsuiker, Corley, Russell, & Nelson, 2005). Atypical monitoring activity in PWS (compared with PWNS) has also been demonstrated neurophysiologically, with higher levels of activity in the anterior cingulate cortex (ACC), a region that detects conflict and modulates cognitive control (S. Brown, Ingham, Ingham, Laird, & Fox, 2005; De Nil, Kroll, Kapur, & Houle, 2000), and larger amplitude of error-related negativity, an electrophysiological component associated with the ACC (Arnstein, Lakey, Compton, & Kleinow, 2011).

On the basis of the VCH, we expected differential dual-task effects depending on stuttering status. Monitoring integrity is assumed to be intact in PWNS but faulty in PWS; thus, the participating PWS should experience a fluency benefit when monitoring resources are engaged, but the PWNS should not. This prediction was not examined by Vasiç & Wijnen (2005), who described only fluency effects for PWS.

Disfluencies as a Result of Explicit Attention

A final approach to stuttering, not considered in the existing literature, is to view disfluent speech as a breakdown in performance caused by heightened self-awareness or excessive attention to the process of speaking. This perspective does not assume involvement of a speech-monitoring mechanism and suggests disruption at the level of motor movements rather than speech planning. An important distinction is drawn in this literature between implicit and explicit forms of information processing, with the former being optimally suited for skilled motor performance but the latter being counterproductive (e.g., Masters, 1992). Whereas explicit processing is rule based and linked to conscious awareness, the implicit system is skill based and involves content not easily verbalized or available within WM (e.g., Maddox & Ashby, 2004). Implicit processing is efficient, resilient, and less vulnerable to distraction (e.g., Beilock et al., 2002); thus, it is ideal for managing well-practiced motor skills and achieving effortless performance patterns (for a review, see Dietrich & Stoll, 2010).

Situations of pressure or demands for excellence often cause performers to consciously focus on internal movements involved in task execution and to apply explicit processing to proceduralized knowledge (McNevin, Shea, & Wulf, 2003). The result is a disruption of automaticity, termed *choking,* in which movement sequences are *dechunked* into independent units that resemble the effortful movement patterns of early motor learning (Baumeister, 1984).

Consistent with this account, skilled performers across a wide variety of tasks execute movements less efficiently when experimental conditions demand skill-focused attention (Beilock & Gray, 2012; Gray, 2004; Snyder & Logan, 2013). Conversely, dual-task studies show enhanced motor outcomes when WM resources are engaged (Masters, 1992; Beilock et al., 2002), implying that forced reliance on implicit processing benefits performance. Studies manipulating the focus of attention further demonstrate that instructing participants to focus on movement effects (external focus) rather than physical movements themselves (internal focus) optimizes performance in various gross motor tasks (McKay & Wulf, 2012; McNevin et al., 2003). On the basis of these findings, Wulf and others proposed the constrained action hypothesis (CAH), which posits that conscious control of movements through an internal attentional focus constrains the movement system and disrupts automatization, whereas external focusing induces more effortless, reflexive performance (Wulf & Lewthwaite, 2010). This effect is also observed in oral motor tasks (Freedman, Maas, Caligiuri, Wulf, & Robin, 2007), suggesting that similar outcomes occur in fine motor movements involving the orofacial system. We therefore expected that predictions related to the CAH could be applied to speech production. We anticipated that increasing reliance on implicit modes of processing (via dual tasking) would induce more effortless speech production, as manifested by a reduction in disfluencies. We further predicted differential effects on the basis of disfluency type, with more pronounced fluency benefits for behaviors representing dechunking, or a breakdown in speech automaticity (atypical disfluencies), compared with behaviors associated with language formulation demands (typical disfluencies).

Dual-Task Paradigm

Three important considerations related to the dual-task paradigm were incorporated into the design of the present study. One critical factor influencing dual-task performance is the degree to which concurrent tasks compete for the same resources. Tasks relying on similar processes typically result in interference (Leclercq, 2002), whereas tasks involving different domains or modalities can often be combined without compromising performance (Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002; Duff & Logie, 2001). These findings are consistent with a multi-component view of WM, as proposed by Baddeley and Hitch (1974). Evidence within the stuttering literature also suggests that relative to PWNS, PWS may be more susceptible to interference when performing concurrent tasks that draw upon similar neurocognitive resources (Bosshardt, 2006; for a review, see Bajaj, 2007).

A second consideration is the extent to which tasks overlap in time. Under the bottleneck explanation for dual-task interference, operations simultaneously vying for dedicated use of central resources typically result in delayed or impaired performance on one or both tasks (Pashler, 1992, 1994). Similar competition occurs between continuous tasks performed simultaneously, although in this case the tasks intermittently compete for WM resources, and the extent of interference depends on how frequently access is required by each task (Pashler, 1992).

Last, dual-task interference is highly influenced by the degree of automaticity associated with each task (Pashler, 1999; Poldrack et al., 2005). Performance is considered automatic when it requires minimal attention and is not affected by other ongoing mental activity. When a task is well practiced, larger pieces of information are chunked and continuous effort is no longer necessary. Studies within the stuttering literature indicate that adults who stutter perform more poorly than those who do not on secondary tasks executed while speaking (Saltuklaroglu, Teulings, & Robbins, 2009; Smits-Bandstra & De Nil, 2009), suggesting that speaking is less automatized for this group and requires more frequent access to attentional resources.

Incorporating these considerations, the present study combined spontaneous speaking with a secondary task that placed demands on WM and included manipulations of domain (verbal vs. spatial), load, and interstimulus interval (ISI). We predicted that speaking under dual-task conditions would affect fluency with distinct patterns of results on the basis of different theoretical assumptions. Under the CRH, conditions that taxed WM were expected to cause a reduction in speech disfluencies (all types) with an accompanying increase in overt speech errors (word and sound errors; for details, see Method). The outcome predicted by the VCH was a differential dual-task effect depending on stuttering status, with dual-task conditions increasing fluency only for PWS (because monitoring integrity is presumably intact in controls). Under the CAH and explicit-attention framework, disfluencies were expected to decrease for all speakers under dual-task conditions, with a greater reduction in disfluency types representing a breakdown in speech automaticity (atypical disfluencies) compared with disfluencies associated with language formulation demands. In accordance with the dual-task literature, we further anticipated differential effects depending on experimental condition, with greater fluency benefits when secondary tasks involved verbal versus spatial WM, higher versus lower WM loads, and shorter versus longer ISIs. Finally, we expected a greater decrement in WM performance under dual- versus single-task conditions for PWS compared with PWNS, on the basis of previous studies suggesting incomplete automatization of speech production (Saltuklaroglu et al., 2009; Smits-Bandstra & De Nil, 2009) and greater overall susceptibility to dual-task interference (Bosshardt, 1999, 2006) in PWS.

Method

Participants

Twenty adults who stutter (AWS) and 20 adults who do not stutter (AWNS), all aged 18–35 years, participated in the study. Participants were recruited via flyers distributed at college campuses, speech-language clinics, and stuttering support groups in and around New York City. Participants were given an overview of experimental procedures prior to scheduling, with additional details provided in print at the time of their session. Informed consent was obtained before experimental procedures began. Flyers, consent forms, and all procedures were approved by the City University of New York Institutional Review Board. Participants were each compensated $40 for their time. Data from one AWS was excluded due to poor speech intelligibility in recorded samples.

All participants met the following inclusionary criteria: (a) nonverbal intelligence within at least the average range (standard score = 85) on the Test of Nonverbal Intelligence, Fourth Edition (L. Brown, Sherbenou, & Johnsen, 2010); (b) expressive vocabulary within at least the average range (standard score = 85) on the Expressive One-Word Picture Vocabulary Test, Fourth Edition (Martin & Brownell, 2011); and (c) absence of significant medical history, learning disability, hearing loss, head injury, and cognitive impairment, as determined by a screening questionnaire and interview. Participants completed computer-based operation- and symmetry-span tasks (Unsworth, Heitz, Schrock, & Engle, 2005) to verify that groups did not differ in WM capacity for verbal and spatial stimuli. A preliminary dual task in which continuous speech production was combined with a simple vigilance task was also administered to confirm that speaker groups were comparable in their basic ability to manage dual-task demands (see Vigilance Task and Preliminary Analyses for details). All participants spoke English as their primary language. Table 1 gives a summary of participant demographics, scores on standardized cognitive and language measures, and preliminary dual-task results.

The participating PWS all self-identified as AWS and were previously diagnosed by certified speech-language pathologists. Formal measures of stuttering severity were obtained via the Stuttering Severity Instrument–Fourth Edition (Riley, 2009). The Overall Assessment of Stuttering Experience Survey was also completed by all AWS to obtain subjective measures of the impact of stuttering (Yaruss & Quesal, 2006). Stuttering severity ranged from very mild to severe as measured by the Stuttering Severity Instrument; stuttering impact ranged from mild–moderate to moderate–severe as measured by the Overall Assessment of the Speaker's Experience of Stuttering. All AWS reported childhood onset of stuttering, and all but one had received prior speech therapy for their stuttering.

Table 1. Participant characteristics.

Characteristic	AWNS	AWS
Age (years)	25.60 (4.58)	27.21 (4.18)
Men/Women	11/9	12/7
Handedness (right/left)	17/3	16/3
SSI-4 (raw score)	4.11 (0.32)	15.78 (8.64)
TONI-4 (standard score)	107.85 (12.90)	107.56 (12.54)
EOWPVT-4 (standard score)	104.60 (9.81)	102.11 (12.45)
SSpan (absolute score)	20.65 (7.04)	19.61 (12.52)
OSpan (absolute score)	45.85 (16.16)	42.56 (17.39)
Vigilance task (accuracy)	96.67 (2.42)	96.75 (3.49)

Note. Cells with parentheses provide mean values, with standard deviations provided in the parentheses. AWNS = adults who do not stutter; AWS = adults who stutter; SSI-4 = Stuttering Severity Instrument–Fourth Edition; TONI-4 = Test of Nonverbal Intelligence, Fourth Edition; EOWPVT-4 = Expressive One-Word Picture Vocabulary Test, Fourth Edition; SSpan = symmetry span; OSpan = operation span.

Preliminary analyses indicated no between-groups difference in age, $t(37) = -1.14$, $p = .26$; nonverbal IQ, $t(36) = 0.07$, $p = .94$; and expressive vocabulary, $t(36) = 0.69$, $p = .50$. The two groups also showed similar distributions of gender and handedness, $X^2(1, N = 39) = 0.27$, $p = .61$, and $X^2(1, N = 39) = 0.005$, $p = .95$, respectively. Group performance patterns on span measures (using an overall absolute scoring procedure; see Unsworth et al., 2005) revealed no between-groups differences on either span task—operation span: $t(36) = 0.61$, $p = .55$; symmetry span: $t(36) = 0.32$, $p = .75$.

General Procedure

The experimental procedure incorporated four sets of tasks: three baseline tasks, a preliminary dual task, two full dual tasks, and a final baseline speaking task. Participants were required to meet a performance criterion of 80% on baseline WM and preliminary dual tasks to proceed to full dual-task conditions. Task sets were administered in the same sequence to all participants; tasks within each set and trials within each task were randomized. All tasks were administered on a desktop computer. E-Prime 2.0 software (Psychology Software Tools, Inc., Sharpsburg, PA) was used to present stimuli and record button-press responses from a standard keyboard. Spoken output was recorded via an adjustable head-worn unidirectional microphone (Shure SM10A, Niles, IL) connected to a preamplifier (Switchcraft 308TR, Chicago, IL). Supplementary video recording was obtained during all speech-production tasks, as recommended by Rousseau, Onslow, Packman, and Jones (2008). Experimental tasks were completed in a single session, with breaks as needed.

Baseline Tasks

Speaking Task

In the baseline speaking task, participants produced continuous, spontaneous speech over a 60-s period for each of four topic prompts (e.g., *Describe a recent vacation*; *What's in a name?*; *Thoughts or memories related to pets*). Prompts were largely adapted from the Student Opinion question list maintained by the Learning Network blog section of the *New York Times* (Slotnik & Schulten, 2012). Prompts for individual trials were presented in a set of three, allowing participants to select topics according to personal preference and experiences. Spoken output for each trial was automatically recorded via E-Prime, saved to individual audio files, and subjected to extensive off-line coding and analysis (see Data Processing). The first speaking trial was considered practice and was omitted from analyses. Once a prompt was selected, it was withheld from subsequent experimental trials. The baseline speaking task was administered twice—once at the beginning of the experiment (pretest baseline) and once at the end (posttest baseline)—to help account for possible order and practice effects.

WM Tasks

WM tasks were modeled after those of Salthouse, Babcock, and Shaw (1991) and adapted to examine the effects of three experimental manipulations: domain (verbal vs. spatial), load (2, 3, and 4), and ISI (long vs. short). Trials within both WM domains were identically structured and designed (after piloting) to be simple enough that participants could maintain high levels of accuracy even during dual-task conditions. Feedback was provided throughout the task to optimize performance. All stimuli were presented for 3,000-ms durations.

In the verbal domain, participants viewed a start number followed by a series of individually presented single-digit addition operations (e.g., +3) and a prompt to enter the correct numeric outcome. All start numbers were between 0 and 4, operation numbers were between 1 and 4, and outcomes were below 10. Individual operations were never repeated more than once during a single trial. Each load condition resulted in four potential numeric outcomes (2, 3, 4, or 5 for a WM load of 2; 4, 5, 6, or 7 for a load of 3; and 6, 7, 8, or 9 for a load of 4), with each outcome having an equal probability (25%) in each experimental condition.

In the spatial domain, a single colored circle was presented in one of four cells within a 2 × 2 grid. This stimulus was followed by a series of individually presented directional arrows and a prompt to enter the number representing the final location of the circle, according to a numbered on-screen grid. As in the verbal WM task, individual operations (directions) were never repeated more than once during a single trial and each WM load condition resulted in four potential position outcomes that had equal probability (25%) within any experimental condition.

WM load was manipulated by varying the number of sequential operations (two, three, or four) to be performed on the initial number or circle stimulus. ISI varied between individual operations, with intervals of 3,000 and 1,000 ms for long and short conditions, respectively. The purpose of the ISI manipulations was to vary the relative frequency with which WM resources were accessed. The shorter ISI (1,000 ms) was expected to increase competition between simultaneous tasks for dedicated use of the same resource and has been used in previous dual-task studies involving continuous speech production (Declerck & Kormos, 2012; Oomen & Postma, 2002).

Within each domain, two practice items were presented, both representing a WM load of 3 and ISI of 1,000 ms. Practice items were followed by 60 test items, with 10 items for each unique combination of WM load and ISI (see Table 2). Accuracy of keyboard responses was automatically recorded for subsequent analysis. Figure 1 presents a sample series of screens representing the baseline verbal and spatial WM tasks.

Vigilance Task

The vigilance task was a preliminary dual-task condition in which participants produced spontaneous speech while responding to simple stimuli (colored circles) appearing at varying locations on the screen. The task was designed with minimal attention demands and no WM load (other than task instructions), to obtain a basic measure of participants' ability to manage dual-task demands. As in baseline speaking tasks, participants selected one of three prompts and produced a 60-s sample of spoken output for each of four speaking trials. The first trial was considered practice and was omitted from analyses. While speaking, participants responded as quickly as possible to on-screen circle stimuli by pressing designated buttons on the keyboard (J or F) when stimuli appeared on the right or left side of the screen, respectively. Twenty circle stimuli were presented during each speaking trial. All circles remained visible for 1,000 ms; ISI randomly varied between 1,000, 1,500, 2,000, 2,500, and 3,000 ms, with each ISI occurring four times per speaking trial. Multiple variations of ISI were included to minimize the predictability of target responses. Responses were considered accurate if they occurred during the 1,000-ms period of stimulus presentation and corresponded to the side on which the stimulus was presented. Spoken output was recorded to ensure that speech production remained continuous.

Dual Tasks

In the full dual-task conditions, participants performed the speaking task and each WM task (verbal or spatial) simultaneously. Tasks in dual conditions were identical to baseline counterparts; however, the total number of WM trials per speaking trial varied depending on load and ISI conditions. Certain conditions required incomplete WM trials due to timing constraints imposed by the uniform 60-s speaking trial. During incomplete trials, an initial stimulus was presented, followed by one or more operations, followed by a large X, which signaled participants to disregard previous stimuli for that trial and begin again. Regardless of load and ISI conditions, speech was recorded continuously, and two keyboard responses were entered for each speaking trial. The order of WM trials with and without outcome responses was randomized during each speaking trial. Spoken output was saved when response prompts for WM appeared, resulting in two audio files per speaking trial.

Table 2. Experimental conditions for working-memory (WM) task.

<table>
<tr><th>Domain</th><th>ISI</th><th>WM load</th><th colspan="5">Sample trial</th><th>Number of trials</th></tr>
<tr><td>Verbal</td><td>Long (3,000 ms)</td><td>4</td><td>3</td><td>+1</td><td>+2</td><td>+1</td><td>+1</td><td>10</td></tr>
<tr><td></td><td></td><td>3</td><td>1</td><td>+1</td><td>+2</td><td>+3</td><td></td><td>10</td></tr>
<tr><td></td><td></td><td>2</td><td>2</td><td>+3</td><td>+2</td><td></td><td></td><td>10</td></tr>
<tr><td></td><td>Short (1,000 ms)</td><td>4</td><td>3</td><td>+2</td><td>+1</td><td>+1</td><td>+2</td><td>10</td></tr>
<tr><td></td><td></td><td>3</td><td>1</td><td>+1</td><td>+2</td><td>+3</td><td></td><td>10</td></tr>
<tr><td></td><td></td><td>2</td><td>2</td><td>+3</td><td>+1</td><td></td><td></td><td>10</td></tr>
<tr><td>Spatial</td><td>Long (3,000 ms)</td><td>4</td><td><table><tr><td></td><td></td></tr><tr><td>•</td><td></td></tr></table></td><td>→</td><td>←</td><td>↑</td><td>↓</td><td>10</td></tr>
<tr><td></td><td></td><td>3</td><td><table><tr><td></td><td></td></tr><tr><td></td><td>•</td></tr></table></td><td>←</td><td>→</td><td>↑</td><td></td><td>10</td></tr>
<tr><td></td><td></td><td>2</td><td><table><tr><td></td><td>•</td></tr><tr><td></td><td></td></tr></table></td><td>←</td><td>→</td><td></td><td></td><td>10</td></tr>
<tr><td></td><td>Short (1,000 ms)</td><td>4</td><td><table><tr><td></td><td></td></tr><tr><td></td><td>•</td></tr></table></td><td>←</td><td>→</td><td>↑</td><td>←</td><td>10</td></tr>
<tr><td></td><td></td><td>3</td><td><table><tr><td></td><td></td></tr><tr><td></td><td>•</td></tr></table></td><td>←</td><td>→</td><td>↑</td><td></td><td>10</td></tr>
<tr><td></td><td></td><td>2</td><td><table><tr><td></td><td></td></tr><tr><td></td><td>•</td></tr></table></td><td>←</td><td>→</td><td></td><td></td><td>10</td></tr>
</table>

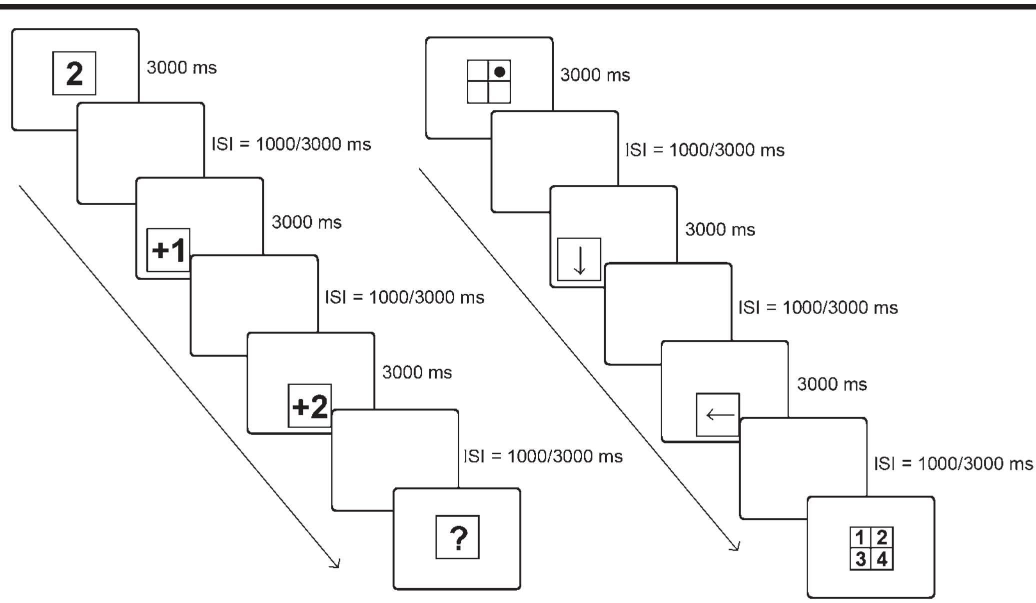
Note. ISI = interstimulus interval.

Participants were instructed to speak continuously while attending to the secondary task. A practice trial was provided within each domain and was repeated until both WM responses were correct. Practice items were followed by five speaking trials for each combination of WM load and ISI, resulting in a total of 30 speaking trials and 60 WM trials within each WM domain.

Data Processing

Dependent variables for WM and vigilance tasks consisted of accuracy scores (proportion of correct responses). Audio files from speaking tasks were orthographically transcribed and coded for speech interruptions, which were categorized as one of three interruption types: typical disfluencies, which included fillers, revisions, repetitions of

Figure 1. Representation of a single trial for the verbal (left) and spatial (right) working-memory (WM) tasks, each shown with WM load of 2. ISI = interstimulus interval.



phrases, and repetitions of multisyllabic words; atypical disfluencies, which consisted of repetitions of monosyllabic words, repetitions of sounds or syllables, prolongations, blocks, and broken words; and speech errors, which consisted of corrected or uncorrected sound and word errors. This categorization system was based on typologies and accepted distinctions between disfluency types within the stuttering literature (Ambrose & Yairi, 1999; Vasiç & Wijnen, 2005; Wingate, 1962; Yaruss, 1997) and speech-monitoring research (Levelt, 1983; Oomen & Postma, 2001).

Transcription and coding were completed independently by two experienced transcribers. A customized script using the qdap package (Rinker, 2013) in R (R Core Team, 2013) was used to automatically generate frequency counts of all spoken syllables (excluding interruptions) and counts of each disfluency and error type for individual speaking trials. Output was used to compute values for four speech-related variables: total number of syllables spoken; total number of typical disfluencies; total number of atypical disfluencies; and total number of speech errors. Counts of disfluencies and errors, with syllable counts, were used to calculate rates of disfluency and errors as a proportion of total syllables spoken.

Thirty-nine speech samples (one per participant) were randomly selected and transcribed by the transcriber who had not originally coded that trial. Interrater reliability for transcription and coding of key dependent variables was assessed via Pearson's product–moment correlations, with results indicating high levels of reliability across outcome measures: $r(37) = 1.00$ for syllable counts, $r(37) = .96$ for typical disfluencies, $r(37) = .92$ for atypical disfluencies, and $r(37) = .88$ for errors.

Data were prepared for statistical analyses using a coding scheme with six dummy variables to represent specific contrasts of interest, on the basis of a priori hypotheses: Disfluency Type, coded with typical disfluencies as the reference level and atypical as the comparison; Task Type, with baseline tasks as reference and dual tasks as the comparison; Domain, with spatial as the reference and verbal as the comparison; ISI, with long ISI as the reference and short ISI as the comparison; WM Load1, with a load of 2 as the reference and a load of 4 as the comparison; and WM Load2, with a load of 3 as the reference and 2 and 4 as the comparison.

Data Analysis

Data were analyzed to examine effects of experimental manipulations and speaker types on speech fluency, speech errors, and secondary-task performance. As already described, fluency variables consisted of counts of typical and atypical disfluencies; WM performance was correct (1) or incorrect (0) for each trial. Generalized multilevel linear models were used because they provide maximum flexibility and robustness in analyzing multilevel experimental data with nonnormally distributed dependent variables (Hoffman & Rovine, 2007). For each outcome measure, individual multilevel models were estimated using the Lme4 package in R (Bates, Maechler, Bolker, & Walker, 2014). Preliminary analyses of fluency and speech-error data indicated that a nonzero inflated negative binomial distribution best fitted disfluency and error counts; therefore, all multilevel generalized linear models for disfluency and error counts utilized a negative binomial link function (Hardin & Hilbe, 2007). In theory, discrete counts of event occurrence are best represented by the negative binomial distribution, which is truncated at zero, inflated at lower counts, and positively skewed. Recent reports within the stuttering literature have also demonstrated that disfluency counts for children who stutter and AWS follow a negative binomial distribution (M. Jones, Dobson, Onslow, & Carey, 2009; Tumanova, Conture, Lambert, & Walden, 2014). To control for unequal durations of speaking trials across experimental conditions, as well as for varying speaking rates, raw counts of disfluencies and errors were converted to a ratio of disfluencies or errors to syllables by including the total number of syllables spoken in each trial as an exposure or offset variable (Hardin & Hilbe, 2007). WM performance was correct (1) or incorrect (0) for each trial and therefore analyzed using a multilevel logistic regression model.

For each analysis, a successive series of increasingly complex models were estimated. Predictors were entered in an a priori selected order such that individual differences were entered first, followed by variables testing primary hypotheses (e.g., Disfluency Type, Task Type), followed by corresponding interactions in increasing complexity (two- then three-way). Nonsignificant variables and interactions were removed at each stage unless they were statistically required to test higher order interactional hypotheses. Secondary-task-related characteristics were added last. Successive models were compared using the log-likelihood ratio, Akaike information criterion, and Bayesian information criterion.

Results

Preliminary Analyses

Comparison of pre- and posttest baseline measures for both typical and atypical forms of disfluency within each group (calculated as a proportion of syllables) indicated no significant difference between the two baseline speaking measures: For typical disfluencies, $t(19) = -1.44$, $p = .17$, and $t(18) = 0.77$, $p = .45$, in AWNS and AWS, respectively; for atypical disfluencies, $t(19) = -0.16$, $p = .86$, and $t(18) = 0.93$, $p = .36$, in AWNS and AWS, respectively. Disfluency outcomes for the two baseline speaking tasks were therefore collapsed and included in models as a single baseline measure without further differentiation.

Performance on the vigilance task was examined to assess whether groups differed in their ability to respond accurately to on-screen stimuli while engaged in continuous speaking. Descriptive data are provided in Table 1. Results indicated no between-groups differences in performance, $t(37) = -0.09$, $p = .93$, suggesting comparable ability to manage dual-task demands.

The results in the following first consider secondary-task performance, because these data provided evidence relating to the integrity of the task design. These outcomes are followed by analyses of dual-task effects on three aspects of speech production: fluency, rate, and speech errors.

Secondary-Task Performance

Secondary-task performance was measured as response accuracy for each WM trial and analyzed via multilevel logistic regression modeling. In the online supplemental materials, Supplemental Appendix A provides descriptive data by group and task condition. A series of successively complex models was used to examine effects of stuttering status and experimental condition on secondary-task performance (see the online supplemental materials, Supplemental Appendix B). Examination of coefficients for the final model (see Table 3) indicated that each experimental manipulation showed an overall effect on secondary-task performance (see Figures 2A–2C). Because these negative binomial regression models utilized a natural-logarithm link function, relative incidence ratios were obtained by exponentiating the regression coefficients (Hardin & Hilbe, 2007). Results indicated that accuracy on the secondary task was 18% lower under dual compared with baseline conditions, $\beta = -1.69$, $SE = 0.08$, $t = 20.46$, $p < .00001$; 72% lower in verbal compared with spatial tasks, $\beta = -0.33$. $SE = 0.04$, $t = 7.71$, $p < .00001$; 113% higher for short compared with long ISI conditions, $\beta = 0.12$, $SE = 0.04$, $t = 2.92$, $p = .004$; and 75% lower for the highest WM load (4) compared with the lowest (2), $\beta = -0.28$, $SE = 0.05$, $t = 5.36$, $p < .00001$. However, there were no interactions between these predictors and no differences between AWS and AWNS. Thus, dual-task manipulations generally showed expected effects on performance; however, the hypothesized group difference in susceptibility to dual-task interference was not supported.

Speech Fluency

As already described, dual-task effects on fluency were analyzed via multilevel generalized linear models with a negative binomial link function and were offset by the total number of syllables in each speaking trial (Hardin & Hilbe, 2007). The dependent variable therefore represented the ratio of disfluencies to fluent syllables. Similar results were obtained when models were offset by the total duration (in seconds) of each speaking trial. Descriptive data for atypical and typical disfluencies are provided in the online

Figure 2. Secondary-task performance (with 95% confidence intervals), based on raw data, by task type and group according to (a) working-memory (WM) domain, (b) WM load, and (c) interstimulus interval (ISI). AWNS = adults who do not stutter; AWS = adults who stutter.

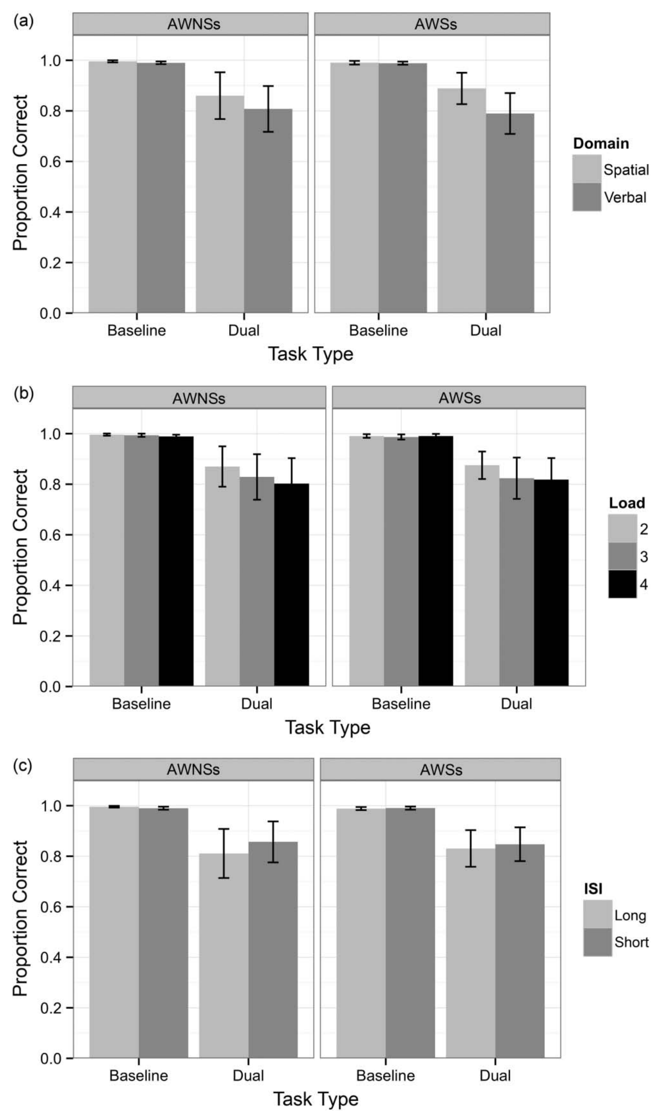


Table 3. Results of final model of secondary-task performance (M11).

Fixed effect	Unstandardized coefficient	Standard error	Incidence ratio	*t*	*p*
Intercept	3.72				
Task	–1.69	0.08	0.18	20.46	< .00001
Domain	–0.33	0.04	0.72	7.71	< .00001
ISI	0.12	0.04	1.13	2.92	.004
Load1	–0.28	0.05	0.75	5.36	< .00001
Load2	0.08	0.04	1.08	1.75	.08

Note. Model was fitted using maximum-likelihood estimation and included the random effect of an individual-level intercept (Var = 1.22). Model was based on 9,360 observations across 39 subjects. ISI = interstimulus interval.

supplemental materials (Supplemental Appendix C). All speaking trials during dual-task conditions were included, regardless of accuracy on secondary task. Identical analyses with erroneous responses excluded (17% of total trials for AWNS, 16% for AWS) produced the same results. A series of successive models (see the online supplemental materials, Supplemental Appendix D) was estimated and compared using the log-likelihood, Akaike information criterion, and Bayesian information criterion.

Final results indicated that a model including Speaker (AWNS vs. AWS), Disfluency (typical vs. atypical), Task (baseline vs. dual), and two two-way interactions (Speaker × Disfluency and Disfluency × Task) provided the best fit (see Table 4). Coefficients of the final model indicated that, controlling for number of syllables (speech rate), atypical disfluencies were significantly less frequent than typical disfluencies, occurring at 28% of the rate, $\beta = -1.26$, $SE = 0.06$, $t = 20.96$, $p < .00001$. The Disfluency × Speaker interaction was also significant, with AWS producing approximately 4 times more atypical disfluencies compared with AWNS, $\beta = 1.43$, $SE = 0.03$, $t = 41.83$, $p < .00001$. Most important, the Disfluency × Task interaction indicated a significant reduction in the frequency of atypical disfluencies under dual compared with baseline tasks. This was true regardless of speaker type (AWS vs. AWNS) or experimental condition (Domain, ISI, Load) within the dual task (two-way Task × Speaker interactions, as well as those between Task and individual experimental manipulations, were not significant). During dual tasks, atypical disfluencies occurred at 70% of the rate at which they occurred under nondual-task conditions, $\beta = -0.35$, $SE = 0.06$, $t = 6.04$, $p < .00001$. Thus, for every 100 atypical disfluencies produced during baseline conditions, 70 were produced during dual-task conditions. Typical forms of disfluency did not show any change as a result of experimental manipulations. This differential dual-task effect depending on disfluency type but not speaker type was consistent with the predictions of the CAH, but not the VCH. Analysis of subject-level factors, including expressive vocabulary, verbal WM (operation span), and stuttering severity (Stuttering Severity Instrument, Overall Assessment of the Speaker's Experience of Stuttering), indicated no interaction between any factor and task type. Thus, dual-task effects were not differentially affected by individual differences in these variables.

Speech Rate

Experimental effects on speaking rate were examined by estimating multilevel regression models using total number of spoken syllables as the dependent variable. Models were offset by the total duration of each trial (in seconds) to control for differences in the length of speaking trials in different experimental conditions. Descriptive data and details related to model comparisons are provided in the online supplemental materials (Supplemental Appendixes E and F, respectively).

Results suggested that a model including Speaker (AWS vs. AWNS), Task (baseline vs. dual), and their interaction best described the data (see Table 5). Coefficients of the final model indicated that AWS produced approximately 28 fewer syllables than AWNS, controlling for total duration, $\beta = -28.22$, $SE = 5.72$, $t = 4.93$, $p = .00002$. The overall effect of Task on the number of syllables spoken was also significant, with a reduction of approximately 50 syllables in dual compared with baseline task conditions, $\beta = -49.61$, $SE = 1.03$, $t = 48.09$, $p < .00001$. The Speaker × Task interaction further indicated that AWNS showed a greater reduction in syllable counts under dual conditions compared with AWS, with AWS demonstrating a smaller dual-task effect by approximately 10 syllables, $\beta = 10.38$, $SE = 1.48$, $t = 7.02$, $p < .00001$.

Speech Errors

Experimental effects on overt speech errors were analyzed via a series of multilevel negative binomial regression models. As in preceding models for disfluency, speech-error models were offset by the total number of syllables in each speaking trial to control for unequal durations of speaking trials across different conditions and varying speaking rates. The dependent variable therefore represented the ratio of speech errors to fluent syllables. Descriptive statistics and details related to model comparisons are available in the online supplemental materials (Supplemental Appendixes E and G, respectively). Results revealed that no model was a significant improvement over a baseline model consisting of a random intercept, indicating that contrary to the predictions of the CRH, speech errors were not influenced by experimental conditions or stuttering status.

Table 4. Results of final model of disfluencies (M5).

Fixed effect	Unstandardized coefficient	Standard error	Incidence ratio	*t*	*p*
Intercept	−3.34				
Disfluency	−1.26	0.06	0.28	20.96	< .00001
Speaker	0.21	0.16	1.23	1.30	.19
Task	0.03	0.04	1.04	0.90	.37
Disfluency × Speaker	1.43	0.03	4.20	41.83	< .00001
Disfluency × Task	−0.35	0.06	0.70	6.04	< .00001

Note. Model was fitted using maximum-likelihood estimation (Laplace approximation) and used a negative binomial logarithmic link function ($\theta = 6.07$). It included the random effect of an individual-level intercept (Var = 0.27). Residual variance was equal to 1.11, and the model was based on 9,826 observations across 39 subjects.

Table 5. Results of final model of speech rate (M4).

Fixed effect	Unstandardized coefficient	Standard error	*t*	*p*
Intercept	140.11	3.99		
Speaker	−28.22	5.72	4.93	.00002
Task	−49.61	1.03	48.09	< .00001
Speaker × Task	10.38	1.48	7.02	< .00001

Note. Model was fitted using maximum-likelihood estimation and included the random effect of an individual-level intercept (Var = 297.7). Residual variance was equal to 486.6, and the model was based on 4,914 observations across 39 subjects. The outcome variable represented the number of syllables spoken offset by the duration of each speaking trial, which varied by condition.

Discussion

The key finding emerging from this study was that dual-task conditions enhanced speech fluency. This outcome was in line with the CAH, which predicts more efficient execution of well-learned motor sequences when WM resources are engaged in unrelated tasks. Also consistent with the CAH, this effect was specific to atypical forms of disfluency, which most clearly represent a breakdown of automaticity. The fluency benefit did not differ with stuttering status and was accompanied by a reduction in speech rate (greater in PWNS than PWS), but there was no corresponding change in the frequency of overt speech errors. Thus, neither speech-monitoring account of stuttering was fully supported. Contrary to predictions from the dual-task literature, fluency effects were not influenced by manipulations within the secondary task (WM domain, load, and ISI). Secondary-task performance was compromised in dual compared with baseline conditions and was affected by each manipulation within the WM task; however, these effects did not differ between the two speaker groups. Thus, our final prediction of group differences in performance decrement on the secondary task was not supported. Results for individual analyses are considered next, followed by a general discussion of their broader theoretical implications.

Secondary-Task Performance Patterns

Results indicated that the WM task generally worked as intended. Overall WM performance was less accurate for dual compared with baseline conditions, verbal compared with spatial WM, high compared with low WM loads, and long compared with short ISI.

Poorer secondary-task performance under dual versus baseline conditions is an expected result of the dual-task paradigm (Pashler, 1994), particularly when concurrent tasks draw upon similar cognitive resources (Leclercq, 2002) or cannot be performed automatically (Pashler, 1999; Poldrack et al., 2005). Previous dual-task studies show similar patterns of performance decrement in secondary tasks executed while speaking (e.g., Bosshardt, 1999; Declerck & Kormos, 2012; Smits-Bandstra & De Nil, 2009), suggesting that producing speech is never completely automatic and requires some degree of ongoing attention in all speakers.

Within the secondary task, WM load showed the clearest effect on performance, with the largest load (4) resulting in significantly poorer accuracy than the smallest load (2). This is consistent with the original study from which the task was derived (Salthouse et al., 1991), as well as general data on set size and updating within the WM literature (Oberauer, 2009; Unsworth & Engle, 2007).

Results showed a main effect of Domain, with verbal tasks being more difficult than spatial tasks overall, but no Domain × Task interaction. Although Salthouse et al. (1991) showed a strong correlation between verbal and spatial versions of the original task, subtle task modifications for the present study may have influenced correct-response probabilities somewhat differently for each domain. Whereas the verbal task required consistent updating to arrive at the correct numeric outcome, the structure of the spatial task made it possible for participants to narrow down correct outcomes to two of four options for any one directional arrow within a sequence of arrows, and to a single correct response after the subsequent arrow. It is not clear whether this type of strategy necessarily contributed to performance differences by domain, particularly because there was no evidence of a Domain × Load interaction.

ISI affected performance, with longer ISIs resulting in poorer recall. The most logical interpretation of this finding is that ISI conditions served as an additional, unintended manipulation of WM, with longer ISIs rendering memory representations more susceptible to decay or interference, as described by many WM models (Baddeley, 2003; Cowan, 2008; Oberauer, 2009).

Dual-task manipulations did not interact with Task Type (baseline vs. dual), possibly due to the relative ease of the secondary task. The WM task was intentionally designed to be simple so that high levels of accuracy could be maintained during dual tasking. This emphasis was based on the consideration that exceedingly difficult secondary tasks might lead to arbitrary prioritization of one task and make results difficult to interpret. Analyses ultimately revealed identical fluency effects when incorrect WM trials were included or excluded, suggesting that these effects were not contingent upon accurate secondary-task performance.

The lack of between-groups differences or interactions with the Group factor differs from previous studies in which AWS showed greater disruption of secondary-task performance (R. M. Jones et al., 2012; De Nil & Bosshardt,

2000) and less automatized speech production compared with AWNS (Saltuklaroglu et al., 2009; Smits-Bandstra & De Nil, 2009). It is possible that secondary-task demands in the present study did not sufficiently strain WM resources to expose group differences, which are often quite subtle (see, e.g., R. M. Jones et al., 2012).

Dual-Task Effects on Fluency

Our primary hypothesis, predicting a reduction in disfluencies under dual-task conditions, was supported. Results further corroborated expected differences in dual-task effects depending on disfluency type, with significant reductions in atypical forms of disfluency but no change in typical disfluencies. Predicted differences in fluency effects on the basis of speaker type and dual-task conditions, however, were not observed. Overall findings were most consistent with the CAH and associated hypotheses related to explicit control of motor performance. In line with this view and its notion of dechunking, the comparable effects across speakers, specifically in atypical disfluencies, suggest that applying more WM resources to the process of speaking contributes to fluency breakdown, regardless of stuttering status.

The lack of a differential effect by WM domain indicated that similar fluency changes occurred regardless of the nature of the stimuli being processed in WM. An alternative possibility is that spatial tasks may have unintentionally taxed verbal WM if verbal rehearsal strategies were utilized. It is not clear whether most participants applied such strategies, but if they did, both secondary tasks may have essentially taxed verbal WM resources, with similar fluency effects being a logical outcome.

Results also indicated similar fluency changes regardless of the extent to which WM was engaged or how frequently it was accessed. It is possible that more extreme changes in WM load or ISI were necessary to observe an effect of these manipulations on speech fluency. In a recent dual-task study involving concurrent rhyme judgments and a WM task, for example, vulnerability to interference (in the group of PWS) was apparent only in the most demanding WM condition, which involved recalling items in a set size of five (R. M. Jones et al., 2012).

Dual-Task Effects on Speaking Rate

Rate results indicated that AWS produced speech at a slower rate than AWNS; all speakers showed reduced speech rate during dual compared with baseline tasks; and compared with AWNS, AWS showed a smaller reduction in rate under dual-task conditions. Rate changes in dual-task conditions provide a context for understanding fluency effects and address an important limitation of many previous studies on this topic, which omit these details (Bosshardt, 1999; Oomen & Postma, 2001; Vasiç & Wijnen, 2005). The cause of rate reduction is not entirely clear but may be related to dual-task effects on linguistic productivity (e.g., reduction in number of propositional units per sentence and in complete sentences, as reported, respectively, by Bosshardt, 2006; Jou & Harris, 1992); rate reduction seems to occur specifically when secondary tasks engage attentional or WM resources. Consistent with this interpretation, an early study combining speech production with external distractions (visual and auditory noise) showed no resulting changes in speech rate (Mallard & Webb, 1980); however, generating random sequences of finger tapping while speaking (a manual version of the random-digit generation task, which cannot be performed automatically and demands ongoing attention) significantly reduced syllable counts (Oomen & Postma, 2002).

The smaller reduction in speech rate for AWS was an interesting finding that has not been reported in prior studies. This result likely reflects combined effects of a habitually slower speech rate in AWS (e.g., Bloodstein & Bernstein Ratner, 2008; for relevant details on speech-rate calculation procedures, see also Guitar, 2006, pp. 193–194) and enhanced fluency under dual-task conditions. From the available data, it is not possible to determine whether reductions in rate potentially underlie fluency changes or whether both changes are independently caused by dual-task conditions but are not causally linked. Either way, the findings suggest that if there is an association between speech rate and fluency, this relationship is specific to atypical disfluencies, because typical disfluencies were not affected by dual-task conditions or by reductions in rate.

Dual-Task Effects on Overt Errors

Results showed no change in error frequency under dual compared with baseline conditions. Contrary to the predictions of the CRH (Postma & Kolk, 1993), this finding indicated that performing the WM task did not interfere with speakers' ability to simultaneously monitor and correct covert errors arising during speech planning. Previous findings are inconsistent, with evidence of increased error rates (Oomen & Postma, 2002), reduced error rates (Maraist & Hutton, 1957), and no change at all (Oomen & Postma, 2001) under dual-task conditions. This lack of agreement is at least partly due to the vague specifications of what constitutes an error in each study, but is also likely a result of the wide variation in speaking and secondary tasks, as previously discussed. Another relevant consideration is that error counts in the present data were exceedingly small (approximately 1–2 mean errors per trial across conditions in both groups). Vasiç and Wijnen (2005) likewise reported small error counts and, on the basis of this outcome, discarded error measures from their analyses altogether. These results, combined with previous data on this topic, also highlight the considerable challenge involved in studying speech errors, particularly covert errors, which by definition cannot be reliably detected or measured, and are poorly understood. Overall, the relationship of covert errors to stuttering is supported by minimal evidence from prior research, and although there are data suggesting an association between subtle phonological encoding deficits and stuttering (Anderson & Byrd, 2008; R. M. Jones et al., 2012; Sasisekaran & Weisberg, 2014), the implications of these deficits are not clear, especially because these weaknesses are not functionally significant (for a review, see Nippold, 2002).

Optimal Levels of Cognitive Control: When Less Can Be More and More Can Be Less

The present findings contribute to an expanding literature on attention and performance (Beilock & Gray, 2012; Gray, 2004; Masters, 1992; Wulf & Lewthwaite, 2010) and demonstrate that maximizing explicit forms of processing, which recruit WM resources, does not always benefit task performance. This principle is a central premise of the *matched filter hypothesis* (Chrysikou, Weber, & Thompson-Schill, 2014), which proposes that optimal levels of cognitive control vary depending on task demands. High levels of top-down control are necessary to efficiently filter sensory information in tasks that are rule driven or involve conflict; however, the same form of control hinders performance on procedural tasks (e.g., language acquisition, implicit motor learning) that are best served by subcortical (e.g., basal ganglia) systems. Thus, optimal performance relies on dynamic adjustments to the filtering mechanism according to task requirements.

Our finding that atypical disfluencies decreased under dual-task conditions further suggests that speech production is ideally managed by implicit, rather than explicit, modes of control. Neuroimaging data indicate that early stages of task learning are characterized by the engagement of frontal regions and specific subcortical areas (associative striatum) that receive input from the prefrontal cortex but that these regions show lower activation once a task has been automatized (Ashby, Turner, & Horvitz, 2010; Poldrack et al., 2005). In a similar vein, regions within the basal ganglia show a decrease in activity on trials that form a sequence compared with pseudorandom trials, reflecting the effective chunking of information that accompanies automatization (Poldrack et al., 2005).

Two key findings suggest that stuttering is associated with differences in this neural circuitry. Giraud et al. (2008) found a positive correlation between basal-ganglia activity and stuttering severity, implying that representations were less efficiently organized and resembled early stages of motor learning in these speakers. Imaging studies further reveal increased ACC activity during conflict tasks in PWS compared to PWNS, indicating that PWS demonstrate excessive monitoring activity, even when behavioral performance does not show impairment relative to PWNS (Arnstein et al., 2011; Liu et al., 2014). These findings suggest that stuttering may be associated with inappropriate matching of task demands with the extent or type (implicit/explicit) of cognitive resources utilized to meet these demands.

The results of the present study indicate that greater availability of WM resources was associated with a greater propensity toward stutterlike disfluencies, whereas engaging WM resources in secondary tasks resulted in fluency enhancement. This outcome suggests that speech production, like skilled motor performance across a variety of tasks, benefits from minimized attentional control and maximized automatization. Recruitment of WM resources during speech production may reflect the speaker's transition to an explicit, rule-based processing system in an effort to consciously control articulatory movements. The similar benefit observed for both groups when WM was engaged further suggests that applying WM resources to speech production tends to interfere with fluency and results in dechunking of motor speech sequences, regardless of whether a stuttering disorder is present.

Limitations

Several limitations in the present design require consideration. Within the secondary task, more closely matched stimuli and structures for the verbal and spatial tasks may have resulted in a clearer domain effect, and setting the task at individual capacity (see, e.g., Cocchini et al., 2002) may have ensured that WM was fully engaged in each participant. High levels of accuracy, even under dual conditions, also suggest that WM tasks were performed with some degree of automaticity and may have diminished dual-task effects. Studies incorporating more demanding secondary tasks may demonstrate more pronounced effects on speech-related variables. It is important to note that a control condition in the secondary task with no WM load could clarify whether the effects observed necessarily involve WM or reflect more general resources, such as sustained attention. Although the vigilance task resembled a sustained-attention task, it was not designed as a control condition and was too different in general structure to be used for comparison purposes.

In the speaking task, speech preparation time was not controlled. This allowed selection of topics for each trial but made it difficult to determine whether language formulation and conceptual planning differed between groups or contributed to results. Comparison of mean topic-selection time for the two speaker groups within baseline and dual conditions suggested this was not the case—baseline: $t(37) = -0.15$, $p = .88$; dual: $t(37) = 0.09$, $p = .93$. Nevertheless, individual differences in speech preparation time may have influenced performance in ways not detected in aggregated data. In addition, although a spontaneous speaking task was used to induce naturally occurring stuttering rates, important aspects of spoken output (emotionality of content, vocabulary, syntactic complexity, level of abstraction) could not be controlled and may have introduced noise. Last, prior dual-task studies (e.g., Bosshardt, 2002, 2006) have shown that generating sentences while performing a secondary task resulted in fewer propositions. Dual-task conditions in the present study may have similarly influenced aspects of linguistic productivity (e.g., semantic diversity, syntactic complexity, word frequency) but were not the focus of our analyses and were not considered.

Future Research

Although data implied that fluency effects were related to WM, additional research is needed to determine whether engaging general attentional resources produces similar outcomes. The specific role of processes related to performance monitoring and self-evaluation also remains unclear. The underlying neurocognitive mechanisms contributing to these functions, and potentially to speech disfluency, require further elucidation. An important direction for future studies on this topic is the application of dual-task paradigms to a younger population. Examining whether maladaptive forms of conscious control are evident near the onset of stuttering would clarify the role of these processes in the development of fluency disorders; however, dual-task paradigms would require considerable modification to be adapted for use with younger children. Future studies assessing the role of emotional and attentional biases during early stuttering are also necessary to fully understand how and when such patterns arise.

Clinical Implications

Although preliminary, the present findings suggest that clinical intervention for stuttering may be enhanced by strategies that induce more implicit forms of speech production. Neuroimaging studies of AWS after treatment show a reduction in ACC activity (De Nil & Kroll, 2001; Neumann et al., 2005), suggesting changes in anticipatory behavior and self-monitoring tendencies immediately following intensive treatment. It is not clear, however, whether these changes are maintained. Subjective reports from long-term posttreatment follow-up indicate that many PWS continue to consciously attend to their speech (Boberg & Kully, 1994). Therapeutic applications of dual tasking are emerging in general rehabilitation research (e.g., Plummer, Villalobos, Vayda, Moser, & Johnson, 2014) but have only been examined in one study of AWS (Metten et al., 2011, experiment 3), which was based on subjective reports from an insufficient sample ($N = 3$). Further research is needed to determine whether and how dual-task methods can supplement traditional approaches to stuttering treatment. A possible approach could involve the incorporation of secondary tasks within the structure of a comprehensive stuttering intervention to facilitate automatization of newly learned speech motor patterns that support fluency. Additional clinical research will be needed to determine whether dual-task methods can successfully generalize to functional communication situations.

The present results can also potentially inform early intervention approaches for stuttering. If incipient stuttering at least partly stems from conscious attempts to control articulatory movements, this cyclic process could be interrupted at its start by helping young children divert attention away from the process of speech production. Thus, clarifying the role of cognitive and attentional mechanisms in the emergence of atypical disfluencies could provide parents and clinicians with a novel and practically useful perspective on early stuttering.

Summary and Conclusion

In conclusion, the present study contributes to the growing body of literature on attention and motor performance by extending the concept of "less is more" to the process of speech production. Our findings demonstrate that engaging WM while speech was produced caused a significant reduction in atypical disfluencies in all speakers. Overall, the data did not support speech-monitoring accounts of stuttering (CRH, VCH). Suppressing WM resources resulted in fluency changes but no accompanying changes in the frequency of speech errors, contradicting the CRH, which implies a trade-off between disfluencies and overt errors (Postma & Kolk, 1993). The similarity of fluency outcomes for both groups was also inconsistent with the VCH, which claims that hypervigilant speech monitoring is unique to PWS (Vasiç & Wijnen, 2005). Much like the axiomatic view that stuttering is essentially what a speaker does to avoid stuttering, the findings indicate that speech fluency is enhanced by less effort and compromised by more. Further research is needed to fully understand the cognitive mechanisms involved in this effect and its potential clinical implications.

Acknowledgments

This study was funded by a Graduate School University Center Dissertation Fellowship awarded to the first author. We also wish to acknowledge Scott Yaruss for his valuable input and feedback related to this research, Claire Leake for her assistance with transcription and coding, and all the individuals who participated in this study.

References

Ambrose, N. G., & Yairi, E. (1999). Normative disfluency data for early childhood stuttering. *Journal of Speech, Language, and Hearing Research, 42,* 895–909.

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

Arnstein, D., Lakey, B., Compton, R. J., & Kleinow, J. (2011). Preverbal error-monitoring in stutterers and fluent speakers. *Brain and Language, 116,* 105–115.

Ashby, F. G., Turner, B. O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in Cognitive Sciences, 14,* 208–215.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4,* 829–839. doi:10.1038/nrn1201

Baddeley, A. D., & Hitch, G. (1974). Working memory. *The Psychology of Learning and Motivation, 8,* 47–89.

Bajaj, A. (2007). Working memory involvement in stuttering: Exploring the evidence and research implications. *Journal of Fluency Disorders, 32,* 218–238.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 (Version 1.1-6) [Computer software]. Retrieved from http://cran.r-project.org/package=lme4

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46,* 610–620.

Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied, 8,* 6–16.

Beilock, S. L., & Gray, R. (2012). From attentional control to attentional spillover: A skill-level investigation of attention, movement, and performance outcomes. *Human Movement Science, 31,* 1473–1499.

Blomgren, M. (2013). Behavioral treatments for children and adults who stutter: A review. *Psychology Research and Behavior Management, 6,* 9–19.

Bloodstein, O., & Bernstein Ratner, N. (2008). *A handbook on stuttering* (6th ed.). Clifton Park, NY: Thomson Delmar Learning.

Boberg, E., & Kully, D. (1994). Long-term results of an intensive treatment program for adults and adolescents who stutter. *Journal of Speech and Hearing Research, 37,* 1050–1059.

Bosshardt, H.-G. (1999). Effects of concurrent mental calculation on stuttering, inhalation and speech timing. *Journal of Fluency Disorders, 24,* 43–72.

Bosshardt, H.-G. (2002). Effects of concurrent cognitive processing on the fluency of word repetition: Comparison between persons who do and do not stutter. *Journal of Fluency Disorders, 27,* 93–114.

Bosshardt, H.-G. (2006). Cognitive processing load as a determinant of stuttering: Summary of a research programme. *Clinical Linguistics & Phonetics, 20,* 371–385.

Brocklehurst, P. H. (2008). *Exploring the relationships between perfectionism, speech-monitoring and disfluency in the speech of people who do and do not stutter* (Unpublished master's thesis). The University of Edinburgh, Scotland.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2010). *Test of Nonverbal Intelligence, Fourth Edition*. Austin, TX: Pro-Ed.

Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: An ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping, 25,* 105–117.

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.

Chrysikou, E. G., Weber, M. J., & Thompson-Schill, S. L. (2014). A matched filter hypothesis for cognitive control. *Neuropsychologia, 62,* 341–355.

Cocchini, G., Logie, R. H., Della Sala, S., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition, 30,* 1086–1095.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 169,* 323–338.

Declerck, M., & Kormos, J. (2012). The effect of dual task demands and proficiency on second language speech production. *Bilingualism: Language and Cognition, 15,* 782–796.

De Nil, L., & Bosshardt, H.-G. (2000). Studying stuttering from a neurological and cognitive information processing perspective. In H.-G. Bosshardt, J. S. Yaruss, & H. F. M. Peters (Eds.), *Stuttering: Research, therapy, and self-help. Proceedings of the Third World Congress on Fluency Disorders* (pp. 53–58). Nijmegen, the Netherlands: Nijmegen University Press.

De Nil, L. F., & Kroll, R. M. (2001). Searching for the neural basis of stuttering treatment outcome: Recent neuroimaging studies. *Clinical Linguistics & Phonetics, 15,* 163–168.

De Nil, L. F., Kroll, R. M., Kapur, S., & Houle, S. (2000). A positron emission tomography study of silent and oral single word reading in stuttering and nonstuttering adults. *Journal of Speech, Language, and Hearing Research, 43,* 1038–1053.

Dietrich, A., & Stoll, O. (2010). Effortless attention, hypofrontality, and perfectionism. In B. Bruya (Ed.), *Effortless attention: A new perspective in the cognitive science of attention and action* (pp. 160–178). Cambridge, MA: MIT Press.

Duff, S. C., & Logie, R. H. (2001). Processing and storage in working memory span. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54*(A), 31–48.

Freedman, S. E., Maas, E., Caligiuri, M. P., Wulf, G., & Robin, D. A. (2007). Internal versus external: Oral-motor performance as a function of attentional focus. *Journal of Speech, Language, and Hearing Research, 50,* 131–136.

Giraud, A.-L., Neumann, K., Bachoud-Levi, A.-C., von Gudenberg, A. W., Euler, H. A., Lanfermann, H., & Preibish, C. (2008). Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. *Brain and Language, 104,* 190–199.

Gray, R. (2004). Attending to the execution of a complex sensorimotor skill: Expertise differences, choking, and slumps. *Journal of Experimental Psychology: Applied, 10,* 42–54.

Guitar, B. (2006). *Stuttering: An integrated approach to its nature and treatment* (3rd ed.). Baltimore, MD: Lippincott, Williams & Wilkins.

Hakim, H. B., & Bernstein Ratner, N. (2004). Nonword repetition abilities of children who stutter: An exploratory study. *Journal of Fluency Disorders, 29,* 179–199.

Hardin, J. W., & Hilbe, J. M. (2007). *Generalized linear models and extensions* (2nd ed.). College Station, TX: Stata.

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods, 39,* 101–117.

Jones, M., Dobson, A., Onslow, M., & Carey, B. (2009). Negative binomial mixed models for analysis of stuttering rates. *Computational Statistics & Data Analysis, 53,* 4590–4600.

Jones, R. M., Fox, R. A., & Jacewicz, E. (2012). The effects of concurrent cognitive load on phonological processing in adults who stutter. *Journal of Speech, Language, and Hearing Research, 55,* 1862–1875.

Jou, J., & Harris, R. J. (1992). The effect of divided attention on speech production. *Bulletin of the Psychonomic Society, 30,* 301–304.

Leclercq, M. (2002). Theoretical aspects of the main components and functions of attention. In M. Leclercq & P. Zimmermann (Eds.), *Applied neuropsychology of attention: Theory, diagnosis, and rehabilitation* (pp. 3–55). New York, NY: Psychology Press.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14,* 41–104.

Lickley, R. J., Hartsuiker, R. J., Corley, M., Russell, M., & Nelson, R. (2005). Judgment of disfluency in people who stutter and people who do not stutter: Results from magnitude estimation. *Language and Speech, 48,* 299–312.

Liu, J., Wang, Z., Huo, Y., Davidson, S. M., Klahr, K., Herder, C. L., . . . Peterson, B. S. (2014). A functional imaging study of self-regulatory capacities in persons who stutter. *PLoS ONE, 9*(2), e89891.

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes, 66,* 309–332.

Mallard, A. R., & Webb, W. G. (1980). The effects of auditory and visual "distractors" on the frequency of stuttering. *Journal of Communication Disorders, 13,* 207–212.

Maraist, J. A., & Hutton, C. (1957). Effects of auditory masking upon the speech of stutterers. *Journal of Speech and Hearing Disorders, 22,* 385–389.

Martin, N. A., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test, Fourth Edition*. Novato, CA: Academic Therapy.

Masters, R. S. W. (1992). Knowledge, knerves, and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology, 83,* 343–358.

McKay, B., & Wulf, G. (2012). A distal external focus enhances novice dart throwing performance. *International Journal of Sport and Exercise Psychology, 10,* 149–156.

McNevin, N. H., Shea, C. H., & Wulf, G. (2003). Increasing the distance of an external focus of attention enhances learning. *Psychological Research, 67,* 22–29.

Melnick, K. S., Conture, E. G., & Ohde, R. N. (2003). Phonological priming in picture naming of young children who stutter. *Journal of Speech, Language, and Hearing Research, 46,* 1428–1443.

Metten, C., Bosshardt, H.-G., Jones, M., Eisenhuth, J., Block, S., Carey, B., … Menzies, R. (2011). Dual tasking and stuttering: From the laboratory to the clinic. *Disability and Rehabilitation, 33,* 933–944.

Neumann, K., Preibisch, C., Euler, H. A., von Gudenberg, A. W., Lanfermann, H., Gall, V., & Giraud, A.-L. (2005). Cortical plasticity associated with stuttering therapy. *Journal of Fluency Disorders, 30,* 23–39.

Nippold, M. A. (2002). Stuttering and phonology: Is there an interaction? *American Journal of Speech-Language Pathology, 11,* 99–110.

Oberauer, K. (2009). Design for a working memory. *The Psychology of Learning and Motivation, 51,* 45–100.

Oomen, C. C. E., & Postma, A. (2001). Effects of divided attention on the production of filled pauses and repetitions. *Journal of Speech, Language, and Hearing Research, 44,* 997–1004.

Oomen, C. C. E., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes, 17,* 163–184.

Pashler, H. (1992). Attentional limitations in doing two tasks at the same time. *Current Directions in Psychological Science, 1,* 44–48.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220–244.

Pashler, H. E. (1999). Automaticity, effort, and control. In *The psychology of attention* (Ch. 8, pp. 357–398). Cambridge, MA: MIT Press.

Plummer, P., Villalobos, R. M., Vayda, M. S., Moser, M., & Johnson, E. (2014). Feasibility of dual-task gait training for community-dwelling adults after stroke: A case series. *Stroke Research and Treatment, 2014,* 538602.

Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The neural correlates of motor skill automaticity. *The Journal of Neuroscience, 25,* 5356–5364.

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.

R Core Team. (2013). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing.

Riley, G. D. (2009). *Stuttering Severity Instrument–Fourth Edition*. Austin, TX: Pro-Ed.

Rinker, T. W. (2013). qdap: Quantitative Discourse Analysis Package (Version 2.2.4) [Computer software]. Buffalo, NY: University at Buffalo.

Rousseau, I., Onslow, M., Packman, A., & Jones, M. (2008). Comparisons of audio and audiovisual measures of stuttering frequency and severity in preschool-age children. *American Journal of Speech-Language Pathology, 17,* 173–178.

Ryan, B. P. (1992). Articulation, language, rate, and fluency characteristics of stuttering and nonstuttering preschool children. *Journal of Speech and Hearing Research, 35,* 333–342.

Salthouse, T. A., Babcock, R. L., & Shaw, R. J. (1991). Effects of adult age on structural and operational capacities in working memory. *Psychology and Aging, 6,* 118–127.

Saltuklaroglu, T., Teulings, H.-L., & Robbins, M. (2009). Differential levels of speech and manual dysfluency in adults who stutter during simultaneous drawing and speaking tasks. *Human Movement Science, 28,* 643–654.

Sasisekaran, J., De Nil, L. F., Smyth, R., & Johnson, C. (2006). Phonological encoding in the silent speech of persons who stutter. *Journal of Fluency Disorders, 31,* 1–21.

Sasisekaran, J., & Weisberg, S. (2014). Practice and retention of nonwords in adults who stutter. *Journal of Fluency Disorders, 41,* 55–71.

Slotnik, D. E., & Schulten, K. (2012, June 12). 163 questions to write or talk about [Blog post]. Retrieved from http://learning.blogs.nytimes.com/2012/06/12/163-questions-to-write-or-talk-about/

Smits-Bandstra, S., & De Nil, L. (2009). Speech skill learning of persons who stutter and fluent speakers under single and dual task conditions. *Clinical Linguistics & Phonetics, 23,* 38–57.

Snyder, K. M., & Logan, G. D. (2013). Monitoring-induced disruption in skilled typewriting. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1409–1420.

Tumanova, V., Conture, E. G., Lambert, E. W., & Walden, T. A. (2014). Speech disfluencies of preschool-age children who do and do not stutter. *Journal of Communication Disorders, 49,* 25–41.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114,* 104–132.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498–505.

Vasiç, N., & Wijnen, F. (2005). Stuttering as a monitoring deficit. In R. J. Hartsuiker, R. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 226–247). Hove, United Kingdom: Psychology Press.

Vincent, I., Grela, B. G., & Gilbert, H. R. (2012). Phonological priming in adults who stutter. *Journal of Fluency Disorders, 37,* 91–105.

Weber-Fox, C., Spruill, J. E., III, Spencer, R., & Smith, A. (2008). Atypical neural functions underlying phonological processing and silent rehearsal in children who stutter. *Developmental Science, 11,* 321–337.

Wingate, M. E. (1962). Evaluation and stuttering: III. Identification of stuttering and the use of a label. *Journal of Speech and Hearing Disorders, 27,* 368–377.

Wulf, G., & Lewthwaite, R. (2010). Effortless motor learning?: An external focus of attention enhances movement effectiveness and efficiency. In B. Bruya (Ed.), *Effortless attention: A new perspective in the cognitive science of attention and action* (pp. 75–101). Cambridge, MA: MIT Press.

Yaruss, J. S. (1997). Clinical measurement of stuttering behaviors. *Contemporary Issues in Communication Science and Disorders, 24,* 33–44.

Yaruss, J. S., & Quesal, R. W. (2006). Overall Assessment of the Speaker's Experience of Stuttering (OASES): Documenting multiple outcomes in stuttering treatment. *Journal of Fluency Disorders, 31,* 90–115.