Effects of the SpeechEasy on Objective and Perceived Aspects of Stuttering: A 6-Month, Phase I Clinical Trial in Naturalistic Environments

Ryan Pollard
John B. Ellis
Don Finan
Peter R. Ramig
University of Colorado at Boulder

Purpose: Effects of the SpeechEasy when used under extraclinical conditions over several months were investigated. Primary purposes were to help establish Phase I level information about the therapeutic utility of the SpeechEasy and to compare those results with previous findings obtained in laboratory and clinical settings.

Method: Eleven adults who stutter participated. A nonrandomized ABA group design was utilized. Speech samples were collected every 2 weeks in extraclinical environments. Qualitative data were collected through weekly written logs and an exit questionnaire.

Results: Group analyses revealed a statistically significant effect of the SpeechEasy immediately postfitting but no treatment effect across 4 months' time. Individual responses varied greatly with regard to stuttering frequency and subjective impressions. Relatively more stuttering reduction occurred during oral reading than during formulated speech.

Conclusions: Based on this protocol, Phase II trials are not indicated. However, positive individual responses and self-reports suggest some clinical utility for the SpeechEasy. The use of more challenging sampling procedures strengthened external validity and captured more modest altered auditory feedback effects compared with those previously reported in laboratory settings. Device use coincided more so with positive subjective impressions than with measurable fluency improvement, highlighting challenges facing clinicians when implementing principles of evidence-based practice, including client-based preferences.

KEY WORDS: stuttering, treatment, prosthetic devices, clinical trials, treatment outcomes

V arious modalities of altered auditory feedback (AAF) have been used for several decades to help reduce stuttering (Adamczyk, 1959; Soderberg, 1969; Sutton & Chase, 1961; Stromsa, 1958). Early generations of AAF machines were quite large and heavy and were necessarily confined to the laboratory or clinic room. As technology improved, the application of AAF during everyday life was realized in the

late 1950s with the advent of the first portable electronic device—the Derazne Correctophone (Derazne, 1966). Since then, other commercial devices have appeared on the market employing delayed auditory feedback (DAF), frequency altered feedback (FAF), masking noise, or some combination of the three. These include, for example, the Edinburgh Masker (in the late 1970s), the Vocaltech Clinical Vocal Feedback Device (late 1980s), and the Fluency Master (early 1990s), as well as a handful of others (Molt, 2005).

The SpeechEasy, introduced to consumers in 2001, is one of the more recent additions to the succession of wearable electronic devices designed to increase fluency in persons who stutter (PWS). In contrast to the fairly conspicuous and cumbersome nature of earlier electronic fluency aids, its appearance resembles that of a digital hearing aid. This affords it the unique advantage of being small and cosmetically pleasing and thus more appealing to some consumers. The programmable device employs two forms of AAF simultaneously: DAF and FAF. DAF reproduces an acoustic signal with a small time delay, in effect producing an echo of the speaker's voice that is fed back into their ear. FAF increases or decreases the entire frequency range of the acoustic signal, causing the speaker to hear his or her own voice at an altered pitch. Pairing these two procedures is thought to emulate the choral effect, a phenomenon during which stuttering is often reduced or eliminated while speaking in unison with another speech signal (Johnson & Rosen, 1937).

Immediate AAF Effects

The effectiveness of the SpeechEasy was unknown upon its release. Similar to many therapeutic options for the treatment of speech and language disorders, the device was marketed to the public without formal treatment outcome evidence. It was featured in a very positive light on several prominent television shows (e.g., Good Morning America, NBC Nightly News, The Montel Williams Show, The Oprah Winfrey Show) as well as in other media outlets (e.g., *USA Today*), sometimes portrayed as a "miracle cure" for stuttering (e.g., Hudson, 2003). Due to the lack of available evidence at the time, the SpeechEasy's putative merit rested chiefly on previous research demonstrating the ability of AAF to reduce the occurrence of disfluencies in many who stutter. Following on prior work documenting the efficacy of AAF (e.g., Goldiamond, 1965; Ryan & Van Kirk, 1971), a cumulative body of literature leading to the release of the SpeechEasy was published during the 1990s by researchers associated with the East Carolina University Stuttering Research Laboratory, several of whom were responsible for the device's development. The group effects from this research collectively showed that different permutations of AAF produced immediate reductions in overt stuttering in controlled laboratory environments. These studies also reported very robust responses to AAF across speech tasks, with reductions in stuttering frequency ranging from roughly 70% to 90% for DAF conditions, FAF conditions, and combined DAF and FAF conditions (Armson & Stuart, 1998; Hargrave, Kalinowski, Stuart, Armson, & Jones, 1994; Kalinowski, Armson, Roland-Mieszkowski, Stuart, & Gracco, 1993; Kalinowski, Stuart, Sark, & Armson, 1996; Macleod, Kalinowski, Stuart, & Armson, 1995; Stuart, Kalinowski, & Rastatter, 1997; Stuart, Kalinowski, Rastatter, & Lynch, 2002). Importantly, these researchers demonstrated that a slowed speech rate may not be required for an individual to benefit from AAF (Kalinowski et al., 1993; Stuart et al., 2002). This finding suggested that acceptable speech naturalness could be maintained while using the SpeechEasy.

Immediate SpeechEasy Effects

Distinct from the substantial AAF literature that served as the theoretical foundation for the SpeechEasy's development, more contemporary work has begun to investigate the effects of the device itself. Researchers examining the immediate effects of the device in a laboratory setting found that it produced group reductions in stuttering ranging from 30% to 74%, depending on the speech task performed and whether SpeechEasy fitting protocols were followed (Armson, Kiefte, Mason, & De Croos, 2006). Significantly improved fluency during monologue speech was also reported for roughly half of the participants. This finding is noteworthy because it conflicts with earlier FAF results showing widely variable effects during a monologue task (Armson & Stuart, 1998; Ingham, Moglia, Frank, Costello-Ingham, & Cordes, 1997). Another interesting finding from the 2006 study was that, as with much previous AAF research (e.g., Armson & Stuart, 1998; Chase, Sutton, First, & Zubin, 1961; Ingham et al., 1997; Zerneri, 1966), response to the device was idiosyncratic. In the authors' words:

> The degree and pattern of benefit varied greatly across participants. Although a few participants exhibited both a dramatic reduction in stuttering and relative freedom from stuttering in the device conditions, others showed a modest or minimal reduction and continued to exhibit a relatively high level of stuttering. (Armson et al., 2006, p. 149)

Longitudinal Effects of AAF and SpeechEasy

Very few studies have examined the effects of AAF as utilized for extended periods of time. Dewar, Dewar, Austin, and Brash (1979) reported equivocal findings from

a study of the Edinburgh Masker, a wearable electronic device that delivered binaural masking noise. More recently, Van Borsel, Reunes, and Van den Bergh (2003) had participants wear a headphone DAF device during a series of daily and weekly speech tasks (e.g., telephone calls, monologue, oral reading) for a period of 3 months. After 3 months' exposure to DAF, the percentage of stuttered words had decreased by 50% across tasks for the group, and carryover fluency was observed without the device in place.

Regarding long-term effects of the SpeechEasy itself, relatively little is known. Although the device has been reported to have an "80%–90% success rate" on national television shows (e.g., Sloan, 2007), only two published experiments to date have examined the issue of long-term efficacy. Stuart, Kalinowski, Rastatter, Saltuklaroglu, and Dayalu (2004) had 8 participants wear custom-fit devices for a period of 4 months. They found that, while reading aloud and giving a monologue in quiet laboratory rooms, participants' stuttering rates were significantly reduced with the device in place at the initial fitting and remained so 4 months later. In a later article following up on the same cohort of participants under identical speaking conditions, the group displayed maintained fluency enhancement and speech naturalness with the device at 12 months postfitting (Stuart, Kalinowski, Saltuklaroglu, & Guntupalli, 2006). Neither study reported individual speech or self-report data.

Ecological Validity of Speech Samples

Armson et al.'s (2006) call for "further investigations of the clinical benefits of [the] SpeechEasy that focus on its stuttering reduction properties in formulated speech" (p. 140) illustrates a salient point: namely, that little is presently known about the benefits of the device during speech tasks other than oral reading. Although research has examined the device during monologue speech (e.g., Stuart et al., 2004, 2006), little work to date has studied its effectiveness during conversational speech. In fact, other than Armson et al.'s (2006) report of slightly more improved fluency during conversation than during monologue, no experimental evidence exists for the effects of any modality of AAF during conversational speech (Lincoln, Packman, & Onslow, 2006). There is also a dearth of evidence on the effectiveness of AAF during activities of daily living and in extraclinical settings. In support of the SpeechEasy's ecological utility, fluency induction under AAF has been documented under two challenging conditions: while speaking in front of audiences of varying sizes (Armson, Foote, Witt, Kalinowski, & Stuart, 1997; Kalinowski, Stuart, Wamsley, & Rastatter, 1999) and while engaging in scripted telephone calls (Zimmerman, Kalinowski, Stuart, & Rastatter, 1997).

The Armson et al. and Zimmerman et al. studies have been cited as evidence that "the robust effects of altered auditory feedback occur outside the laboratory environment" (Stuart et al., 2003, p. 233). Others, however, claim that recent reports have generally failed to show that the SpeechEasy effectively and reliably improves fluency in everyday speaking situations (Bothe, Finn, & Bramlett, 2007; Finn, Bothe, & Bramlett, 2005). Another issue involves the consistency of AAF effects over time and whether those effects are perishable (Ingham et al., 1997). Finn et al. (2005) summarize such concerns by stating that "no published clinical research is available about this device that describes its long-term, real-world effectiveness and efficiency with a wide range of persons who stutter" (pp. 178–179). Clearly, much more remains to be known as to the fluency-inducing characteristics of AAF during the types of challenging conditions PWS are likely to encounter in their day-to-day lives. It has not yet been determined whether the mitigating effects of AAF seen in the laboratory and clinic room during mainly oral reading will generalize to more quotidian speech tasks and authentic speaking environments and will be maintained across time.

The present study investigates quantitative and qualitative effects of the SpeechEasy when used under challenging, extraclinical conditions over an extended period of time. The primary purposes are to help establish Phase I level information about the therapeutic viability of AAF as delivered by the SpeechEasy and to compare these results with previous findings obtained in laboratory and clinical settings. Phase I treatment outcome research is designed to establish whether a therapeutic effect exists, to estimate its potential magnitude, and to help identify efficient treatment protocols (Robey, 2004). Individual responses are documented and described in order to better characterize any idiosyncratic effects manifested in overt disfluencies or attitudinal/ emotional aspects of stuttering. Consequently, we report both group and individual data in an effort to facilitate transparency and to provide useful information to clients and clinicians concerned with making individualized treatment decisions. To determine whether prolonged SpeechEasy use affects more than the observable symptoms of stuttering, we included qualitative instruments and participant self-report to augment objective stuttering counts. As Yaruss and Quesal (2006) contend, "there is a compelling and immediate need for research on the outcomes of treatment that addresses aspects of the stuttering disorder beyond the surface speech behaviors" (p. 92). Questions are addressed in the following areas: overall effectiveness of the device, differential effects across speech tasks, variability in response profiles, and congruity between subjective impressions of the device and objective speech performance.

Method

Participants

This study was approved by the University of Colorado Human Research Committee. Eleven adult PWS were recruited from the Colorado Front Range region. Participants were recruited through flyers posted around the Boulder/Denver area, e-mails sent to local speech-language pathologists, postings placed on Craigslist (http://www.craigslist.com), and a local stuttering support group. The sample consisted of 6 males and 5 females, ranging in age from 18 to 62 years (M = 34.2). Other than stuttering, no other speech and language disorders were reported. All participants had received stuttering treatment in the past and, if they wished, were permitted to attend the local stuttering support group while enrolled in the study. Only 1 participant (S5) chose to attend the group, and he did so irregularly. Our criteria for inclusion was as follows: (a) participants could not be enrolled in formal stuttering therapy while involved in the study (S5 was being seen once a week for maintenance therapy and was allowed to participate because he was at the end of his therapy program); (b) participants had to have normal pure-tone audiometric thresholds of 25 dB HL or better at all octave frequencies from 500 to 8000 Hz (American National Standards Institute, 1996); (c) participants had to feel that they received enough benefit from the SpeechEasy at the initial fitting to warrant wearing the device regularly for 4 months; and (d) participants had to commit to wearing the device as much as possible during the treatment phase, the stated guideline being at least 5 hr per day. The first 11 adults who met the selection criteria were enrolled in the study. Informed consent was obtained from all participants prior to enrollment. Participant characteristics are summarized in Table 1.

Apparatus

SpeechEasy In-the-Canal units were used by all participants in the project. The design and operating characteristics are detailed elsewhere (Stuart et al., 2003). The In-the-Canal unit fits into the ear canal of the wearer and comes equipped with an external volume control and dual memory settings. The dual settings are intended to permit wearers to toggle between two distinct settings using a button located on the device's exterior. During this study, the second memory setting was always programmed with no frequency shift and no time delay for those instances when participants wished to listen to others without filtering, such as in a loud environment. Unless otherwise indicated by the participant, the device was always fit into the ear opposite the one used when speaking on the telephone.

The SpeechEasy interfaces with a computer via AudioPro hardware (Micro-DSP). During the study, SpeechMaster software Version 1.4.1 was run on a laptop computer. The computer was connected to the AudioPro hardware via a serial port connection and universal serial bus adapter. Using the SpeechMaster software, the SpeechEasy device can have DAF, FAF, and internal volume settings adjusted according to user preference. All participants used DAF and FAF simultaneously (see Table 1). DAF settings can be programmed to a delay of 0 to 220 ms in 1-ms increments. Regarding FAF, the technology employed by the SpeechEasy shifts the entire acoustic signal up or down in increments of 500 Hz with a maximum shift of 2 kHz. The internal gain of the device can be adjusted in six steps of 5 dB each. The software also includes a 16-band equalizer to adjust attenuation of specific frequencies (200–7800 Hz) up to 20 dB each. This spectral shaping feature was used on 2 participants' devices during the study.

Table 1. Participant characteristics.

Subject	Age	Sex	Severity (SSI-3)	Severity (OASES)	Device settings	Device wear time (hr/day)
S1	62	F	mild (19)	mod to severe (60)	100 ms; +500 Hz	4.4
S2	35	M	very severe (41)	mod to severe (73)	100 ms; +500 Hz	10.4
S3	25	M	very severe (41)	moderate (59)	60 ms; +1000 Hz	5.1
S4	32	F	severe (36)	moderate (57)	60 ms; +500 Hz	2.3
S5	57	M	very mild to mild (17)	moderate (56)	60 ms; +500 Hz	6.3
S6	22	F	severe (36)	moderate (52)	60 ms; +500 Hz	—
S7	18	M	moderate (26)	mild to mod (41)	60 ms; +500 Hz	4.0
S8	27	F	very severe (42)	mod to severe (66)	90, 80 ms; +500, 0 Hz	6.4
S9	54	F	very mild (15)	moderate (48)	60, 50 ms; −500 Hz	6.4
S10	18	M	mild to mod (24)	moderate (52)	40 ms; +500 Hz	4.0
S11	26	M	very mild (8)	mild (28)	50 ms; +500 Hz	1.1

Note. Dashes indicate that data were not obtained. SSI-3 = Stuttering Severity Instrument for Children and Adults–Third Edition; OASES = Overall Assessment of the Speaker's Experience of Stuttering; F = female; M = male; mod = moderate.

All speech samples for the Stuttering Severity Instrument for Children and Adults – Third Edition (SSI-3; Riley, 1994) were recorded in a sound-treated room using a VHS video camera with built-in microphone. All speech samples collected outside the laboratory were recorded using an inconspicuous lavaliere microphone connected to a portable digital recorder. Digital .wav files of each sample were then transferred to a laptop computer for analysis.

Procedure

Design

The study utilized a nonrandomized ABA group design with a predetermined protocol allowing detailed exploration of individual response profiles. The design incorporated a 1-month baseline phase (without device), a 4-month treatment phase (with device), and a 1-month withdrawal phase (without device). Speech samples were collected approximately every 2 weeks (±1 week) throughout the project. The timeline for speech samples was as follows: baseline (B0, B2, B4), treatment (T0, T2, T4, T6, T8, T10, T12, T14, T16), and withdrawal (W0, W2, W4). One month was chosen as the duration for the baseline phase because this is approximately the amount of time it takes for a custom-fit device to be fabricated and returned from the manufacturer.

Baseline phase. At the first visit, the study protocol was reviewed with participants, and informed consent was given. Participants completed a hearing screening and then were led back into the laboratory to complete the SSI-3. Participants were then fit with a demonstration device by a trained SpeechEasy provider (the second author), and SpeechEasy fitting protocols were followed (*SpeechEasy Training Manual,* 2006). During oral reading and conversation with the experimenters, device settings were adjusted according to participant preference. The experimenters then took participants through the department building to engage in conversation with strangers and to experience the device outside of an optimally quiet setting. Afterwards, participants were asked whether they felt they received sufficient benefit from the device to agree to wear it continually for 4 months. To increase the external validity of results, no minimum level of stuttering reduction was required for inclusion in the study. If participants desired to continue on with the study after trying the device, an ear impression was taken by a trained doctoral student in audiology. Last, participants were accompanied to an extraclinical environment to give their first speech sample. During the second and third baseline phase visits (B2 and B4), the Overall Assessment of the Speaker's Experience of Stuttering (OASES; Yaruss & Quesal, 2006) and Perceptions of Stuttering Inventory (PSI; Woolf, 1967) were administered (one instrument at B2, one instrument at B4, counterbalanced across participants). The administration of the two self-report

instruments was distributed over the baseline phase to shorten the duration of the initial visit. After giving their third baseline phase speech sample at B4, participants began the treatment phase.

Treatment phase. Upon returning to the laboratory at B4, participants were fitted with their devices by the second author, a trained SpeechEasy provider, and SpeechEasy fitting protocols were followed precisely to finalize optimal settings and acclimate participants to the device. This procedure is described elsewhere (Armson et al., 2006; *SpeechEasy Training Manual,* 2006). Briefly, it includes instructing the wearer to attend to the second speech signal and teaching several active techniques to alter one's speech pattern, such as easy vocal onsets, prolongations, continuous phonation, starter sounds (e.g., initial /m/ or schwa), and fillers (e.g., /m/ or schwa inserted between words). Participants were told that these active strategies could be introduced at their discretion to help initiate voicing and/or enhance responsiveness to the second speech signal and could be reduced if they felt choral effects and naturalness stabilize (Ramig, Ellis, Pollard, & Finan, in press). No other instructions of a therapeutic nature were given at this time, although participants were informed that the techniques described in their copy of the SpeechEasy manual could be reviewed at subsequent meetings if they wished, and they were encouraged to practice them on their own. If a participant requested further instruction on any active technique(s) during the treatment phase, this instruction was always carried out after a speech sample was collected. In general, briefer DAF delay settings (e.g., 50–60 ms) were encouraged in order to retain acceptable levels of speech naturalness. Final settings, however, were ultimately based on participant preference. Participants were also told that settings could be changed at any point during the treatment phase if they wished. Similar to Van Borsel et al. (2003), participants were encouraged to wear their devices at least 5 hr per day. However, in order to obtain unqualified data on SpeechEasy usage patterns, this was not made a requirement but was instead a guideline. Next, participants completed the SSI-3 in the laboratory while wearing the device. They were then taken to an extraclinical setting, and the first treatment phase speech sample was collected. During subsequent treatment visits (i.e., T2–T16), participants returned every 2 weeks to give speech samples. Weekly written logs were completed during the treatment phase as well. After giving their ninth treatment phase sample at T16, participants were taken to the laboratory to complete the SSI-3 with the device inserted and begin the withdrawal phase.

Withdrawal phase. After completing the SSI-3 at T16, the SpeechEasy was removed, and the withdrawal phase began when participants were accompanied to an extraclinical setting to give their first withdrawal phase

speech sample. The two self-report measures (OASES, PSI) were given to the participants to be taken home and completed within the next day. Participants returned twice more to give speech samples (W2 and W4). At the final visit, participants completed a short exit questionnaire detailing general satisfaction with the device.

Speech Tasks

All speech samples collected outside the laboratory were obtained alternately at the University of Colorado student union and the City of Boulder Public Library. These locations were chosen because of their authentic qualities in comparison to laboratory settings and because of their relative quietness. As ambient noise is a common complaint of SpeechEasy users, sometimes resulting in an inability to wear the device in loud environments (*SpeechEasy Professional Information Packet,* 2006), participants were always asked whether the noise level was acceptable before speech samples were taken. This was done to ensure that common SpeechEasy usage patterns were mirrored as best as possible during sampling. If background noise was deemed by the participant to be too bothersome, the sample was collected in a quieter section of the building. This scenario occurred very rarely.

During each speech sample, participants completed three speech tasks (counterbalanced across and within subjects): reading, conversation, and asking a question to a stranger. During reading and conversation, participants sat opposite the experimenter(s) at an available table or at a vacant sitting area. For the reading task, participants read one of three passages from a middle school–level text that were similar in theme and syntactic complexity (Hemingway, 1998). During the conversation task, participants engaged in conversation with one or two experimenters (the first or second author) on preselected topics or on a topic of their choice. For the question task, participants were given one of three predetermined questions to ask to a stranger, or they could ask a question of their choice, provided it was of sufficient length. The planned questions were as follows: "Excuse me, could you please tell me where the nearest restroom is?" (15 syllables); "Could you tell me how to access the Internet in this building?" (16 syllables); and "I was wondering what time you're going to close today?" (14 syllables). These predetermined questions were used roughly 92% of the time. To ensure that the exchange was as realistic as possible, participants were told to ask the question using their own phraseology rather than repeating it verbatim, and the experimenter(s) did not accompany the participants while the question was being asked. The average length of all questions asked was 21.8 syllables (*SD* = 14.6; range = 7–88). Although this is a relatively modest sample size for evaluating speech behavior beyond clinic, we felt that requiring participants

to ask several questions to many different strangers would have been too daunting or, in some cases, an impossible expectation. Some participants had little or no experience with desensitization exercises commonly used to lessen the fear and avoidance of stuttering, making merely one question a sufficiently stressful undertaking.

Stuttering Counts

Stuttering frequency was calculated in percentage of syllables stuttered (%SS) by dividing the stuttering count by the total number of scored syllables. For the reading and conversation tasks, the first 300 syllables were scored. For the question task, all syllables that were part of a question were scored. Statements, asides, and other types of utterances besides questions were not scored.

A *disfluency* was defined as a part-word, whole-word, or phase repetition; a part-word prolongation; an inaudible postural fixation; a starter; or a filler (e.g., "um," "uh," "you know," "like"). A starter or filler was counted as a disfluency only if it was determined that the participant used it habitually to postpone the next word or as a means to say the desired word fluently (Guitar, 2005). A syllable could be counted as stuttered only once. For instance, "*um, um, um,* I want to *g-,* to *g-,* to go" would be counted as two disfluencies.

All speech samples were scored by the first author, a doctoral student in speech-language pathology specializing in fluency disorders. Interrater reliability was established by the second author, a certified speech-language pathologist, rescoring 10% of samples for each speech task chosen at random. The calculated Pearson correlation coefficient was .998, indicating that interjudge agreement was high. The percentage of agreement between raters by speech task ranged from 87% to 94% (reading = 87.0%; conversation = 91.2%; question = 93.5%).

Self-Report Instruments

Two standardized self-report instruments were given during the baseline and withdrawal phases. The OASES (Yaruss & Quesal, 2006) collects information about the total impact of stuttering on one's life, including (a) general perspectives about stuttering; (b) affective, behavioral, and cognitive reactions to stuttering; (c) functional communication difficulties; and (d) impact of stuttering on the speaker's quality of life. One hundred items are organized into four sections entitled General Information, Reactions to Stuttering, Communication in Daily Situations, and Quality of Life. The PSI (Woolf, 1967) inquires about an individual's struggle behaviors, avoidance patterns, and expectancy to stutter. A total of 60 statements are presented, and the person is to check the statement if it is a characteristic of him or her. These two instruments were included to provide a fuller picture

of the overall effects of the SpeechEasy by supplementing more traditional measures of surface speech behaviors. The OASES "collects information about the totality of the stuttering disorder" (Yaruss & Quesal, 2006, p. 90) by assessing the largely unobservable symptoms of stuttering. Reliability and validity of this instrument have been established (Yaruss & Quesal, 2006). Although its psychometric integrity is less well established, the PSI has been used previously in SpeechEasy outcomes research (Stuart et al., 2006) and "can give clinicians and researchers a more complete picture of the speaker's experience of the stuttering disorder" (Yaruss & Quesal, 2006, p. 92). Also, like the OASES, the PSI addresses many hidden components of stuttering, particularly those associated with anxiety.

Participants kept weekly logs while wearing their devices. The logs were e-mailed to the participants at the start of each week. Daily device wear time was recorded (see Table 1), and participants gave a summary evaluation of their speech for that week. The experimenters instructed the participants to comment on any notable experiences with, opinions about, and/or reflections on wearing the SpeechEasy. The exact wording as printed on the logs was as follows: "Summary evaluation of your speech this week. Write as much as you would like and include any comments regarding your experiences with the device this past week."

Upon completion of the study, participants were given an exit questionnaire containing five open-ended questions probing overall satisfaction with the device: "What did you especially like about the device?"; "What did you especially not like about the device?"; "In what particular situation(s) was this device helpful?"; "In what particular situation(s) was this device not helpful?"; and "After this research project is completed, will you continue to wear the device? Why or why not?" The questions were devised based on clinical intuition and recent SpeechEasy literature reporting usage patterns and feedback given by device users (Molt, 2006c; *SpeechEasy Professional Information Packet,* 2006). Participants were encouraged to write as much as they liked in response to each question.

Preparation and Analysis of Qualitative Data

The log and exit questionnaire data were subjected to thematic coding procedures similar to those described in Anderson and Felsenfeld (2003). The first author carefully read through the entire corpus of participant comments and highlighted all information-rich quotes. Quotes were selected as being information-rich if they were judged to pertain in any way to use of the SpeechEasy. As the weekly log instructions and exit questionnaire items referred specifically to the device, nearly all of the quotes

examined were judged to be information-rich. A total of 288 quotes was extracted, given an identifying code to facilitate referencing, and compiled in a database using Microsoft Access. To operationalize the categories, a "bottom-up" procedure was applied wherein thematic categories were developed from the quotes themselves. For example, the categories *Attending to the Second Speech Signal* and *Using Fluency Techniques With Device* were created due to several participants commenting on those topics. Information that was thematically related was grouped together, and quotes could be placed into more than one category (e.g., a quote could refer to both *Increased Confidence* and *Carryover Effects* if the participant reported that his or her confidence remained when the device was removed). The investigators then jointly discussed in detail the preliminary categories to determine whether any needed to be clarified, collapsed to reduce redundancy, or expanded to enhance precision (e.g., *Poor Device Effects* was expanded to *Insufficient Effects to Warrant Buying Device, Variable and/or Diminishing Effects Over Time,* and *Device Is Generally Ineffective*). The only difference in interpretation concerned the categories of *Background Noise* and *Hearing/Understanding Self and/or Others.* After further study of the quotes, agreement was reached that the former category involved specific references to "noise," "loud" environments, and "annoying sounds," whereas the latter involved comments such as "blocks my hearing," "harder to hear others," and "not being able to understand people." Altogether, 20 thematic categories were preserved. One category, *Attending to Second Speech Signal,* was further divided into a subcategory related specifically to "tuning out" the device after continued use. All participants were represented, though not equally, as some were more verbose in their self-reports than others.

Coding reliability was established by the second author and a graduate student in speech-language pathology, who recoded 15% of the quotes. Forty-three quotes were randomly chosen, and the scorers were given the 20 thematic categories and asked to apply the most appropriate categories to the quotes. Because some quotes could be placed into more than one category, there was a total of 68 codes that could be applied to the selected quotes. The scorers were not told how many total codes could be applied to the selected quotes. Interrater agreement between the first and second author was 94.1% (i.e., both raters independently assigned the same category to the same quote on 64 out of 68 opportunities), and agreement between the first author and the graduate student was 95.6%.

Outcome Measures

Primary outcome measures for this study were %SS scores and scores on pre- and posttreatment instruments.

Secondary outcomes included themes derived from weekly log and exit questionnaire comments, as well as participants' ultimate purchasing decision.

Results
Group Data
Quantitative Data

For each participant, stuttering counts were pooled into a single data point representing each phase and speech task. In other words, a single baseline, treatment, and withdrawal score was calculated individually for the three speech tasks. The data from S4 and S6 were not included in the analysis because insufficient data were collected from these participants. A 3 (experimental phase) × 3 (speech task) repeated measures analysis of variance (ANOVA) run on the pooled data revealed a significant main effect of speech task, $F(2, 16) = 4.72$, $p = .024$; no significant main effect of phase, $F(2, 16) = 2.82$, $p = .090$; and no significant interaction effect of Phase × Speech Task, $F(4, 32) = 1.72$, $p = .169$. Since the ANOVA failed to show a significant interaction effect, no post hoc t tests were run on the group data. Figure 1 shows group stuttering frequencies for each experimental phase and speech task.

In order to more directly compare our data with previous reports of stuttering reduction under AAF and while wearing the SpeechEasy, the percentage of fluency improvement (i.e., stuttering reduction) compared with baseline was also calculated from the %SS scores of the pooled group data. Figure 2 shows the results of this analysis, which was carried out merely as a means of

Figure 1. Interval plot of speech task by phase ($n = 9$; 95% CI for the mean). Group data from pooled stuttering counts. Data from S4 and S6 are not included because insufficient data were collected from these participants. CI = confidence interval; %SS = percentage of syllables stuttered; base = baseline; treat = treatment; with = withdrawal.
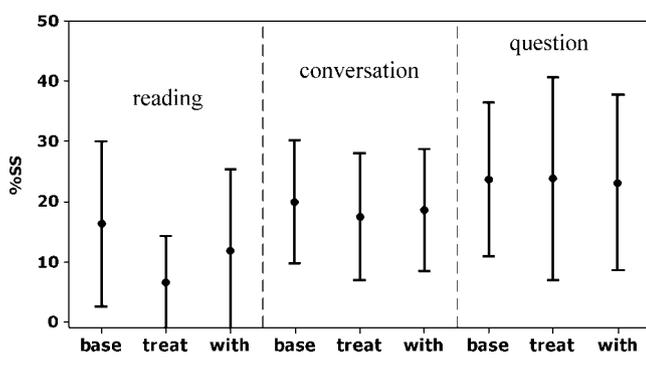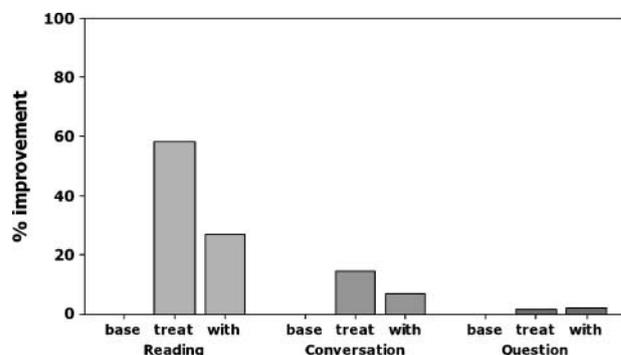

Figure 2. Group of percent improvement compared to baseline ($n = 10$). Data from all participants except S6 were included in treatment phase calculations. Data from all participants except S6 and S4 were included in withdrawal phase calculations.


comparison with past research, as there was not a significant main effect of phase.

For oral reading, there was 58.3% less stuttering during the treatment phase and 27.2% less stuttering during the withdrawal phase. For the conversation task, there was 14.5% less stuttering during the treatment phase and 6.8% less stuttering during withdrawal. For the question task, stuttering decreased 1.9% during the treatment phase and 2.3% during withdrawal. Also, to compare our data with the percentage of overall reduction in stuttered syllables reported by Stuart et al. (2004) through 4 months of continuous SpeechEasy use, we collapsed treatment phase data (T0–T16) across speech task and time and compared it with collapsed baseline fluency levels (B0–B4). We report overall stuttering reductions of 22% with the device in place during speaking in extra-clinical environments.

Immediate effects of the device were also analyzed at the group level. Only data from the final baseline speech tasks (B4) and immediately after device fitting (T0) were included in the analysis (Note: The data from S6 were used because samples were collected from her pre- and post-fitting). A 2 (experimental phase) × 3 (speech task) repeated measures ANOVA revealed a significant main effect of speech task, $F(2, 20) = 5.57$, $p = .012$; a significant main effect of phase, $F(1, 10) = 16.57$, $p = .002$; and a nonsignificant interaction effect of Phase × Speech Task, $F(2, 20) = 1.20$, $p = .323$. Because the ANOVA failed to show a significant interaction effect, no post hoc t tests were run on the immediate pre- and postfitting group data. Figure 3 shows group stuttering frequencies by speech task immediately before and after device fitting.

Group scores on the pre- and posttreatment self-report instruments and the SSI-3 are represented graphically in Figure 4. A series of planned paired-sample t tests were run in order to determine whether any of the

Figure 3. Interval plot of pre/post-fitting effects (*n* = 11; 95% CI for the mean). Immediate device effects. "pre" = samples collected at B4; "post" = samples collected at T0; read = reading; con = conversation; ques = question.

Figure 4. Graph of pre/posttreatment measures (*n* = 9; 95% CI for the mean). All data are raw scores. *p < .05.
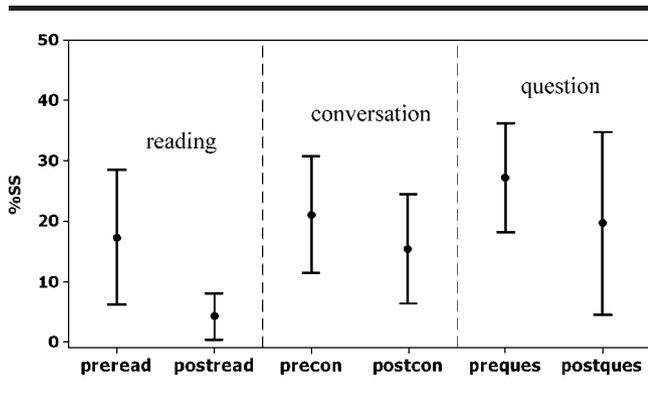
pre- and posttreatment measures showed significant improvement after 4 months of continuous exposure to AAF. For the SSI-3, the score taken at B0 was compared with that taken at T16. Analyses revealed that the posttreatment PSI scores were significantly improved compared with pretreatment, $t(16) = 3.13$, $p = .014$. No other tests reached significance ($p > .05$ for all other comparisons).

Qualitative Data

The thematic coding procedures applied to the comments recorded in the weekly logs and exit questionnaires yielded 20 thematic categories. The final 17 thematic categories on which more than 2 participants commented are presented in Table 2. The conceptual scheme offered by the exit questionnaire items (i.e., "likes" and "dislikes," situations where device was helpful or unhelpful) was used to arrange the categories. Next to each category is the number of different participants who commented on that category.

The most commonly reported "likes" of the device were increased confidence in speaking (*n* = 6) and overall improvement in fluency (*n* = 6). The most commonly reported "dislikes" of the device were irritating background noise (*n* = 8) and being unable to hear/understand one's self and/or others (*n* = 5). The most common situations in which the device was reported to be helpful were using the telephone (*n* = 9) and speaking with strangers (*n* = 5). The most common situations in which the device was reported to be unhelpful were speaking in a restaurant and/or bar (*n* = 6), during physical and/or psychological stress (*n* = 5), and speaking in "noisy" and/or "crowded" settings (*n* = 5).

Of the 11 participants who originally began the study, 4 (36%) elected to purchase their devices at 60% off the retail purchase price of $4,500 after finishing the project. A further 3 out of 11 (27%) chose not to purchase
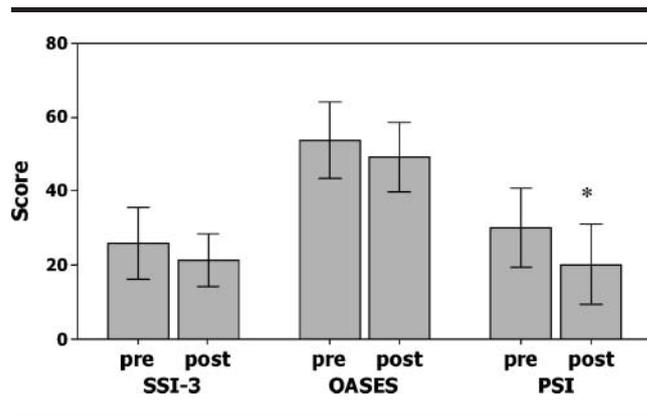
their devices at the substantial discount but reported that they would have continued to use the SpeechEasy if it were given to them free of charge. Three out of 11 (27%) chose not to purchase their devices and also reported that they would not have continued to use the device if it were given to them free of charge. One participant who withdrew from the study immediately after receiving her device (S6) could not be contacted to answer whether she would have continued to use the device if it were given to her at no cost.

Individual Data
Stuttering Counts

Individual responses are represented graphically in Figure 5, showing stuttering frequencies for each participant by time and speech task. For each participant, all speech samples taken postbaseline were compared with the mean for that participant's baseline. In order to demonstrate significant fluency improvement, a speech sample had to score at least two standard deviations below the baseline mean. This criterion was chosen to increase the meaningfulness of the results and because previous SpeechEasy research has reported robust stuttering reductions of 75%–85% while wearing the device (Stuart et al., 2006).

Two data points are missing for the question task of S3 (T4 and T10) because, upon listening to the samples, it was determined that the participant said the question aloud but was not actually speaking to another person. S9 lost her device at Week 15, and her final two treatment phase speech samples were therefore not collected. Similarly, S8 lost her device at Week 6. After receiving a replacement device 4 weeks later, her previous settings were confirmed, and she completed the treatment and withdrawal phases. S4 and S6 have incomplete data because they chose to withdraw from the study. S4 withdrew after wearing the device for approximately 1 month,

Table 2. Conceptual scheme and thematic categories extracted from participant quotes.

Likes	Dislikes	Situations in which device helped	Situations in which device did not help	Other comments
Increased confidence (n = 6)	Background noise (n = 8)	Telephone (n = 9)	Restaurant and bar (n = 6)	Attending to second speech signal (n = 8)
Improved fluency (n = 6)	Hearing/understanding self and/or others (n = 5)	Speaking with strangers (n = 5)	During physical and/or psychological stress (n = 5)	"Tuning out" device after prolonged use (n = 5)
Less severe blocks (n = 5)	Insufficient effects to warrant buying device (n = 4)	Small group (n = 3)	"Noisy" and/or "crowded" setting (n = 5)	Acclimating to device over time (n = 7)
Device facilitated slower speech rate (n = 4)	Variable and/or diminishing effects over time (n = 4)	Job interview (n = 3)	One-on-one conversation (n = 3)	Using fluency techniques with device (prolongation, nonspecific) (n = 7)
Carryover effects (n = 4)	Device is generally ineffective (n = 3)	Quiet setting (n = 3)		Wearing regimen (n = 5)
Became more focused on own speech (n = 3)	Hygienic issues (cerumen, itching, sweat) (n = 3)			Lack of carryover effects (n = 3)
	Cost of device (n = 3)			

Note. Numbers in parentheses are the number of different participants who commented on that category.

citing intolerable background noise as the principal reason. S6 did not show up for any meetings after she received her device and would not respond to subsequent attempts to contact her. Consequently, it could not be determined why she withdrew from the study. Her individual data are not reported here because only one sample was collected during the treatment phase.

Pre- and Posttreatment Measures

Table 3 displays pre- and posttreatment scores on the two self-report instruments (OASES, PSI) for the participants who completed the study (n = 9). Also shown are scores on the three administrations of the SSI-3 at B0 (first visit), T0 (postfitting), and T16 (end of treatment phase).

Weekly Logs

The average time each participant wore his or her device throughout the treatment phase was as follows: S1 = 4.4 hr/day; S2 = 10.4 hr/day; S3 = 5.1 hr/day; S4 = 2.3 hr/day; S5 = 6.3 hr/day; S7 = 4.0 hr/day; S8 = 6.4 hr/day; S9 = 6.4 hr/day; S10 = 4.0 hr/day; S11 = 1.1 hr/day (also refer to Table 1). The average wear time for the group (n = 10) was 5.0 hr/day.

Discussion
Overall Results

The results of this study represent data from the first Phase I clinical trail of the SpeechEasy under challenging, relatively naturalistic conditions. Group stuttering reduction data suggest that, through 4 months of use, our sample did not benefit significantly from the device as compared with baseline. Inspection of individual responses indicate that prolonged SpeechEasy use produced variable levels of fluency enhancement and that the effects varied by speech task as well. Moreover, participants were also widely diverse in their subjective responses to the device. In their own words, some participants felt that "Using the SpeechEasy device is wonderful for me. It's like day and night. I don't have to think about how I am talking all the time" (S9); whereas others reported that "It did not significantly improve my fluency and was unhelpful in conversational speech" (S4). Responses like these represent two ends of a continuum along which the opinions of the participants were spread.

Group analyses indicated that the device was most beneficial while participants read a passage aloud (58% less stuttering), was less so during conversation (15% less stuttering), and performed poorest during the often challenging task of asking a question to a stranger (2% less stuttering). It is important to note, however, that no statistically significant treatment effect was found for any of the three speech tasks used in this study. The relatively weaker effects of the SpeechEasy during formulated speech are noteworthy, given Armson and Stuart's (1998) statement that "a prerequisite to successful use of [AAF] in situations of daily living is that it reduce stuttering during self-formulated speech" (p. 487). For our sample of PWS, that requirement was not met over the long term, as neither the conversation nor question

Figure 5. Stuttering frequencies for each participant by time and speech task. Dotted line = baseline mean; solid line underneath = 2 standard deviations below baseline mean; vertical lines delineate experimental phases (baseline, treatment, withdrawal).
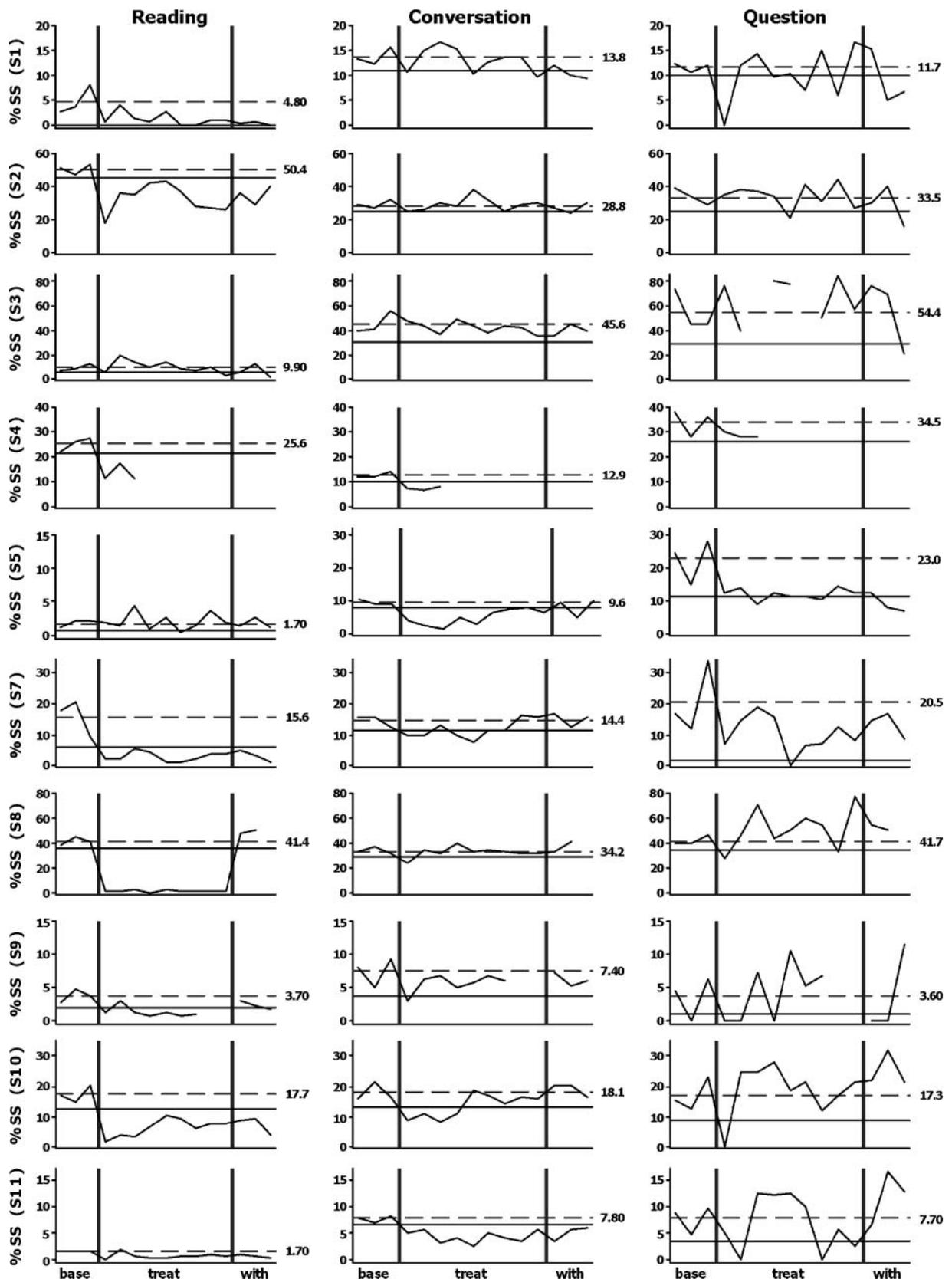
Table 3. Scores on pre- and/posttreatment instruments.

Subject	SSI-3 (B0)	SSI-3 (T0)	SSI-3 (T16)	OASES (pre)	OASES (post)	PSI (pre)	PSI (post)
S1	Mild	Very mild	Mild	Mod to sev	Mild to mod	24/60	15/60
S2	Very severe	Very severe	Very severe	Mod to sev	Mod to sev	49/60	43/60
S3	Very severe	Sev to very sev	Mod	Mod	Mild to mod	44/60	14/60
S5	Very mild to mild	Very mild	Very mild	Mod	Mod	40/60	40/60
S7	Mod	Very mild to mild	Very mild	Mild to mod	Mild to mod	11/60	6/60
S8	Very severe	Mod	Mod	Mod to sev	Mod to sev	38/60	31/60
S9	Very mild	Very mild	Very mild	Mod	Mod	29/60	10/60
S10	Mild to mod	Very mild	Mild	Mod	Mod	26/60	14/60
S11	Very mild	Very mild	Very mild	Mild	Mild to mod	10/60	9/60

Note. All data are raw scores. PSI = Perceptions of Stuttering Inventory; sev = severe.

tasks showed statistically significant fluency improvement over 4 months' time. To directly compare our data with the 81% overall reduction in stuttered syllables reported by Stuart at el. (2004) through 4 months of continuous SpeechEasy use, we collapsed treatment phase data and compared it with collapsed baseline fluency levels. We report a more modest overall stuttering reduction of 22% with the device in place; but, again, this reduction in stuttering was not significantly different from baseline due to high within-group variability.

There were two areas in which statistically significant group effects were found. First, similar to recent findings (e.g., Armson & Kiefte, 2008; Armson et al., 2006; Stuart et al., 2004, 2006), we report a significant immediate effect of the device across speech tasks. However, our immediate effects of 75% and 27% reductions in stuttered syllables for reading and conversation, respectively, are less robust than the 90% and 67% stuttering reductions reported by Stuart et al. (2004) for reading and monologue. Secondly, group effects were seen in pre- and posttreatment PSI scores. The PSI was designed to inventory the maladaptive, compensatory behaviors PWS often use when speaking or anticipating speaking. Our findings of significantly improved scores on this instrument at posttreatment suggest that secondary characteristics such as word or situational avoidance, extraneous movements, and/or anticipatory behaviors or feelings may be reduced with prolonged exposure to the SpeechEasy. It is worth noting here that many PSI items refer to stuttering phenomena assumed to be associated with anxiety. It is therefore possible that improved PSI scores at posttreatment may have reflected general anxiolytic effects of the SpeechEasy that lessened the negative experience of stuttering, while failing to reduce the frequency of stuttered syllables under challenging conditions. A similar finding of improved PSI scores compared with baseline has been reported at 12 months postfitting (Stuart et al., 2006). The baseline in that study, however,

was established by asking participants at 12 months postfitting to recall as best they could how they felt and spoke before receiving their devices. This retrospective method raises serious methodological issues of recall, reliability, and bias that preclude consideration of those findings. Therefore, further evidence is needed to establish the reliability of our results regarding improved PSI scores following continuous SpeechEasy use.

Individual Responses

Although group results failed to show a significant effect of the SpeechEasy over time compared with baseline, between-subjects variability was conspicuous. The relatively large standard deviations of %SS data across speech tasks indicated that certain participants responded more favorably to the SpeechEasy than others. Several of our participants showed significant fluency enhancement across one or more speech tasks for all or most of the treatment phase. S5 and S11, for instance, maintained significantly improved fluency compared with baseline throughout the treatment phase for the conversation task. Similarly, half of our participants (S2, S7, S8, S9, S10) maintained improvement during oral reading. Although no individual showed sustained improvement during the question task, 2 participants (S1, S9) were significantly more fluent for nearly half of their speech samples while asking questions to strangers.

The collective comments from the weekly logs and exit questionnaire suggest that a promising feature of the device is its effectiveness while using the telephone. Although we did not directly test the SpeechEasy under this condition, 90% of our participants mentioned the telephone as a speaking situation in which they felt their fluency was improved while wearing the device. Several participants reported that this was the activity of daily living during which they believed the SpeechEasy's effects were most noticeable. Similar subjective impressions

of the device's effectiveness while on the telephone have also been reported by Molt (personal communication, November 17, 2006). This is of no small consequence, as PWS often report telephone use as one of the most stressful and feared speaking situations (Georgieva, 1994; Silverman, 1997). Our results thus provide support for the findings reported by Zimmerman et al. (1997) of significant stuttering reduction under AAF during scripted telephone calls.

Generally, SpeechEasy use was associated more so with greater subjective satisfaction and improved scores on posttreatment instruments than with objective fluency enhancement. This trend was evidenced by several participants. For example, although %SS data indicated that S3 received no benefit from wearing the device, the majority of his log comments were positive, and all of his scores on the posttreatment instruments were dramatically improved. S9's PSI score improved from 29/60 to 10/60, and most of her log comments were positive, despite only 29% of her formulated speech samples meeting the improvement criterion. S8 also displayed this trend, with almost no increased fluency during formulated speech but substantially improved posttreatment scores on the SSI-3 and PSI. It is interesting to note that the opposite case was also seen within our sample. Two participants (S4, S11) demonstrated measurable fluency enhancement while wearing the SpeechEasy during conversation and reading but nevertheless reported very negative opinions of the device and had the lowest device wear times of the group (2.3 hr/day and 1.1 hr/day, respectively).

This dissociation between overt fluency scores and more qualitative responses to the device is something that is beginning to emerge as an intriguing phenomenon in SpeechEasy research. Reporting interim data on a relatively large sample of 20 participants, Molt (2006a) similarly found occasional incongruity between attitudinal measures and objective fluency scores. Relatedly, Runyan, Runyan, and Hibbard (2006) found that although 5 out of 9 participants no longer wore their devices at 2–3 years follow-up, all of them reported that they would choose to purchase the SpeechEasy again. Perhaps germane to this issue is evidence showing that device use may increase perceived confidence for many individuals, often coinciding with decreases in reported anxiety and avoidances (Cook & Smith, 2006; Molt, 2006a; Runyan et al., 2006). Many of our participants' log and exit questionnaire comments suggest this to be the case, as 60% of our sample mentioned increased confidence in speaking as a benefit of device use. Although the extant evidence is not definitive, it appears that this perceived boost in confidence may, in part, account for the curious discrepancy sometimes seen between still fairly high stuttering levels and a favorable opinion of the device.

Purchasing Decision and Response Profiles

Statistics related to treatment effects and reports of subjective responses are invaluable to any experimental assessment of treatment efficacy. However, in the final analysis, the individual's decision whether to purchase the product under scrutiny is perhaps the most telling piece of evidence. In total, 4 of the original 11 participants (36%) who began the study elected to purchase their devices at a substantial discount (Group 1). Three participants (27%) did not purchase the device but indicated that they would continue to use the SpeechEasy if it were given to them free of charge (Group 2). Last, 3 participants (27%) did not buy the device and reported that they would not continue to use it even if it were given to them at no charge (Group 3). This grouping of participants by ultimate outcome (i.e., purchasing decision) approximately agrees with anecdotal reports and our own clinical experience suggesting that roughly one third of those who try the SpeechEasy experience substantial benefit, one third experience some degree of fluency improvement, and one third are helped very little or not at all (Ramig et al., in press). Although this coincidence is admittedly far from reliable, it, coupled with the data reported herein, does raise questions about the "80%–90% success rate" reported in the national media (e.g., Sloan, 2007) and the 80% satisfaction rating reported for the device by Rainmaker & Sun Integrated Marketing (in *SpeechEasy Professional Information Packet,* 2006). It is worth noting, however, that Rainmaker and Sun surveyed individuals who had sought out and purchased a SpeechEasy, whereas our sample consisted of individuals who were recruited to try the device at no cost and thus could conceivably have been less motivated and emotionally invested in a positive outcome.

Another interesting finding pertains to the composition of the three groups. Although all the participants in Group 1 purchased their devices, none of them displayed sustained fluency enhancement across speech tasks while wearing the SpeechEasy. The same can be said of Group 2 and Group 3. Additionally, the response profiles of all three groups were quite diverse. In other words, none of the three groups possessed within-group homogeneity, and those who chose to purchase their devices had response profiles (i.e., %SS scores and scores on posttreatment measures) that were essentially indistinguishable from those who did not. This indicates that there may have been unaccounted-for variables that could have determined participants' ultimate opinions of the device. The following summative statement from S7 may help elucidate some of the factors left undetected by common measures, such as self-report instruments and stuttering counts: "I'm starting to realize that I'm not going to buy the device…. I gain in my speech improvement,

but I lose in being able to understand other people or it just being really annoying to hear an echo of everything in my world." As this comment demonstrates, an individual's complete judgment of the SpeechEasy can be multifaceted and even conflictive. Further research is therefore needed to determine which factor or constellation of factors bears most strongly on purchasing decision.

A fairly clear response profile that some participants displayed was evident in the %SS data for S5, S7, and S10. Each of these participants showed initial increases in fluency during conversation before gradually relapsing back to baseline fluency levels by roughly the third or fourth month of device use. The declension of AAF effects over time is an issue that has been raised previously by stuttering researchers (e.g., Ingham et al., 1997), and this pattern of initial improvement with the SpeechEasy followed by gradual relapse has also been reported by Molt (2006b). He found that improvements in fluency scores and on qualitative measures held for the majority of participants after 3 months of SpeechEasy use but were less pronounced at 6 and 12 months postfitting. These data conflict with previous research reporting that the reduction in disfluencies experienced at initial fitting is maintained at 4 and 12 months follow-up in the absence of intervening treatment (Stuart et al., 2006).

Limitations and Strengths of Our Study

A limitation of this study was that speech samples were collected over several months by the same experimenter(s), rather than by different experimenters. During stuttering treatment, it is not unusual for symptoms to decrease as a client becomes more comfortable with his or her clinician, irrespective of the actual effects of the therapy (Conture & Curlee, 2007; Guitar, 2005). Although the experimenters did not technically fill the traditional role of therapist during the study, they were in a similar authoritative position of trust and developed a rapport with the participants due to frequent contact over the 6-month period during which data were collected. This could potentially have confounded the true effects of the SpeechEasy and likely contributed to the difficulty in interpreting the data from many participants' withdrawal phases. Although group analyses failed to show an overall carryover effect for the device, several individuals appeared to retain significantly increased fluency posttreatment for certain speech tasks. Relatedly, some participants reported in their logs increased confidence or increased stuttering upon removal of the device during the week, neither of which could be confirmed empirically. This issue is significant because a previous report suggested that continued exposure to AAF produces carryover fluency in the absence of the stimulus (Van Borsel et al., 2003), whereas Stuart et al. (2004) offered no evidence of a carryover effect. S1 (conversation),

S2 (reading), S5 (question), S7 (reading), S10 (reading), and S11 (conversation) all showed continued fluency enhancement for at least one speech task after the device was removed. What is unclear, however, is the proportion of maintained fluency enhancement that was due to a real carryover effect versus the proportion due to participants becoming more comfortable with the researchers, the speech tasks, and/or the environments over the course of the study. Further research better controlling for such confounds is needed to more clearly disentangle device from familiarity effects.

A constraint that may arguably hinder the generalizability of our findings was our small sample size. However, the sample size was comparable to other work investigating the SpeechEasy (Armson et al., 2006; Stuart et al., 2004, 2006), and importantly, it was appropriate and ethical for a Phase I trial designed to identify whether sufficient treatment effects exist that indicate the need for further clinical trials (Robey & Schultz, 1998).

Two distinguishing features of this study were our inclusion criteria and our sampling methods. Regarding the former, our criteria were somewhat less restrictive than those commonly employed in other AAF research (e.g., Kalinowski et al., 1993; Molt, 2006a, 2006b; O'Donnell, Armson, & Kiefte, 2008). Participants mainly had to have a subjective impression during initial fitting that they received enough benefit from the SpeechEasy to want to wear it continually for 4 months. Past AAF research has typically included requirements that prospective participants meet a certain level of stuttering reduction under AAF or display minimum levels of disfluency while reading as a prerequisite for inclusion. Though likely making for a more homogenous sample, this practice may result in decreased external validity by preemptively eliminating a portion of all those who stutter. By contrast, our sample was likely composed of a broader sample of all PWS, as it potentially included some who would have ordinarily been unlikely to purchase the SpeechEasy and/or were less intrinsically motivated than typical SpeechEasy buyers.

Our sampling methods that differed from past AAF research were the collection of speech samples outside the laboratory in naturalistic environments, the inclusion of a conversation and question task, and our collection of multiple speech samples. As Braun et al. (1997) stated:

> One of the most consistent observations in the evaluation of individuals who stutter has been that situation- or task-specific variations in symptom intensity represent a salient feature of the disorder. Typically, stuttering occurs during spontaneous interpersonal communication and may be exacerbated by stress. (p. 762)

Indeed, the nature of stuttering is such that the overt stuttered speech patterns as well as the covert struggles

of the individual often fluctuate dramatically in relation to speaking environment, listener identity, speaking task, speech content, and levels of internal and external stress (Guitar, 2005; Van Riper, 1982). The context-dependent nature of the disorder was therefore the kernel around which our research design was fashioned and what led us to sample outside the laboratory and under challenging conditions. Additionally, we believe that collecting repeated measures frequently over several months, rather than single measures at lengthy intervals, provides data that are probably more indicative of participants' habitual stuttering behaviors. As the frequency of stuttering symptoms can vary across time, this method of sampling may be better suited to capture the fluctuations in severity intrinsic to the disorder, particularly during a longitudinal study such as this. Previous longitudinal AAF research, by contrast, sampled at intervals of 3 or 4 months to 8 months (e.g., Stuart et al., 2004, 2006; Van Borsel et al., 2003).

We feel that the combination of less restrictive inclusion criteria and repeated, naturalistic sampling procedures increases the external validity of our results. By testing the effects of the SpeechEasy under challenging conditions with a broad range of PWS, we have sought to address some of the limitations of past AAF research carried out within quiet laboratories and clinic rooms using a restricted variety of speech tasks. We concur with Lincoln et al. (2006) that "The true test of the effectiveness of AAF as a treatment for stuttering would involve measuring speech in everyday speaking situations such as workplaces, restaurants, homes and social settings in general" (p. 79). All of our samples were collected in naturalistic settings. Also, two of our speech tasks were situations that PWS are likely to encounter frequently in everyday life—conversing with an acquaintance and asking a question to a stranger. To our knowledge, this is the first experimental evidence of the effects of AAF in naturalistic settings during those two challenging and clinically relevant tasks. It is encouraging to note, however, that other research groups are beginning to examine AAF effects under naturalistic conditions as well. A recent multiple single-subject study (O'Donnell et al., 2008) similarly describes SpeechEasy effects during situations of daily living. Although they did not perform statistical analyses of group data, their findings are comparable to ours in that they also report variable between-subjects effects and changeable response patterns over time, including an adaptation effect for several participants after prolonged use.

Conclusions and Future Directions

Our group findings showing no treatment effect for the device suggest that Phase II trials of the SpeechEasy are not warranted; however, given the parameters of participant recruitment and methodological protocols employed in the study, this is a qualified conclusion. As some participants benefited clinically on certain speech tasks and/or reported subjective satisfaction with the device, clinicians may wish to probe for device effects with clients on an individual basis. Importantly, our group results conflict with previous reports of dramatic fluency induction under AAF conditions, both immediately upon application and over an extended period of time (Kalinowski et al., 1996; Stuart et al., 1997, 2004). Our results are instead more aligned with recent work reporting more equivocal findings than those obtained by Kalinowski and colleagues (Armson & Kiefte, 2008; Armson et al., 2006; Cook & Smith, 2006; Molt, 2006a, 2006b, 2006c; O'Donnell et al., 2008; Runyan et al., 2006). Our results, though based on a relatively small sample, also support Bothe et al.'s (2007) interpretation of recent SpeechEasy research: "initial promising results from very brief studies conducted by one research group, followed by growing evidence that outcomes achieved using longer trials of spontaneous or conversational speech, conducted by other researchers, are much less positive than the original developers maintained" (p. 79). Additionally, our findings support previous studies indicating that the effects of AAF are more robust during oral reading than during formulated speech (Armson et al., 2006; Armson & Stuart, 1998; Ingham et al., 1997).

Our finding that prolonged SpeechEasy use coincided more so with positive subjective impressions than with measurable reductions in stuttering ought to be viewed in the larger context of evidence-based practice (American Speech-Language-Hearing Association [ASHA], 2004). Given ASHA's commitment to embracing this medical model of endorsing only research-based interventions, how should our results be interpreted by practicing clinicians and their clients? This is difficult to answer, as each clinician's notion of what constitutes an "efficacious" stuttering treatment is probably as diverse as our participants' opinions on the subject. To illustrate, 1 participant (S3) showed almost no reduction in overt disfluencies during the treatment phase yet attributed the device to a remarkable decrease in fear, anxiety, and avoidance behaviors that arguably represent the most handicapping features of the disorder (Guitar, 2005; Ramig & Dodge, 2005; Van Riper, 1982). Conversely, another participant (S4) demonstrated measurable fluency enhancement while wearing the SpeechEasy yet returned the device after 1 month because she felt the drawbacks were "rather annoying." Both responses are equally valid, both demonstrate the importance of client perception in outcomes research, and both highlight precisely why "consumers' feedback about the perceived effect of devices is of critical importance" (Lincoln et al., 2006, p. 78). Yet the issue remains: How much consideration

should clinicians give to perceived benefits in the absence of appreciable improvements in fluency, or vice versa? We cannot say for sure. Suffice it to say that our findings highlight the complexities of integrating client–patient values with clinical expertise and current best evidence while attempting to give equal weight to each element of evidence-based practice (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000).

Regarding reductions in overt stuttering behaviors, our failure to find positive Phase I treatment outcome evidence for the SpeechEasy in naturalistic settings is perhaps compelling. It suggests that the "passive inhibition of stuttering" and "sense of invulnerability to stuttering" (Saltuklaroglu et al., 2003, p. 121) presumably engendered by AAF did not hold up under challenging conditions for our group. Further, our results do not support the assertion that

> [by] simply placing a device in a person's ear; offering good therapeutic instructions for a brief time; and keeping in contact with that person to maintain care of the device, answer questions, and give further instructions, one was able to provide efficient and effective reduction in stuttering at initial fitting and 4 months post-fitting. (Stuart et al., 2004, p. 110)

Rather, our findings suggest that substantial training in active fluency techniques, such as initial sound prolongation and easy onset of voicing, may be necessary to optimize performance of the SpeechEasy. A caveat to the possible necessity of using active fluency techniques to enhance device performance is that if such techniques are needed by some or many SpeechEasy users, might those techniques in and of themselves be sufficient to increase fluency independent of the device? Currently no research exists that addresses this question or related questions pertaining to amount or type of stuttering treatment needed to best augment device performance. We hope that future inquiries will provide insight into those important issues.

Another potential area for future research relates to the perceived drawback of background noise reported here and elsewhere (*SpeechEasy Professional Information Packet,* 2006). The manufacturers of the SpeechEasy have recently developed a modified device specifically intended to address this issue. Further study is needed to assess its effectiveness in attenuating background noise within naturalistic speaking tasks and environments.

Finally, this project was intended to assess the effectiveness of AAF as delivered by the SpeechEasy. Because of this aim, no individualized stuttering treatment was provided beyond periodically reviewing SpeechEasy protocols. These protocols included the use of prolongations on initial sounds and instructions to attend to the second speech signal, for example. For ideal results, the manufacturer recommends that traditional stuttering treatment be implemented in conjunction with device use (*SpeechEasy Professional Information Packet,* 2006). Although we cannot conclude from our findings that device use coupled with stuttering treatment provided by a certified clinician would have produced greater effects, recent research indicates that the manufacturer's suggestion of pairing the two ought to be heeded if and when circumstances permit (Armson et al., 2006). It would be worthwhile to compare the outcomes of such an approach with those that we report here.

References

Adamczyk, B. (1959). Use of instruments for the production of artificial feedback in the treatment of stuttering. *Folia Phoniatrica, 11,* 216–218.

American National Standards Institute. (1996). *Specifications for audiometers* (ANSI S3.6-1996). New York: Author.

American Speech-Language-Hearing Association. (2004). *Evidence-based practice in communication disorders* [Technical report]. Available from www.asha.org/policy.

Anderson, T. K., & Felsenfeld, S. (2003). A thematic analysis of late recovery from stuttering. *American Journal of Speech-Language Pathology, 12,* 243–253.

Armson, J., Foote, S., Witt, C., Kalinowski, J., & Stuart, A. (1997). Effect of frequency altered feedback and audience size on stuttering. *European Journal of Disorders of Communication, 32,* 359–366.

Armson, J., & Kiefte, M. (2008). The effect of SpeechEasy on stuttering frequency, speech rate, and speech naturalness. *Journal of Fluency Disorders, 33,* 120–134.

Armson, J., Kiefte, M., Mason, J., & De Croos, D. (2006). The effect of SpeechEasy on stuttering frequency in laboratory conditions. *Journal of Fluency Disorders, 31,* 137–152.

Armson, J., & Stuart, A. (1998). Effect of extended exposure to frequency altered feedback on stuttering during reading and monologue. *Journal of Speech, Language, and Hearing Research, 41,* 479–490.

Bothe, A., Finn, P., & Bramlett, R. (2007). Pseudoscience and the SpeechEasy: Reply to Kalinowski, Saltuklaroglu, Stuart, and Guntupalli (2007). *American Journal of Speech-Language Pathology, 16,* 77–83.

Braun, A. R., Varga, M., Stager, S., Schulz, G., Selbie, S., Maisog, J. M., et al. (1997). Altered patterns of cerebral activity during speech and language production in developmental

stuttering: An H215O positron emission omography study. *Brain, 120,* 761–784.

Chase, R. A., Sutton, S., First, D., & Zubin, J. (1961). A developmental study of changes in behavior under delayed auditory feedback. *Journal of Genetics and Psychology, 49,* 101–112.

Conture, E., & Curlee, R. F. (Eds.). (2007). *Stuttering and related disorders of fluency* (3rd ed.). New York: Thieme Medical Publishers.

Cook, M. J., & Smith, L. M. (2006, November). *Outcomes for adult males using the SpeechEasy fluency device for one year.* Poster presented at the annual convention of the American Speech-Language-Hearing Association, Miami, FL.

Derazne, J. (1966). Speech pathology in the U.S.S.R. In R. W. Rieber & R. S. Brubaker (Eds.), *Speech pathology* (pp. 140–156). Philadelphia: Lippincott.

Dewar, A., Dewar, A. D., Austin, W. S., & Brash, H. M. (1979). The long term use of an automatically triggered auditory feedback masking device in the treatment of stuttering. *British Journal of Disorders of Communication, 14,* 219–229.

Finn, P., Bothe, A. E., & Bramlett, R. (2005). Science and pseudoscience in communication disorders: Criteria and applications. *American Journal of Speech-Language Pathology, 14,* 172–186.

Georgieva, D. (1994). Speech situations increasing stuttering by 13–16 year old persons. In C. W. Starkweather & H. F. M. Peters (Eds.), *Stuttering: Proceedings from the First World Congress on Fluency Disorders* (pp. 259–263). Nijmegen, the Netherlands: University Press Nijmegen.

Goldiamond, I. (1965). Stuttering and fluency as manipulatable operant response classes. In L. Krasner & L. Ullman (Eds.), *Research in behavior modification* (pp. 106–156). New York: Holt, Rinehart and Winston.

Guitar, B. (2005). *Stuttering: An integrated approach to its nature and treatment* (3rd ed.). Baltimore: Williams & Wilkins.

Hargrave, S., Kalinowski, J., Stuart, A., Armson, J., & Jones, K. (1994). Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rates. *Journal of Speech and Hearing Research, 37,* 1313–1319.

Hemingway, E. (1998). *The complete short stories of Ernest Hemingway.* New York: Simon & Schuster.

Hudson, D. A. (Executive Producer). (2003, February 3). *The Oprah Winfrey show* [Television broadcast]. Chicago: Harpo Productions.

Ingham, R. J., Moglia, R. A., Frank, P., Costello-Ingham, J. C., & Cordes, A. K. (1997). Experimental investigation of the effects of frequency-altered auditory feedback on the speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 40,* 361–372.

Johnson, W., & Rosen, L. (1937). Studies in the psychology of stuttering: VII. Effects of certain changes in speech pattern upon the frequency of stuttering. *Journal of Speech Disorders, 2,* 105–109.

Kalinowski, J., Armson, J., Roland-Mieszkowski, M., Stuart, A., & Gracco, V. L. (1993). Effects of alterations in auditory feedback and speech rate on stuttering frequency. *Language and Speech, 36,* 1–16.

Kalinowski, J., Stuart, A., Sark, S., & Armson, J. (1996). Stuttering amelioration at various auditory delays and speech rates. *European Journal of Disorders of Communication, 31,* 259–269.

Kalinowski, J., Stuart, A., Wamsley, L., & Rastatter, M. P. (1999). Effect of monitoring condition and altered auditory feedback on stuttering frequency. *Journal of Speech, Language, and Hearing Research, 42,* 1347–1354.

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders, 31,* 71–89.

Macleod, J., Kalinowski, J., Stuart, A., & Armson, J. (1995). Effect of single and combined altered auditory feedback on stuttering frequency at two speech rates. *Journal of Communication Disorders, 28,* 217–228.

Molt, L. (2005). *A brief historical review of assistive devices for treating stuttering.* Paper presented at the Eighth International Stuttering Awareness Day Online Conference (ISAD8). Retrieved September 2, 2007, from http://www.mnsu.edu/comdis/isad8/isadcon8.html.

Molt, L. (2006a, November). *SpeechEasy AAF device long-term clinical trial: Attitudinal/perceptual measures.* Poster presented at the annual convention of the American Speech-Language-Hearing Association, Miami, FL.

Molt, L. (2006b, November). *SpeechEasy AAF device long-term clinical trial: Speech fluency and naturalness measures.* Poster presented at the annual convention of the American Speech-Language-Hearing Association, Miami, FL.

Molt, L. (2006c, November). *SpeechEasy AAF device long-term clinical trial: Usage patterns and satisfaction ratings.* Poster presented at the annual convention of the American Speech-Language-Hearing Association, Miami, FL.

O'Donnell, J. J., Armson, J., & Kiefte, M. (2008). The effectiveness of SpeechEasy during situations of daily living. *Journal of Fluency Disorders, 33,* 99–119.

Ramig, P. R., & Dodge, D. (2005). *The child and adolescent stuttering treatment and activity resource guide.* Clifton Park, NY: Thomson Delmar Learning.

Ramig, P. R., Ellis, J. B., Pollard, R., & Finan, D. (in press). Application of the SpeechEasy to stuttering treatment: Introduction, background, and preliminary observations. In B. Guitar & R. McCauley (Eds.), *Treatment of stuttering: Conventional and emerging interventions.* Baltimore: Lippincott, Williams & Wilkins.

Riley, G. D. (1994). *Stuttering Severity Instrument for Children and Adults* (3rd ed.). Austin: Pro-Ed.

Robey, R. R. (2004). A five-phase model for clinical-outcome research. *Journal of Communication Disorders, 37,* 401–411.

Robey, R. R., & Schultz, M. C. (1998). A model for conducting clinical-outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology, 12,* 787–810.

Runyan, C. M., Runyan, S. E., & Hibbard, S. (2006, November). *The Speech Easy* [sic] *device: A three year study.* Poster presented at the annual convention of the American Speech-Language-Hearing Association, Miami, FL.

Ryan, B. P., & Van Kirk, B. (1971). *Programmed conditioning for fluency: Program book.* Monterey, CA: Behavioral Sciences Institute.

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-based medicine:*

How to practice and teach EBM. Edinburgh, Scotland: Churchill Livingstone.

Saltuklaroglu, T., Kalinowski, J., Dayalu, V. N., Guntupalli, V. K., Stuart, A., & Rastatter, M. P. (2003). A temporal window for the central inhibition of stuttering via exogenous speech signals in adults. *Neuroscience Letters, 349,* 120–124.

Silverman, F. H. (1997). Telecommunication relay services: An option for stutterers. *Journal of Fluency Disorders, 22,* 63–64.

Sloan, D. (Executive Producer). (2007, January 10). *ABC primetime: Medical mysteries* [Television broadcast]. New York: ABC News.

Soderberg, G. A. (1969). Delayed auditory feedback and the speech of stutterers. *Journal of Speech and Hearing Disorders, 34,* 20–29.

SpeechEasy professional information packet. (2006). Available from Janus Development Group, Inc., 112 Staton Road, Greenville, NC 27834. Phone: 866-551-9042. E-mail: customerserv@janusdevelopment.com.

SpeechEasy training manual. (2006). Available from Janus Development Group, Inc., 112 Staton Road, Greenville, NC 27834. Phone: 866-551-9042. E-mail: customerserv@ janusdevelopment.com.

Stromsa, C. (1958). *The effects of altering the fundamental frequency of masking on the speech performance of stutterers* [Technical report Project B-1331]. Rockville, MD: National Institutes of Health.

Stuart, A., Kalinowski, J., & Rastatter, M. (1997). Effects of monaural and binaural altered auditory feedback on stuttering frequency. *The Journal of the Acoustical Society of America, 101,* 3806–3809.

Stuart, A., Kalinowski, J., Rastatter, M., & Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *The Journal of the Acoustical Society of America, 111,* 2237–2241.

Stuart, A., Kalinowski, J., Rastatter, M., Saltuklaroglu, T., & Dayalu, V. (2004). Investigations of the impact of altered auditory feedback in-the-ear devices on the speech of people who stutter: Initial fitting and 4-month follow-up. *International Journal of Language and Communication Disorders, 39,* 93–113.

Stuart, A., Kalinowski, J., Saltuklaroglu, T., & Guntupalli, V. K. (2006). Investigations of the impact of altered auditory feedback in-the-ear devices on the speech of people who stutter: One-year follow-up. *Disability and Rehabilitation, 28,* 1–9.

Stuart, A., Xia, S., Jiang, Y., Jiang, T., Kalinowski, J., & Rastatter, M. P. (2003). Self-contained in-the-ear device to deliver altered auditory feedback: Applications for stuttering. *Annals of Biomedical Engineering, 31,* 233–237.

Sutton, S., & Chase, R. A. (1961). White noise and stuttering. *Journal of Speech and Hearing Research, 4,* 72.

Van Borsel, J., Reunes, G., & Van den Bergh, N. (2003). Delayed auditory feedback in the treatment of stuttering: Clients as consumers. *International Journal of Language and Communication Disorders, 38,* 119–129.

Van Riper, C. (1982). *The nature of stuttering* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Woolf, G. (1967). The assessment of stuttering as struggle, avoidance, and expectancy. *The British Journal of Disorders of Communication, 2,* 158–171.

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.

Zerneri, L. (1966). Attempts to use delayed speech feedback in stuttering therapy. *Journal Francais d'Oto-Rhino-Laryngologie et Chirurgie Maxillo-Faciale, 15,* 415–418.

Zimmerman, S., Kalinowski, J., Stuart, A., & Rastatter, M. (1997). Effect of altered auditory feedback on people who stutter during scripted telephone conversations. *Journal of Speech, Language, and Hearing Research, 40,* 1130–1134.