Research Article

Two to Tango or the More the Merrier? A Randomized Controlled Trial of the Effects of Group Size in Aphasia Conversation Treatment on Standardized Tests

Gayle DeDe,^a Elizabeth Hoover,^b and Edwin Maas^a

Purpose: Group conversation treatment has the potential to improve communication and reduce social isolation for people with aphasia. This project examined how 2 conflicting hypotheses—treatment dosage and group dynamics—affect treatment outcomes.

Method: Forty-eight participants with chronic aphasia were randomly assigned to either a dyad, a large group, or a delayed control group. Conversation group treatment was provided for an hour, twice per week, for 10 weeks. Individual goals were developed by each participant and addressed in the context of thematically oriented conversation treatment. Standardized testing across language domains was completed pretreatment (Time 1), posttreatment (Time 2), at a 6-week maintenance point (Time 3), and at 11-month follow-up for the experimental groups.

Results: Treatment groups showed greater changes on standardized measures than the control group posttreatment.

Dyads showed the most changes on measures of language impairment, whereas changes on the self-reported functional communication measure (Aphasia Communication Outcome Measure) and connected speech task only showed significant changes in the large group.

Conclusions: This randomized controlled trial on conversation treatment indicated that both treatment groups—but not the delayed control group—showed significant changes on standardized tests. Hence, conversation treatment is associated with changes in measures of language impairment and quality of life. Dyads showed the most changes on measures of language impairment, whereas changes on the functional communication measure (Aphasia Communication Outcome Measure) and discourse production only showed significant changes in the large group. Thus, group size may be associated with effects on different types of outcome measures.

Aphasia is the loss or reduction of language after brain damage. It is one of the most debilitating but least well-known consequences of stroke (e.g., Simmons-Mackie, Code, Armstrong, Stiegler, & Elman, 2002). In the United States, conservative estimates reveal that over 2.5 million Americans are living with aphasia (Simmons-Mackie, 2018). Many individuals with aphasia (IWAs) recover some function after stroke, but communication difficulties are often chronic and limit the ability to return to work or previous hobbies. Aphasia and the resultant

communication barriers can also worsen social relationships, causing stigma and social isolation (Parr, 2007). Social isolation has been correlated with negative health outcomes including infection (Cohen, Doyle, Skoner, Rabin, & Gwaltney, 1997), depression (Heikkinen & Kauppinen, 2004), cognitive decline (Barnes, Mendes de Leon, Wilson, Bienias, & Evans, 2004; Wilson et al., 2007), and morbidity and mortality (Brummett et al., 2001; Seeman, 2000). As a result, the potential consequences of aphasia are expansive and severe. It is critical to establish which treatment approaches maximize communication outcomes and whether those approaches minimize the social isolation associated with aphasia. This study focuses on group conversation treatment for aphasia and specifically addresses the question of optimal group size for such treatment.

There are many treatment options available to remediate aphasia. One treatment approach used in many aphasia centers across the United States is conversation group

^aDepartment of Communication Sciences and Disorders, Temple University, Philadelphia, PA

^bDepartment of Speech, Language and Hearing Sciences, Boston University, MA

Correspondence to Gayle DeDe: gayle.dede@temple.edu

Editor-in-Chief: Sean Redmond

Editor: Carl Coelho

Received October 4, 2018

Revision received December 22, 2018

Accepted January 15, 2019

https://doi.org/10.1044/2019_JSLHR-L-18-0404

Disclosure: Gayle DeDe and Edwin Maas are salaried employees of Temple University. Elizabeth Hoover is a salaried employee of Boston University and is also on the Board of Directors of Aphasia Access.

treatment (Elman, 2007). Although previously considered an adjunct to individual treatment or as a more cost-effective means of service delivery, group treatment has become increasingly popular because it seems to have certain advantages over individual treatment (Basso, 2005; Elman, 2007; Elman & Bernstein-Ellis, 1999a, 1999b; Pulvermüller et al., 2001). Studies in support of group treatment efficacy and/or effectiveness are growing in number (Aten, Caligiuri, & Holland, 1982; Bollinger, Musson, & Holland, 1993; Elman & Bernstein-Ellis, 1999a, 1999b; Hoover, Caplan, Waters, & Budson, 2015; Wertz et al., 1981). Wertz et al. (1981) reported the first large-scale, randomized controlled trial (RCT) to investigate the efficacy of group treatment of aphasia. Sixty-seven patients, 4 weeks after onset of a single stroke, were randomly assigned to one of two treatment groups for a 44-week treatment program. Wertz et al. compared 4 hr of traditional, stimulus-response treatments in all communicative modalities (reading, writing, auditory comprehension, and verbal expression) plus 4 hr of "machine-assisted treatment and speech-language drill" to 4 hr of clinician-led conversation group treatment with no direct treatment of speech or language deficits, supplemented by 4 hr of recreational activities. Results showed that groups in both cohorts made statistically significant gains on the overall communicative ability scale from the Porch Index of Communicative Ability, word fluency, and a conversation rating scale. To compare the effects of individual and group treatments, the authors completed analyses of covariance in which the intake score at onset served as the covariate for each subsequent period. Both groups changed over time at the same rate, though the individual treatment group demonstrated significantly better performance on the Porch Index of Communicative Ability graphic tasks. This difference is easily explained given that writing was not treated in the conversation group.

Elman and Bernstein-Ellis (1999b) enrolled 24 persons with aphasia into two treatment groups (immediate and deferred) to determine the effect of speech-language group treatment compared to socialization group treatment over a 4-month period. The treatment condition consisted of 5 hr per week of language group treatment across 2 days. Groups were conversational in format and encouraged increased communicative initiation and overall effectiveness using multiple modalities. Elman and Bernstein-Ellis found improved scores on language-dependent measures (SPICA and Western Aphasia Battery) only after treatment; no significant change was noted on the linguistic measures during the baselines when no treatment was provided.

Studies also show that group treatments, along with other opportunities to interact with peers who have aphasia, positively affect quality of life. van der Gaag et al. (2005) evaluated the impact of attending an aphasia center by quantifying participants' perceptions of their communication and quality of life. Over 20 weeks, 28 participants with aphasia received an average of 1.7 hr a week of various group language therapies plus a total of 8 hr of counseling. Outcome measures included two quantitative ratings of quality of life: EuroQoL (EQ-5D; The EuroQol Group, 1990) and the Stroke and Aphasia Quality of Life Measure (Hilari, Byng,

Lamping, & Smith, 2003). A communication rating scale, the Communication Effectiveness Index (Lomas et al., 1989), was also administered. At the end of the 6-month period, there were statistically significant changes on one of the quality of life measures (EuroQoL) and on the communication rating scale. Although the Stroke and Aphasia Quality of Life Measure did not yield statistically significant changes, the direction of change was positive. These results suggest that participation in an aphasia center, which provided group treatments and psychosocial support, positively impacted the individuals' perceptions of their quality of life and overall communication.

Conversation treatment, in contrast to drill-based types of therapies, provides opportunities for language improvisation. Participants are often required to respond to changing topics and novel conversational prompts, which is important for training creative language use (Elman, 2004). Aphasia conversation groups are a prime example of a social approach to intervention because they frequently address practical goals in the context of authentic conversations (Elman, 2011a, 2011b) and are consistent with the Life Participation Approach to Aphasia (Elman, 2011a, 2011b; LPAA Project Group, 2008; Simmons-Mackie, 2008; Simmons-Mackie, Elman, Holland, & Damico, 2007).

Perhaps one of the most powerful benefits of group treatment is the opportunity for members to find a sense of community. Collaboration and psychosocial reinforcement found with group treatment have been shown to positively impact overall health and well-being (Elman, 2007; Hilari & Northcott, 2006; Ryff, Singer, & Love, 2004; Vickers, 2010). Elman and Bernstein-Ellis (1999b) analyzed data taken from qualitative interviews with participants at the Aphasia Center of California during the mixed-methods RCT study described earlier. Participants reported several psychosocial benefits: an enjoyment from being with others, feeling supported by others with aphasia, making friends, an ability to help others, seeing others improving, and a feeling of increased confidence. Linguistically, participants reported noticing that their talking, reading, and writing had improved. Caregivers also noted comparable psychosocial gains such as improved confidence, improved social skills, improved independence, and increased motivation. Some have claimed that, if begun early in the rehabilitation continuum, group membership may reduce the isolation that commonly occurs with aphasia, which in turn may minimize or prevent depression (Elman, 2007).

Despite commonalities in the philosophy and rationale for conversation treatment, there are many differences in its delivery. Simmons-Mackie, Savage, and Worrall (2014) completed a qualitative literature review of 36 studies describing conversation therapy in aphasia. The authors noted that the delivery of conversation therapy differed in many key elements including the therapy participants (e.g., persons with aphasia, professional, volunteer), the overall number of participants (individual, dyad, or group), the focus of the intervention (individualized, problem/solution, or compensatory), and the type of methods used for training (explicit/implicit, external/internal). As part of the conversation therapy

taxonomy, the authors also identified differences in the theoretical orientations or “principal roots” to the intervention group (Simmons-Mackie et al., 2014): conversational analysis, the social model, functional, behavioral, and relationship oriented. Conversational analysis emphasizes the structural features involved in conversation, such as negotiating meaning, taking turns, and repairing breakdowns (Hutchby & Wooffitt, 2008; Simmons-Mackie et al., 2014). The social model, also referred to as a *participation model*, focuses directly on reengagement in life situations for persons with aphasia. In the context of conversation treatment groups, this approach aims to provide ongoing opportunities for conversation and social interaction. Functional, behavioral orientations focus on practical training to change behavior within the conversation interaction between the person with aphasia and another person. This training is often described as conversation coaching of family member training. Finally, relationship-oriented approaches are focused on the relationship between members in the conversation and often employ counseling or couples therapy techniques in the process. These orientations are not mutually exclusive, and indeed, conversation treatment may include aspects from each orientation at several points in the process.

There is evidence that conversation therapy provides significant and lasting benefits to language, communication, and well-being; however, conversation therapy may be administered in many different ways, and we do not yet understand the critical ingredients of this treatment. Further research is warranted to better define treatment protocols, as well as the relative contributions of group therapy to successful participation in life for different profiles of aphasia (Layfield, Ballard, & Robin, 2013).

In order to better understand the underlying mechanism of recovery in conversation treatment, we investigated one ingredient: the number of participants. Conversation treatment can be offered in group (six to eight IWAs), dyadic (two IWAs), or individual formats, but most studies have focused on groups (Simmons-Mackie et al., 2014). Individual or dyadic conversation treatments typically focus on partner training (Hopper, Holland, & Rewega, 2002) or have a more structured format (e.g., cooperative learning principles; Avent, 1997). Savage, Donovan, and Hoffman (2014) reported positive outcomes for two IWAs after individual conversation treatment. However, their study had significant flaws including possible carryover effects from comparison treatment and rising baselines. In summary, there is more evidence regarding the benefits of conversation treatment in the group format than in individual or dyadic formats.

There is little direct evidence regarding the ideal number of participants in conversation treatment, but principles derived from other fields can inform decision making. Group dynamics is a multidimensional, theoretical framework that originated in psychotherapy and identifies beneficial features of group treatments (e.g., Ewing, 2007; Toseland & Rivas, 2017; Yalom & Leszcz, 2005). Meta-analyses of group and individual psychotherapy outcomes suggest both formats are effective in treating a variety of mental illnesses (Burlingame et al., 2015; McRoberts, Burlingame, &

Hoag, 1998). Group therapy brings together individuals with similar interests or problems and provides a place for members to share resources and gain support (Yalom & Leszcz, 2005). Some benefits of group membership, such as emotional support, hope, and universality, relate to minimizing social isolation (Yalom & Leszcz, 2005). Groups engender vicarious learning, such as observing other IWAs use strategies to communicate successfully (Holmes & Kivlighan, 2000; Luterman, 1991). Groups also allow interpersonal learning. For example, IWAs may have more credibility than clinicians when suggesting strategies to one another based on their own experiences.

Additional factors may contribute to group dynamics in conversation treatment. Groups offer a greater variety of communicative functions or speech acts used among members and provide a wider array of communicative partners (Elman, 2007; Elman & Bernstein-Ellis, 1999a). As a result, skills developed in a group may show increased generalization to other environments. All of these factors share the common idea that a group of six to eight individuals with shared experiences offers benefits less available in smaller groups of individuals engaging in the same treatment protocol.

The apparent benefits of group dynamics conflict with a learning principle that is well known in aphasiology: Greater treatment intensity is associated with better treatment outcomes (Basso, 2005; Bhogal, Teasell, Foley, & Speechley, 2003; Cherney, Patterson, & Raymer, 2011; Cherney, Patterson, Raymer, Frymark, & Schooling, 2008, 2010). Intensity, or *cumulative intervention intensity*, as defined by Warren, Fey, and Yoder (2007), is a function of the time over which the intervention is provided (*total intervention duration*; Warren et al., 2007), the number and duration of weekly sessions (*dose frequency*; Warren et al., 2007), and the number of practice trials per session (*dose*; Warren et al., 2007). Most existing studies about treatment intensity in aphasia confound number of treatment hours (Total Intervention Duration × Dose Frequency) with dose because participants who have a higher treatment dose typically also spend more time in treatment. Studies of motor learning suggest that increasing the number of practice trials improves learning, even when the number of sessions is controlled (Maas et al., 2008). It seems intuitive that individual and dyadic treatment sessions would provide a greater number of practice trials than group treatment sessions, even when the total time spent in treatment (in terms of number and duration of sessions) is the same. In the context of conversation treatment, a practice trial is a conversation turn. Thus, individual or dyadic conversation treatment may be advantageous due to the increased opportunity to take conversation turns.

This study compared two conflicting predictions regarding the optimal number of clients in conversation treatment sessions. The group dynamics hypothesis claims that processes inherent to larger groups provide clients with benefits that are not available in smaller groups, meaning that the group treatment condition should outperform the dyadic treatment condition on all outcome measures. In contrast,

the treatment dosage hypothesis states that it is better to have a small number of IWAs in a treatment session because it maximizes each participant's dosage. The treatment dosage hypothesis predicts that the dyadic treatment condition would outperform the group treatment condition on all outcome measures. These competing predictions were tested in an RCT examining effects of conversation treatment on individuals enrolled in large groups or dyads. We contrasted IWAs enrolled in groups with dyadic rather than individual treatment because both dyads and groups include peer-to-peer conversation turns. Importantly, IWAs in both these formats have equal roles in the discourse given their reduced communicative competencies. In addition, a delayed treatment group was included as a no-treatment control condition.

This article presents changes on secondary outcome measures including standardized tests of language function and self-reported measures of functional communication and social networks. Further analyses of discourse, including the primary outcome measure of correct information units (Nicholas & Brookshire, 1993), will be reported separately. Here, we ask whether conversation treatment in large groups or dyads is associated with greater changes on standardized test scores and patient-reported outcome measures.

Method

Participants

Twenty-four IWAs were recruited from both Temple University (TU) and Boston University (BU), for a total of 48 participants for this multicenter, parallel-group, unblinded RCT design efficacy study with a delayed treatment control group and balanced randomization. Figure 1 shows the total number of participants at each site at each stage of the study. Table 1 presents basic descriptive data about the participants. Inclusion criteria were (a) age of 18 years or older, (b) at least 5 months postonset from a stroke, (c) English as a first language (based on self-report), and (d) presence of aphasia (based on clinical judgment and scores on the Comprehensive Aphasia Test [CAT]; see below). Using simple randomization, participants were randomly assigned to one of three conditions when they came for their pre-treatment evaluation, based on a computer-generated random number sequence. The randomization sequence was determined by a statistician who did not interact with participants and provided to each site in an Excel spreadsheet. The three conditions were as follows: large group (six to eight IWAs), dyad (two IWAs), or delayed control group. One-way analyses of variance (ANOVAs) showed that the groups were equivalent with respect to age and years of education (all F s < 1, p s > .40).

After random assignment, dyads were matched on the basis of clinician judgment of compatibility. The decision to assign dyads in this way was made to maximize the clinical relevance of the study. Blinding of participants and clinicians was not possible given the behavioral nature of the intervention, and blinded outcome assessment was not

possible due to resource limitations. Recruitment took place during Spring 2017. The intervention for the large group and dyad conditions took place during Summer 2017; the delayed control condition received treatment during Fall 2017. The protocol was approved by the institutional review board at BU (institutional review board of record) and TU. Informed written consent was obtained from all participants.

Treatment

Participants in the experimental conditions (dyads and large group) received treatment first. Participants allocated to the delayed control condition received treatment (large group) after the first maintenance period. Treatment sessions were 1-hr sessions held twice weekly, for 10 weeks. Group and dyads were facilitated by trained graduate student clinicians from TU and BU, under the supervision of licensed speech-language pathologists (first and second authors) with previous experience administering conversation treatment.

Conversation treatment is inherently unstructured, as it is based on the principle of naturalistic conversation. No two conversation treatment groups are identical; however, the philosophical orientation of the treatment can be described using the taxonomy detailed by Simmons-Mackie et al. (2014). Our conversation treatment approach primarily followed a socially oriented approach. Participants were encouraged to engage in meaningful, authentic conversations surrounding functional topics. Opportunities to participate in conversation were scaffolded through the use of supported conversation techniques (Kagan, Black, Duchan, Simmons-Mackie, & Square, 2001; Simmons-Mackie, King, & Beukelman, 2013). The intervention was individualized to the participants; the supports and opportunities were tailored to the communication needs of those involved in the group.

We took several steps to maximize the reproducibility of the study and the consistency of treatment across sessions and sites. These included developing consistent supports and procedures to be used across sites and offering a consistent, detailed training to group facilitators. We also checked the fidelity of implementation during and after the intervention period.

The first method to enhance consistency was to establish a set of shared topics for treatment sessions. We generated a list of conversation topics that fell under five categories: personal history (family, where you grew up, education/vocation/occupation, hobbies/interests), dining (cooking, eating out, food that travels, holiday food), travel (transportation, sights in Boston/Philadelphia, places in the United States, international travel), news/events (politics, sports, natural disasters, people who inspire you), and entertainment (books/films, TV, art, music). Subthemes of the topics were discussed each session over the 10-week treatment period allowing for each broad topic to be discussed over four sessions (see Appendix). For each topic, a PowerPoint slide show was developed with visual cues and questions to prompt conversation. An outline of potential follow-up questions was also developed, as well as relevant supports. For example,

Figure 1. Study recruitment and attrition. Tx/tx = treatment; pt = participant; F/U = follow-up; mos = months; exp = experimenter.

for the topic of “family,” prompts included “In my family, I have...,” with follow-up questions such as “Do you have any sisters or brothers” and “What is your sister’s name?” Relevant supports included blank family trees and visual depictions of relationships such as wife, husband, mother, father, uncle, cousin, and so on. Materials were developed by the graduate student clinicians and were used for the large groups and dyads at both sites as natural supports for the conversation when necessary.¹ In addition to topic-specific supports, a standard set of visual supports were

available during all conversation treatment sessions (Simmons-Mackie et al., 2013). These included writing implements (pen and paper or whiteboards), maps, picture dictionaries, and a laptop or tablet computer as a digital librarian. Thus, all conversation treatment sessions had the same set of materials available to prompt and support conversation as necessary based on the needs of the participants.

Each session began with a general introductory question such as “How is everyone today?” Then, conversation was facilitated throughout the session using the materials described above. It is important to note that, once the conversation topic was introduced, the clinicians supported the conversation turns of the participants. Clinicians did

¹Examples of stimuli are available from the contact author by request.

Table 1. Demographic data by group and site.

	Boston (<i>n</i> = 24)			Temple (<i>n</i> = 22)			All (<i>N</i> = 46)		
	Dyad	Large group	Delay	Dyad	Large group	Delay	Dyad	Large group	Delay
Age (years)	70.5 (9.7)	61.4 (20.9)	71.4 (9.0)	60.8 (6.6)	61.4 (5.9)	58.4 (11.7)	65.6 (9.4)	61.4 (14.3)	65.8 (11.9)
Education (years)	16.4 (3.5)	15.1 (3.1)	16.9 (2.1)	13.7 (1.5)	12.6 (2.2)	13.8 (3.7)	15.0 (2.9)	13.8 (2.9)	15.6 (3.2)
Gender (M/F)	7/1	5/3	8/0	5/3	6/1	1/6	12/4	11/4	9/6
CAT naming total score ^a	29.9 (27)	50.6 (21.8)	45.5 (23.6)	52.1 (24.5)	49.8 (21.5)	59.5 (16.7)	41 (27.4)	50.2 (20.9)	51.5 (21.5)
Aphasia syndrome									
Anomic	1	2	1	4	4	2	5	6	3
Broca's	1	3	2	2	4	1	3	7	3
Conduction	2	2	3			2	2	2	5
Sv MNF			2	1			1		2
TCM	1					1	1		1
TCS	1						1		
Wernicke's	2	1		1			3	1	

Note. Aphasia syndromes were based on CAT scores and clinician judgment. Boston = Boston University; Temple = Temple University; M/F = male/female; CAT = Comprehensive Aphasia Test; Sv MNF = severe mixed nonfluent; TCM = transcortical motor; TCS = transcortical sensory.

^aValues are mean (standard deviation).

not direct the flow of the topic but rather followed the natural course of the conversation and only used prompts to move the conversation forward as appropriate. The clinicians tried to ensure that all IWAs had the opportunity to contribute conversation turns and IWAs were cued to use supports as appropriate. Clinicians modeled use of strategies in order to normalize multimodal communication. The clinicians used gestures or wrote key words to supplement speech at least three times per session. In addition, strategies such as writing or gesturing key words and/or repeating/rephrasing key information were used to facilitate group understanding.

Finally, each participant had two individual goals that were identified based on testing performance and the IWAs' goals, which were discussed before initiating treatment. Individual goals were targeted by creating opportunities to practice the communication target within the natural flow of the conversation. For example, one participant's goal was to produce complete subject-verb-object sentences. He or she was provided with a personalized visual cue. In addition, the clinicians frequently recast sentence fragments as complete sentences or, if appropriate, cued the participant to reformulate a conversation turn as a complete sentence. For example, if another group member needed to hear something again, the participant might be asked to repeat an utterance as a complete sentence.

Treatment Fidelity

The second method used to maximize the consistency of treatment was for all graduate student clinicians to receive the same training. All clinicians received a treatment manual with information about conversation treatment, cueing hierarchies, and weekly topics. They completed a 2-hr orientation via Cisco WebEx to review procedures and goals. Performance was reviewed each week to ensure correct implementation of the protocol. All clinicians provided treatment in both conditions (large group and dyad).

In addition, trained undergraduate students used a checklist to record variables related to fidelity in each session. To ensure consistency in the implementation of groups and dyads, we recorded the number of clinician models for multimodal communication strategies.

Another goal of fidelity coding was to confirm assumptions about groups and dyads. To verify that dyads provided a higher dosage of conversation turns than groups, the observer counted the number of IWAs in the session (in case participants missed sessions), the total number of turns, and the number of conversation turns for each IWA and clinician. Any attempt at communication, regardless of modality or accuracy, was counted as a conversational turn. An independent observer viewed 20% of the treatment sessions to determine reliability of the online coding.

Assessment Protocol

After random assignment to group, participants completed the initial evaluation. Participants were also evaluated

immediately after the 10-week treatment period and again 6 weeks later. For the dyads and large group, these testing sessions corresponded to pretreatment, posttreatment, and 6-week maintenance testing. In addition, maintenance for the dyads and large group is also reported 11 months after initial testing. IWAs did not participate in concurrent speech-language treatment from the point of initial testing through the 6-week maintenance period. After that time, it was considered unethical to restrict their involvement in speech-language treatment.

All testing was completed by a licensed speech-language pathologist or by a graduate student clinician under supervision of a licensed speech-language pathologist. The primary outcome measure was percent correct information units (Nicholas & Brookshire, 1993), based on the average score across four speech samples. In the present article, we present results from the secondary outcome measures, which include the standardized language tests and patient-reported outcome measures. Data were entered into REDCap (Research Electronic Data Capture), hosted at BU (Harris et al., 2009).

Standardized Tests of Language in Aphasia

The CAT (Swinburn, Porter, & Howard, 2004) documented the presence of aphasia and served as a measure of general treatment effects. The CAT is a psycholinguistic test battery that targets expressive and receptive language in oral and written modalities. We administered five sections: auditory comprehension, reading comprehension, verbal expression, repetition, and picture description. Subsection scores were analyzed to examine changes as a function of treatment.

The Northwestern Assessment of Verbs and Sentences (NAVS; Cho-Reyes & Thompson, 2012) Sentence Production Priming Test measured production of canonical and noncanonical sentences. It provided a standardized measure of whether treatment led to improved sentence production. Given the variability in severity across the IWAs, the standardized administration was modified such that the test was discontinued if a participant scored 0 on one subsection. The dependent measure was number correct, with a maximum score of 32.

The NAVS (Cho-Reyes & Thompson, 2012) Verb Naming Test (VNT) is a single-word confrontation naming test of verbs used to evaluate changes in naming ability. The dependent measure was number correct, with a maximum score of 22.

The Philadelphia Naming Test (Short Form; Walker & Schwartz, 2012) is a single-word confrontation naming test of nouns used to evaluate changes in naming ability. The dependent measure was number correct, out of 30 items.

Standardized Patient-Reported Outcomes

The Aphasia Communication Outcome Measure (ACOM; Hula et al., 2015) is a psychometrically validated measure of the impact of aphasia on communication in daily life. The ACOM asks questions such as how effectively the IWA talks to family members and strangers or how

effectively the IWA writes a short to-do list. We administered the adaptive short form of the ACOM, which uses the IWA's response to each item to select subsequent items.

The Lubben Social Network Scale-6 (LSNS-6; Lubben et al., 2006) is a relatively simple questionnaire that has been used to quantify social isolation in many studies of aging adults. Tang et al. (2005) showed that lower LSNS-6 scores were associated with higher rates of poststroke depression. The LSNS-6 asks about the number of family and friends who the participant (a) has been in touch with at least once a month, (b) would ask for help, and (c) would discuss private matters with. The examiner read questions aloud to the participant when necessary. It is scored out of 30 points.

Results

Effects of Treatment

Given the attrition of participants across groups (see Figure 1), the study was underpowered for ANOVAs to study treatment effects. For this reason, the data were analyzed using nonparametric statistics. Outcome measures at Time 1 (pretreatment) were analyzed using Kruskal–Wallis tests to determine whether the groups differed before treatment. There were no significant effects of group (all p s > .20), suggesting that the randomization resulted in statistically equivalent groups.

To examine treatment effects, data were analyzed separately by group using nonparametric Wilcoxon signed-ranks tests² in R Version 3.4.2. Data were compared for pretreatment versus posttreatment and for pretreatment versus maintenance. This corresponds to the first testing session versus the second one and the first testing session versus the third testing. The delay control condition did not receive any treatment during this period. For the large group and dyad conditions, data were also compared from pretreatment to the 11-month follow-up. Holm corrections were used to correct for multiple comparisons between groups. The full set of results is presented in Table 2. Table 2 denotes all comparisons that were significant at $p < .05$ and comparisons that survived the Holm correction. In the interest of brevity, here, we discuss only dependent variables that were associated with significant changes that survived the Holm correction.

CAT Repetition Scores

The dyads showed a significant improvement from pre- to posttreatment testing in the CAT repetition total score, $V = 4.5$, $p = .003$, which was maintained at the 6-week follow-up, $V = 6.5$, $p = .004$. The improvement was maintained 11 months later, but the effect was not significant, $V = 16$, $p = .08$. No other groups showed a significant change in repetition performance.

²Note that R reports a V statistic (rather than W) for the Wilcoxon signed-ranks tests. V represents the sum of all positive ranks and is equivalent to W .

CAT Naming Scores

The dyads showed a significant improvement on the CAT naming score from pre- to posttreatment, $V = 15$, $p = .007$. The scores at follow-up showed a trend toward improved performance, but the effects did not survive the Holm correction (6 weeks: $V = 19.5$, $p = .04$; 11 months: $V = 14$, $p = .03$).

Participants in the large group also showed a significant improvement on CAT naming scores from pre- to posttreatment, $V = 7$, $p = .019$. The change was not significant at the 6-week follow-up testing, $V = 5$, $p = .15$, or 11 months later, $V = 10$, $p = .08$. The delayed control group showed no significant changes on the CAT naming scores.

CAT Picture Description Total Score

The CAT picture description total score combines number of information carrying words as well as ratings of grammaticality, speech rate, and diversity of syntactic structures. Participants in the large group showed a significant improvement from pre- to posttesting on the CAT picture description total score, $V = 6$, $p = .017$. This improvement was maintained at the 6-week follow-up, $V = 3$, $p = .024$, and 11 months later, $V = 3$, $p = .014$. No other groups showed a significant change.

VNT

The dyads showed a significant improvement from pre- to posttreatment testing, $V = 6.5$, $p = .012$. This change was not significant at the 6-week follow-up, $V = 9$, $p = .12$, though there was also a significant change 11 months after initial testing, $V = 2$, $p = .01$.

The delayed control group also showed an improvement on VNT scores from pre- to posttreatment testing (though they received no treatment in this period), but the effect did not survive the Holm correction, $V = 11$, $p = .03$. The large group showed no significant changes in the VNT scores.

Adaptive Aphasia Outcome Measure (ACOM)

Participants in the large group showed a significant improvement from pre- to posttesting on the ACOM, $V = 2$, $p = .011$. At the 6-week follow-up, the improvement did not survive the Holm correction, $V = 7$, $p = .037$. The change was nonsignificant at 11 months posttesting, $V = 7$, $p = .56$.

The dyads showed an improvement from pretreatment to the 6-week follow-up testing, though the effect did not survive the Holm correction, $V = 21$, $p = .05$, and was not significant at 11 months posttesting, $V = 13$, $p = .81$. There were no other significant changes in the ACOM.

Treatment Fidelity

Attendance

Mean, standard deviation, and range for number of sessions attended by IWAs in groups and dyads are presented in Table 3, separately by test site. A 2×2 between-participants ANOVA with attendance as the dependent variable revealed no significant differences by condition (group

Table 2. Results of standardized testing for all conditions and time points.

Measure	Delay			Dyads				Large group			
	Pre	Post	6 wks	Pre	Post	6 wks	11 months	Pre	Post	6 wks	11 month
Comprehensive Aphasia Test section scores											
Auditory comprehension	47.8 (9.9)	52.0* (6.3)	50.5 (9.1)	43.8 (14.7)	45.8 (16.3)	40.3 (20.6)	48.1 (15.9)	48.4 (11.0)	50.5 (12.0)	49.7 (11.3)	50.5 (10.7)
Reading comprehension	44.6 (7.5)	43.3 (9.6)	46.1 (9.1)	39.3 (16.2)	37.6 (16.2)	34.4 (20.6)	40.7 (14.2)	44.9 (10.9)	45.6 (10.9)	49.2* (8.1)	47.4 (11.1)
Repetition	48.9 (19.7)	51.9 (19.1)	52.0 (19.7)	44.6 (23.3)	50.0** (22.5)	46.2** (23.2)	48.4 (22.6)	52.3 (14.1)	55.0 (14.9)	52.5 (13.1)	55.4 (15.8)
Naming	51.9 (21.7)	49.0 (22.1)	49.2 (18.6)	41.0 (27.4)	45.8** (28.7)	41.9* (29.1)	45.6 (33.6)	47.5 (21.6)	56.7** (24.4)	53.1 (22.2)	56.3 (27.2)
Oral reading	43.1 (18.7)	41.9 (18.7)	44.1 (18.8)	33.6 (24.4)	35.4 (26.3)	33.8 (25.0)	36.4 (25.9)	44.1 (18.0)	45.7* (19.1)	43.8 (20.7)	48.8 (19.9)
Pic descript total score	18.0 (11.4)	22.8 (11.2)	25.1 (13.9)	16.9 (15.0)	20.1 (12.9)	15.4 (24.1)	20.5 (20.1)	14.1 (9.1)	18.1** (10.3)	21.1** (11.8)	23.9** (15.0)
Northwestern Assessment of Verbs and Sentences											
SPPT	14.2 (12.0)	15.1* (11.5)	13.3 (11.0)	10.5 (10.9)	12.3* (11.5)	10.1 (11.5)	12.6 (11.2)	10.6 (11.3)	11.6 (10.5)	11.8 (12.5)	14.4 (12.1)
VNT	14.9 (5.4)	15.1 (5.7)	16.2* (6.0)	11.0 (8.1)	13.1** (8.7)	11.6 (8.6)	13.8** (8.2)	15.5 (6.9)	15.9 (6.0)	16.2 (5.9)	15.2 (7.3)
Confrontation naming test											
Philadelphia Naming Test	22.9 (6.7)	25.6 (6.3)	25.8 (7.3)	18.1 (11.9)	19.8 (11.5)	17.6 (13.3)	19.3 (13.0)	21.9 (7.9)	22.0 (7.7)	23.8 (7.0)	24.2 (8.0)
Patient-reported outcome measures											
Lubben	16.4 (8.1)	14.7 (6.7)	15.7 (6.5)	18.3 (7.9)	17.4 (7.0)	16.6 (6.0)	16.7 (7.5)	17.1 (6.5)	16.0 (6.5)	15.9 (6.5)	15.8 (6.3)
ACOM	51.7 (8.8)	52.4 (9.2)	54.0 (8.3)	47.7 (10.4)	51.2 (10.7)	50.9* (7.4)	51.6 (7.5)	48.3 (8.8)	52.9** (9.8)	52.7* (11.5)	51.4 (9.5)

Note. Effects that survived the Holm correction are bolded and denoted with two asterisks (**). One asterisk (*) indicates that the effect was significant at $p < .05$ but did not survive the Holm correction. Pre = pretreatment; Post = posttreatment; wks = weeks; Pic descript = picture description; SPPT = Sentence Production Priming Test; VNT = Verb Naming Test; Lubben = Lubben Social Network Scale; ACOM = Adaptive Aphasia Communication Outcome Measure.

Table 3. Treatment fidelity.

	Dyads	Large group
Attendance		
Boston	19.0 (<i>SD</i> = 0.8, range: 18–20)	18.0 (<i>SD</i> = 1.5, range: 15–19)
Temple	17.9 (<i>SD</i> = 3.1, range: 11–20)	17.3 (<i>SD</i> = 2.6, range: 13–20)
Number of conversational turns		
Boston	89.4 (<i>SD</i> = 19.1, range: 70.3–119.0)	31.2 (<i>SD</i> = 13.1, range: 14.7–49.2)
Temple	151.6 (<i>SD</i> = 18.8, range: 124.1–180.2)	87.8 (<i>SD</i> = 29.6, range: 50.9–129.9)
Number of clinician models		
Boston	18.6 (<i>SD</i> = 9.8, range: 4–51)	19.4 (<i>SD</i> = 23.4, range: 0–97)
Temple	12.2 (<i>SD</i> = 4.7, range: 3–21)	13.3 (<i>SD</i> = 4.6, range: 6–20)

vs. dyad) or site (BU vs. TU). All *F* values were less than or equal to 1.25, and all *p* values were greater than or equal to .28.

Number of Conversation Turns

Two independent coders' counts of conversation turns were compared for 20% of the treatment sessions. Coding of conversation turns was a challenge given overlapping turns, interruptions, and so on. For this reason, we asked whether the counts of conversation turns differed by more than 10 for any individual in any session. For 83.6% of the observations (across IWAs and sessions), the coders' counts of conversation turns met the criterion. On this basis, the reliability of conversation turn coding was deemed acceptable.

As expected, the average number of speaking turns differed as a function of condition (see Table 3). A 2×2 between-participants ANOVA revealed a main effect of group size, $F(1, 24) = 59.72, p < .0001$, with more turns in the dyads than in the larger group. There was also a main effect of site, $F(1, 24) = 58.63, p < .0001$, with participants having fewer speaking turns at TU than BU. There was no interaction between condition and site, $F(1, 27) = 0.12, p = .73$.

Number of Clinician Models

Clinicians modeled use of strategies at least three times in 93.5% of the treatment sessions. Sessions with fewer than three models occurred in dyads in which the participants were relatively mildly impaired, making multimodal strategies less appropriate. The average number of models varied widely, as seen in Table 3. A 2×2 ANOVA showed a main effect of condition, $F(1, 19) = 29.08, p < .0001$, with more strategies being modeled in the dyads than in the larger groups. The effect of site was not significant, $F(1, 19) = 3.3, p = .09$, nor was the interaction between condition and site, $F(1, 19) = 0.22, p = .64$.

Discussion

This RCT examined whether the group dynamics or the treatment dosage hypothesis best predicted treatment outcome data for IWAs who completed 10 weeks of conversation treatment. This study reports secondary outcome measures. An important finding was that conversation treatment was associated with gains on standardized measures of language. This finding is consistent with previous studies (e.g., Elman & Bernstein-Ellis, 1999b) and suggests that effects of conversation treatment are robust and replicable. However, with respect to our specific hypotheses, the data point to a more complex picture than we predicted, with the dyad and large group conditions showing changes on different types of measures.

Effects of Conversation Treatment

Both treatment conditions showed significantly improved performance from pre- to posttreatment on the CAT naming subtest, whereas the delayed control group showed a nonsignificant downward trend. The CAT naming subtest comprises verbal fluency, object naming, and action naming. The scoring criteria are sensitive to both accuracy and speed of performance. These results suggest that conversation treatment is associated with improved naming performance, regardless of group size. More generally, the treatment conditions showed a greater number of significant improvements in performance from pre- to posttesting than the delayed control group did over the same period. In fact, the delayed control group showed no significant changes that survived the Holm correction. In this way, our data are consistent with previous studies (e.g., Elman & Bernstein-Ellis, 1999a) showing that conversation treatment is associated with improved performance on standardized measures of language performance.

To our knowledge, this study is the first to report 11-month follow-up data for participants in conversation treatment. The only effect that was significant at all time points was picture description total scores for the large group. This result should be interpreted cautiously given the fact that participants may have received other treatments during the follow-up interval and that we are only reporting data from one picture description task (cf. Boyle, 2014). However, it is interesting to note that it was a discourse measure that showed persistent changes. We did not track the extent to which the IWAs participated in other individual or group treatment after the study. Where treatment effects do not persist, it would be important to know whether continued involvement in groups or targeted booster sessions would extend effects that were observed immediately posttreatment.

Effects of Group Size

There was some indication that, beyond the CAT naming scores, the two treatment conditions were associated with improvement on different types of dependent variables. The dyads seemed to show more improvement on

task-specific standardized tests, that is, the CAT Repetition subtest total score and the VNT. This is consistent with the treatment dosage hypothesis. However, the large group showed significantly improved performance on the CAT picture description total score and the ACOM. These results are consistent with the group dynamics hypothesis. The ACOM and picture description tasks may be viewed as reflecting functional language to a greater extent than more discrete tasks such as repetition or naming. Taken together, the results suggest that increased practice trials are most important when the treatment outcome is improvement on discrete linguistic tasks. On the other hand, group dynamics may facilitate greater changes in more functional measures, including the participants' own perceptions of their communication abilities.

Many of the standardized measures were not sensitive to changes in either the dyadic or large group condition. For discrete language tasks, it may be that there were insufficient participants with deficits in some domains to show significant treatment effects. For example, many of our participants scored at or near ceiling on the CAT auditory comprehension task. Participants with more severe impairments may have shown improvements in this domain, but it might not have been enough to achieve statistical significance given the number of participants at ceiling. Nonetheless, we did observe changes in several of the standardized measures of expressive language, including repetition and naming tasks.

One question is why conversation treatment would be expected to generalize to standardized test scores, particularly of discrete linguistic tasks such as naming. This type of generalization might be predicted by a complexity account (Elman, 2011c; Thompson & Shapiro, 2007) in which treatment directed at a more complex task (retrieving words in conversational discourse) may generalize to a less complex task (e.g., retrieving single words in a picture naming task). However, generalization to discrete language tasks may require more practice trials, which might explain why the dyad condition showed more generalization to discrete language tasks than the larger group condition.

We were surprised that the LSNS did not show significant changes as a function of treatment. We anticipated that participants in both treatment conditions might report increased involvement with family members and new friendships. In retrospect, we underestimated the variability in potential interpretations of the term *friendship*. Informal conversations after the final testing session suggested that most participants thought about rekindling friendships that preceded the onset of aphasia, rather than forming new friendships during treatment. When asked directly, many participants did report forming or strengthening friendships during treatment. The LSNS was also insensitive to changes in friendship in a study that investigated the effects of a psychosocial group treatment on older adults' self-ratings of loneliness or social network size (Routsalo, Tilvis, Kautiainen, & Pitkala, 2008). The LSNS was chosen because it is a well-established measure. However, it might not be sensitive to changes associated with psychosocial treatments. In the future, a modified administration or

another measure may be preferable to investigate the effects of group treatment on social networks.

A question for future studies is whether a different pattern of results would be observed with individual conversation treatment rather than dyads or groups. We chose to focus on dyadic rather than individual conversation treatment because dyads were thought to promote a more naturalistic conversation than discourse with a clinician. In individual therapy, IWAs may rely on the clinician to direct the conversation and repair communication breakdowns (Avent, 1997, 2004). In contrast, IWAs in dyads had conversations with one another, with the clinician acting as a facilitator. Given our interest in comparing relatively naturalistic conversations between IWAs, our study did not maximize dosage relative to individual sessions, as IWAs might have even more practice trials in an individual session. Similarly, our study did not maximize the group dynamics difference between conditions, as dyads also offer a degree of peer-to-peer support.

Although not significant, there were interesting numeric differences in attendance between conditions and treatment sites that warrant some discussion in relation to implementing conversation treatment. Attendance was generally more consistent in the dyadic than group condition. Inspection of the data in Table 3 shows that the minimum attendance was lower in the group condition (18 vs. 13 at BU and 13 vs. 11 at TU). One possible reason is that dyads, but not groups, could be rescheduled if one participant was aware of scheduling conflicts in advance. Additionally, participants in dyads may have felt more individual responsibility to their partner than participants in the larger group. Thus, there may be reason to expect better attendance in dyads than larger groups.

Another trend was that attendance was numerically higher at BU than TU. The minimum number of sessions attended was 15 at BU but 11 at TU. This trend may reflect demographic differences between the sites. Although not formally assessed, several of the IWAs at TU (but not BU) relied on paratransit or lengthy rides on public transportation (more than two transfers) to attend treatment sessions. As a result, IWAs at TU sometimes missed treatment sessions when their paratransit was more than 1 hr late or when weather made the trip unreasonable. This observation points to some of the challenges faced by clients who do not have the means to arrange expeditious transportation.

In this study, treatment was conducted by trained graduate student clinicians under the supervision of speech-language pathologists with experience implementing conversation treatment. It may be that more experienced clinicians, rather than graduate students, would effect greater treatment changes. As a result, this study may underestimate the potential magnitude of effects associated with conversation treatment. However, given that the same clinicians conducted both dyadic and large group treatments, it is unlikely that this variable influenced differences between the conditions. In addition, using graduate student clinicians increases the external validity of our study: If the success of conversation treatment relied on having very experienced clinicians

facilitate the groups, it would be difficult to implement conversation treatment in real-world settings. Showing that graduate student clinicians can effect change suggests that novice clinicians can facilitate effective groups, albeit with support from more experienced colleagues.

Another point of discussion is the nature of our conversation groups. As Simmons-Mackie et al. (2014) noted, there are many different models for conversation groups. They identified different “principal roots,” or theoretical origins, of conversation therapy. These principal roots are not mutually exclusive but differ in broad strokes. Our conversation treatment was motivated primarily by the “life participation” root, which is based in a social model and focuses on opportunities for authentic conversation. In its purest form, a “life participation” group might not have a set topic for a session. Instead, clinicians might open the group by asking clients what they want to talk about that day. For the purpose of a research study, especially a two-site study, it was important that we maximize the consistency of our sessions across sites and the replicability of our study by other clinicians and researchers. To achieve this goal, we identified topics for each meeting and had topic-specific supports that were shared across sites, groups, and dyads. Nonetheless, the conversations about those topics were directed by the interests of the participants, and the clinicians worked to foster genuine interactions among group members. In this way, the group conformed to the “life participation” root described by Simmons-Mackie et al. (2014).

The clinicians also used strategies to facilitate the conversation. Many of the strategies were based on Simmons-Mackie et al.’s (2007) descriptions of well-run groups. For example, the clinicians encouraged, but did not require, all group members to participate in the conversation using eye gaze and silence and by asking questions of quieter group members. The clinicians also sometimes encouraged group members to use strategies such as writing or external aids or reminded clients to rephrase utterances in complete sentences (if that was the client’s individual goal). These aspects of facilitation share some features with other types of conversation treatment, such as the functional/behavioral and conversation analysis roots identified by Simmons-Mackie et al. (2014). These principal roots focus more explicitly on teaching strategies for conversational repair or on other strategies that may improve conversational interactions. However, these strategies were used carefully and not at the expense of natural conversation turns. Thus, even though aspects of our conversation treatment may overlap in some ways with other “principal roots,” our conversation treatment was primarily grounded within a social model/life participation approach.

Limitations and Future Directions

This study had a number of limitations. First, though we recruited a relatively large number of participants, attrition resulted in a relatively small number of participants per condition for an RCT. As a result, we could not directly compare performance across groups and time points. Although

we did control for multiple comparisons in our nonparametric analyses, we were not able to directly compare effect sizes in the three groups, which limits the strength of conclusions we can draw.

Second, we only report data from one picture description task. Boyle (2014) has shown that narrative analysis for people with aphasia is most reliable when multiple narratives are incorporated. Here, we reported results of the coding scheme for the CAT, which has been standardized on an independent sample of people with aphasia (Swinburn, Porter, & Howard, 2005). Swinburn et al. showed that test–retest reliability was .74 ($n = 21$) at 10 weeks and .90 ($n = 48$) at 6–12 months. According to Boyle (2014), these values are sufficient for group research studies. In addition, we will be able to report results from multiple discourse samples in future analyses of our data.

As in most behavioral clinical trials, blinding of participants and clinicians was not possible. We did not use blinded outcome assessors due to resource limitations. However, the outcome measures reported here were all standardized measures, rather than subjective outcome ratings (except perhaps for the ACOM and LSNS). Furthermore, for the comparison between dyads and groups, we maintained equipoise by formulating alternative hypotheses and never presenting one condition as likely to be more effective than another to participants nor to treating clinicians.

Coding of conversation turns proved difficult, resulting in acceptable but not strong reliability for this measure of fidelity. For the purposes of this study, it is unlikely that this issue significantly influenced the results given the degree of difference between the two treatment conditions. Despite its limitations, the coding of conversation turns did appear to validate the expectation that IWAs in dyads would have more speaker turns than IWAs in a larger group. Future work, with a larger number of participants, could examine whether there is a relationship between session variables such as number of sessions attended or number of speaking turns and treatment outcomes.

Conclusion

In summary, this study examined effects of conversation treatment in groups and dyads on standardized test performance. Consistent with previous studies, our results suggest that conversation treatment is associated with at least short-term and some long-term gains in performance for individuals with chronic aphasia. In addition, this study provided more detailed information about how we implemented conversation treatment, in an effort to facilitate replication in clinical and research settings. The results expand on previously reported studies by including a longer term follow-up assessment and by providing information about the potential benefits of the components of conversation treatment. Specifically, the data showed that group dynamics seemed to be associated with more changes in functional measures of performance, whereas a larger number of practice turns in dyadic sessions was associated with increased task-specific performance.

Acknowledgments

This work was funded by Grant R21DC015859 (awarded to Elizabeth Hoover and Gayle DeDe) from the National Institute on Deafness and Other Communication Disorders. Preliminary findings were reported at the 2018 Clinical Aphasiology Conference (Austin, TX, June 2018) and the International Aphasia Rehabilitation Conference (Aveiro, Portugal, September 2018) and are published in abstract form (Hoover, DeDe, & Maas, 2018). The authors would like to thank David Caplan, Audrey Holland, and Ting Dai for their help at the initial stages of this project. They also thank the students who assisted in the study, including Alexandra McFee, Rachael Campbell, Andrea Chang, Leela Rao, Rebecca Goldenberg, Chrissy Sandilos, Jaclyn Navarro, and Mary Mash, and also William Hula for sharing the ACOM materials. Finally, they thank the participants and their families, without whom this would not have been possible. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Aten, J. L., Caligiuri, M. P., & Holland, A. L. (1982). The efficacy of functional communication therapy for chronic aphasic patients. *Journal of Speech and Hearing Disorders*, 47, 93–96.
- Avent, J. R. (1997). *Manual of cooperative group treatment for aphasia*. Oxford, United Kingdom: Butterworth-Heinemann Medical.
- Avent, J. R. (2004). Group treatment for aphasia using cooperative learning principles. *Topics in Language Disorders*, 24(2), 118–124.
- Barnes, L. L., Mendes de Leon, C. F., Wilson, R. S., Bienias, J. L., & Evans, D. A. (2004). Social resources and cognitive decline in a population of older African Americans and whites. *Neurology*, 63, 2322–2326.
- Basso, A. (2005). How intensive/prolonged should an intensive/prolonged treatment be? *Aphasiology*, 19(10–11), 975–984.
- Bhagal, S. K., Teasell, R. W., Foley, N. C., & Speechley, M. R. (2003). Rehabilitation of aphasia: More is better. *Topics in Stroke Rehabilitation*, 10, 66–76.
- Bollinger, R. L., Musson, N. D., & Holland, A. D. (1993). A study of group communication intervention with chronically aphasic persons. *Aphasiology*, 7, 301–313.
- Boyle, M. (2014). Test–retest stability of word retrieval in aphasic discourse. *Journal of Speech, Language, and Hearing Research*, 57(3), 966–978.
- Brummett, B. H., Barefoot, J. C., Siegler, I. C., Clapp-Channing, N. E., Lytle, B. L., Bosworth, H., . . . Mark, D. B. (2001). Characteristics of socially isolated patients with coronary artery disease who are at elevated risk for mortality. *Psychosomatic Medicine*, 63(2), 267–272.
- Burlingame, G. M., Gleave, R., Erikson, D., Nelson, P. L., Olsen, J., Thayer, S., & Beecher, M. (2015). Differential effectiveness of group, individual, and conjoint treatments: An archival analysis of OQ-45 change trajectories. *Psychotherapy Research*, 26(5), 556–572.
- Cherney, L. R., Patterson, J. P., Raymer, A., Frymark, T., & Schooling, T. (2008). Evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language, and Hearing Research*, 51, 1282–1299.
- Cherney, L. R., Patterson, J. P., Raymer, A., Frymark, T., & Schooling, T. (2010). *Updated evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia*. Rockville Pike, MD: American Speech-Language-Hearing Association.
- Cherney, L. R., Patterson, J. P., & Raymer, A. M. (2011). Intensity of aphasia therapy: Evidence and efficacy. *Current Neurology and Neuroscience Reports*, 11, 560–569.
- Cho-Reyes, S., & Thompson, C. K. (2012). Verb and sentence production in aphasia: Northwestern assessment of verbs and sentences (NAVS). *Aphasiology*, 26, 1250–1277.
- Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M., Jr. (1997). Social ties and susceptibility to the common cold. *Journal of the American Medical Association*, 277, 1940–1944.
- Elman, R. J. (2004). Group treatment and jazz: Some lessons learned. In J. F. Duchan & S. Byng (Eds.), *Challenging aphasia therapies: Broadening the discourse and extending the boundaries* (pp. 144–147). London, United Kingdom: Psychology Press.
- Elman, R. J. (2007). The importance of aphasia group treatment for rebuilding community and health. *Topics in Language Disorders*, 27, 300–308.
- Elman, R. J. (2011a). *The complexity hypothesis: Aphasia research & ideas about conversation*. Paper presented at the American Speech-Hearing Convention, San Diego, CA.
- Elman, R. J. (2011b). Social and life participation approaches to aphasia intervention. In L. LaPointe (Ed.), *Aphasia and related neurogenic language disorders* (4th ed., pp. 171–184). New York, NY: Thieme.
- Elman, R. J. (2011c). Starting an aphasia center? *Seminars in Speech and Language*, 32(3), 268–272.
- Elman, R. J., & Bernstein-Ellis, E. (1999a). The efficacy of group communication treatment in adults with chronic aphasia. *Journal of Speech, Language, and Hearing Research*, 42, 411–419.
- Elman, R. J., & Bernstein-Ellis, E. (1999b). Psychosocial aspects of group communication treatment. Preliminary findings. *Seminars in Speech and Language*, 20, 65–71.
- Ewing, S. (2007). Group process, group dynamics, and group techniques with neurogenic communication disorders. In R. Elman (Ed.), *Group treatment for neurogenic communication disorders: The expert clinician's approach* (2nd ed., pp. 11–23). San Diego, CA: Plural.
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381.
- Heikkinen, R. L., & Kauppinen, M. (2004). Depressive symptoms in late life: A 10-year follow-up. *Archives of Gerontology and Geriatrics*, 38, 239–250.
- Hilari, K., Byng, S., Lamping, D. L., & Smith, S. C. (2003). The Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of acceptability reliability and validity. *Stroke*, 34, 1944–1950.
- Hilari, K., & Northcott, S. (2006). Social support in people with chronic aphasia. *Aphasiology*, 20, 17–36.
- Holmes, S. E., & Kivlighan, D. M., Jr. (2000). Comparison of therapeutic factors in group and individual treatment processes. *Journal of Counseling Psychology*, 47, 478–484.
- Hoover, E. L., Caplan, D., Waters, G., & Budson, A. (2015). Effects of impairment-based individual and socially oriented group therapies on verb production in aphasia. *Aphasiology*, 29, 781–798.
- Hoover, E. L., DeDe, G., & Maas, E. (2018). Effects of group size on conversation treatment outcomes: Results of standardized testing. *Aphasiology*, 32(Suppl. 1), 93–95.
- Hopper, T., Holland, A., & Rewega, M. (2002). Conversational coaching: Treatment outcomes and future directions. *Aphasiology*, 16, 745–761.
- Hula, W. D., Doyle, P. J., Stone, C. A., Austermann Hula, S. N., Kellough, S., Wambaugh, J. L., . . . St Jacques, A. (2015). The aphasia communication outcome measure (ACOM):

- Dimensionality, item bank calibration, and initial validation. *Journal of Speech, Language, and Hearing Research*, 58, 906–919.
- Hutchby, I., & Wooffitt, R. (2008). *Conversation analysis* (2nd ed.). Cambridge, United Kingdom: Polity.
- Kagan, A., Black, S. E., Duchan, F. J., Simmons-Mackie, N., & Square, P. (2001). Training volunteers as conversation partners using “supported conversation for adults with aphasia” (SCA): A controlled trial. *Journal of Speech, Language, and Hearing Research*, 44(3), 624–638.
- Layfield, C. A., Ballard, K. J., & Robin, D. A. (2013). Evaluating group therapy for aphasia: What is the evidence? *EBP Briefs*, 7(5), 1–17.
- Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989). The communicative effectiveness index: Development and psychometric evaluation of a functional measure for adult aphasia. *Journal of Speech and Hearing Disorders*, 54, 113–124.
- LPAA Project Group. (2008). Life participation approach to aphasia: A statement of values for the future. In R. Chapey, (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 279–289). Baltimore, MD: Lippincott Williams & Wilkins.
- Lubben, J., Blozik, E., Gillmann, G., Iliffe, S., von Renteln Kruse, W., Beck, J. C., & Stuck, A. E. (2006). Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. *The Gerontologist*, 46, 503–513.
- Luterman, D. (1991). *Counseling the communicatively disordered and their families*. Austin, TX: Pro-Ed.
- Maas, E., Robin, D. A., Austermann Hula, S. N., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (Eds.), (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, 17, 277–298.
- McRoberts, C., Burlingame, G. M., & Hoag, M. J. (1998). Comparative efficacy of individual and group psychotherapy: A meta-analytic perspective. *Group Dynamics: Theory, Research, and Practice*, 2, 101–117.
- Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research*, 36(2), 338–350.
- Parr, S. (2007). Living with severe aphasia: Tracking social exclusion. *Aphasiology*, 21, 98–123.
- Pulvermüller, F., Neining, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, 32, 1621–1626.
- Routsalo, P. E., Tilvis, R. S., Kautiainen, H., & Pitkala, K. H. (2008). Effects of psychosocial group rehabilitation on social functioning, loneliness and well-being of lonely, older people: Randomized controlled trial. *Journal of Advanced Nursing*, 65(2), 297–305.
- Ryff, C. D., Singer, B. H., & Love, G. D. (2004). Positive health: Connecting well-being with biology. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 359, 1383–1394.
- Savage, M. C., Donovan, N. J., & Hoffman, P. R. (2014). Preliminary results from conversation therapy in two cases of aphasia. *Aphasiology*, 28, 616–636.
- Seeman, T. E. (2000). Health promoting effects of friends and family on health outcomes in older adults. *American Journal of Health Promotion*, 14, 362–370.
- Simmons-Mackie, N. (2008). Social approaches to aphasia intervention. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 290–318). Baltimore, MD: Lippincott Williams & Wilkins.
- Simmons-Mackie, N. (2018). *The state of aphasia in North America: A white paper*. Moorestown, NJ: Aphasia Access.
- Simmons-Mackie, N., Code, C., Armstrong, E., Stiegler, L., & Elman, R. J. (2002). What is aphasia? Results of an international survey. *Aphasiology*, 16(8), 837–848.
- Simmons-Mackie, N., Elman, R. J., Holland, A. L., & Damico, J. S. (2007). Management of discourse in group therapy for aphasia. *Topics in Language Disorders*, 27(1), 5–23.
- Simmons-Mackie, N., King, J., & Beukelman, D. (Eds.), (2013). *Supporting communication for adults with acute and chronic aphasia*. Baltimore, MD: Brookes.
- Simmons-Mackie, N., Savage, M. C., & Worrall, L. (2014). Conversation therapy for aphasia: A qualitative review of the literature. *International Journal of Language & Communication Disorders*, 49, 511–526.
- Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive Aphasia Test (CAT)*. London, United Kingdom: Psychology Press.
- Swinburn, K., Porter, G., & Howard, D. (2005). *The Comprehensive Aphasia Test*. Hove, United Kingdom: Psychology Press.
- Tang, W. K., Chan, S. S., Chiu, H. F., Ungvari, G. S., Wong, K. S., Kwok, T. C., . . . Ahuja, A. T. (2005). Poststroke depression in Chinese patients: Frequency, psychosocial, clinical, and radiological determinants. *Journal of Geriatric Psychiatry and Neurology*, 18(1), 45–51.
- The EuroQol Group. (1990). EuroQol—A new facility for the measurement of health-related quality of life. *Health Policy*, 16, 199–208.
- Thompson, C. K., & Shapiro, L. P. (2007). Complexity in treatment of syntactic deficits. *American Journal of Speech-Language Pathology*, 16(1), 30–42.
- Toseland, R. W., & Rivas, R. F. (2017). *An introduction to group work practice* (8th ed.). London, United Kingdom: Pearson Education Limited.
- van der Gaag, A., Smith, L., Davis, S., Moss, B., Cornelius, V., Laing, S., & Mowles, C. (2005). Therapy and support services for people with long-term stroke and aphasia and their relatives: A six-month follow-up study. *Clinical Rehabilitation*, 19, 372–380.
- Vickers, C. P. (2010). Social networks after the onset of aphasia: The impact of aphasia group attendance. *Aphasiology*, 24, 902–913.
- Walker, G. M., & Schwartz, M. F. (2012). Short-form Philadelphia Naming Test. *American Journal of Speech-Language Pathology*, 21, S140–S153.
- Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, 13, 70–77.
- Wertz, R. T., Collins, M. J., Weiss, D., Kurtzke, J. F., Friden, T., Brookshire, R. H., . . . Resurreccion, E. (1981). Veterans Administration cooperative study on aphasia: A comparison of individual and group treatment. *Journal of Speech and Hearing Disorders*, 24, 580–594.
- Wilson, R. S., Krueger, K. R., Arnold, S. E., Schneider, J. A., Kelly, J. F., Barnes, L. L., . . . Bennett, D. A. (2007). Loneliness and risk of Alzheimer disease. *Archives of General Psychiatry*, 64, 234–240.
- Yalom, I. D., & Leszcz, M. (2005). *The theory and practice of group psychotherapy* (5th ed.). New York, NY: Basic Books.

Appendix

Schedule of Conversation Topics

Week	Personal history	Session 1	Session 2
1	Personal history	Family	Where did you grow up?
2	Dining	Dining in/cooking	Dining out
3	Travel	Transportation	Sights in Boston/Philadelphia
4	News/events	Politics	Sports
5	Arts/entertainment	Books/films	Television
6	Personal history	Education/occupation/vocation	Hobbies/interests
7	Dining	Food that travels	Holiday food
8	Travel	Places in America	Places worldwide
9	News/events	Global natural disasters	Who inspires you?
10	Arts/entertainment	Art	Music