Improved Language in Chronic Aphasia After Self-delivered iPad Speech Therapy

Brielle C Stark

University of Cambridge, Department of Clinical Neurosciences

Addenbrookes Hospital, R3 Neurosciences Box 83, Hills Road, Cambridge, CB20QQ

+44 07450 021057

brielle.stark@gmail.com

Elizabeth A Warburton, DPhil, FRCP

University of Cambridge, Department of Clinical Neurosciences

Addenbrookes Hospital, R3 Neurosciences Box 83, Hills Road, Cambridge, CB20QQ

eaw23@medschl.cam.ac.uk

Published in: Neuropsychological Rehabilitation, February 2016, DOI:

10.1080/09602011.2016.1146150

Abstract

Self-delivered speech therapy provides an opportunity for individualized dose and a complement to speech-therapy regime in the long term rehabilitation pathway. Few apps for speech therapy have been subject to clinical trials, especially on a self-delivered platform. In a crossover design study, the Comprehensive Aphasia Test (CAT) and Cookie Theft Picture Description (CTPD) were used to measure untrained improvement in a group of chronic expressive aphasics after using a speech therapy app. A pilot study (n=3) and crossover design (n=7) comparing the therapy app with a non-language mind-game were conducted. Patients self-selected their training on the app, with a recommended dose of 20 minutes per day. There was significant post-therapy improvement on the CAT and CTPD but not significant improvement after mind-game intervention, suggesting language-specific effects following therapy app usage. Improvements on the CTPD, a functional measurement of speech, suggest that a therapy app can produce practical, important changes for speech. The improvements post-therapy were not due to type of language category trained or amount of training on the app, but an inverse relationship with severity at baseline and post-therapy improvement was shown. This study suggests that self-delivered therapy via an app is beneficial for chronic expressive aphasia.

Introduction

Recent estimates suggest that 33% of people suffering a stroke develop aphasia, oftentimes chronic, continuing to affect patients a year or more after their initial stroke. Aphasia has a significant impact on all aspects of the patients life, as well as that of their carers (Royal College of Physicians, 2012): a large survey of stroke survivors and their carers cited research into recovery from aphasia as one of the top priorities for the research community (Pollock, St George, Fenton, & Firkins, 2012).

Provision of outpatient speech therapy in the UK as well as USA for chronic aphasia is limited and variable across location (Code & Heron, 2003; Code & Petheram, 2011; Richard C Katz et al., 2000; Verna, Davidson, & Rose, 2009). It is largely accepted in the field that this small and infrequent amount of therapy does not support the evidence available, which suggests that intensive therapies (those that provided greater hours of therapy per week) is most effective (Bhogal, Teasell, & Speechley, 2003; Kelly, Brady, & Enderby, 2010).

Since the early 1980s, computerized therapy has been available as a tool for speech-language therapists in inpatient settings and, though less prevalent, as a means of patient-driven, at-home therapy. Neuropsychological research has investigated computerized therapy, showing as early as 1983 that computer-based treatment in acute and chronic aphasia was effective (Cherney, Halper, Holland, & Cole, 2008; Cherney, Halper, & Kaye, 2011; Doesborgh et al., 2004; R C Katz & Wertz, 1997; Richard C. Katz & Nagy, 1983, 1985; Richard C. Katz & Wertz, 1992; Lee, Fowler, Rodney, Cherney, & Small, 2010; Manheim, Halper, & Cherney, 2009; Marshall et al., 2013; Mortley, Wade, Enderby, & Hughes, 2004; Palmer et al., 2012; Palmer, Enderby, & Paterson, 2013; Varley, Windsor, & Whiteside, 2005; Wade, Mortley, & Enderby, 2003).

A recent systematic review of computerized therapy for aphasia by Zheng et al 2015 included seven studies ((Cherney, 2010); Doesborgh et al 2004; Katz & Wertz 1992, 1997; Palmer et al 2012; Thompson et al 2010 and Loverso et al 1992) and of those seven studies, none used entirely self-delivered therapy (Zheng, Lynch, & Taylor, 2015). A recent study, which sends therapist-specialized apps home with the patient for self-delivery, has shown some improvement on iPad-based tasks and variable improvements on standardized tests (Kiran, Roches, Balachandran, & Ascenso, 2014). However, no study has yet looked at an entirely patient-selected and self-delivered app outcome.

Self-delivery method holds great potential for the chronic aphasia patient looking to supplement their rehabilitation regime or continue therapy on their own, especially with the large amount of apps currently on the market (The Tavistock Trust for Aphasia, 2015). The self-delivery method, which also allows for self-selection of types of language training by the patient, is self-paced, enables patients to take an active role in their own treatment, helps achieve individualized dose, can be personalized to individual needs, and is user-friendly, inexpensive, and widely available (RCSLT, 2009). In a subacute cohort, pro-rata cost of providing treatment per hour per client for a computer therapy model was found to be approximately 30% cheaper compared to the standard service in Queensland, Australia (Wenke et al., 2014), and this figure would be arguably more economical for tablet-based and at home platforms.

Can self-delivery methods provide enough dose? While intensity of dose—providing greater hours of therapy in a short window—has been shown as important (Bhogal et al., 2003), there is no standard definition of 'intensity.' Cherney notes that the notion of "more is better" is not necessarily supported by the evidence and that the optimal 'intensive' dose will vary depending on the type of therapy, stimuli delivered and response requirement of the patient (Cherney, 2013). The characteristics of the patient, such as motivation, and environmental variables, complicate this further (Baker, 2012; Cherney, 2013). Self-delivery could be a powerful complement to a speech-language therapy regime that allows the patient to control their own dose, coupled with a recommendation from a speech therapist.

As software has become more advanced, speech-language therapists have become able to remotely monitor progress and increase the difficulty for each patient. Self-learning app progressions that keep difficulty levels in line with patient performance have also been designed (Kiran et al., 2014; Mortley, Davies, & Enderby, 2003). This ability to remotely monitor and machine-learn, coupled with self-delivery, provides an intriguing way to better distribute the rehabilitation pipeline resources.

While some have argued that the ageing population is not comfortable with computer use on their own, Rosemary Varley argues that this view underestimates the pervasiveness in society of computers and portable tablets across all age groups (Varley, 2011). Presently, the prevalence of tablet usage in the adult and older adult population has been steadily increasing since the advent of the tablet in 2010. In the USA, almost half (49%) of adults aged 35-44 now own a tablet computer, significantly more than any other age group, while older adults aged 65 and above have greatly increased their usage, now making up 18% of tablet ownership (PEW Research Center, 2014).

Currently, there are very few apps supported by clinical research for self-delivery efficacy. Aphasia Software Finder lists 61 speech-language therapist verified apps for aphasia use (The Tavistock Trust for Aphasia, 2015). Julie Sidock, a speech-language therapist, notes that evidence for app-based therapies are largely driven by expert opinion (Sidock, 2011). Jessica Snape and Brittany Maiolo, of Independent Living Centre WA, ranked apps across several categories from 4 to 1. 4, the best possible score, were apps that demonstrated sound scientific research foundations. Of 52 apps that they reviewed, nine apps received a score of 3 or better, four of which focused on aphasia rehabilitation in adults (Snape & Maiolo, 2013).

Self-delivered tablet-based speech therapy, should it prove effective, could combat what Katz and Code have cited as issues with resources in outpatient speech therapy for chronic aphasia by providing speech-specific, individualized dosage, self-directed therapy (Richard C Katz et al., 2000).

The aim of this study was to investigate the effectiveness and feasibility of self-delivered and directed iPad-based speech therapy in patients with chronic aphasia following a left MCA-territory stroke.

Methods and Design

Patients

200 patients were screened via the Addenbrookes Hospital stroke service. Patients were contacted if the screening indicated one-time left MCA-territory stroke, presence of aphasia, absence of pre-stroke neurodegenerative condition, first language competence in British English and at least one year post-stroke.

This study was interested in patients whose structural lesion anatomy, and whose language, would not spontaneously improve during the course of the study. This one-year cut-off for spontaneous change is supported in the research (Cherney & Robey, 2001; Koenig-Bruhin, Kolonko, At, Annoni, & Hunziker, 2013; Sarno, 1991), though some disagreement remains about the definition of spontaneous recovery. Because this study did not employ a double-baseline design to assess stability, by making one-year post-stroke the chronic inclusion parameter, any change observed during the study in neural function or behavioral language could then be attributable to the conditions applied.

25 potential patients were contacted and 16 patients were screened for cognitive impairments and performance on the Comprehensive Aphasia Test (Swinburn, Porter, & Al, 2004) and CTPD (Goodglass, Kaplan, & Barresi, 2000). Cognitive examinations included the Addenbrookes Cognitive Examination III (Neuroscience Research Australia (NeuRA), 2012), Apraxia Battery for Adults (van Heugten & Geusgens, 2006), cognitive portion of the CAT and Edinburgh Handedness Test (Oldfield, 1971).

This study was interested in expressive aphasia with intact comprehension. For this reason, patients were included if they scored above the aphasia cut-off on the Comprehension portion of the CAT. It is understood in the literature that patients with expressive aphasia show a heterogeneous phenotype, with components of intact speech alongside severely restricted components of speech. Therefore, patients were included if they scored beneath the aphasia cut-off on at least one of five subtests of the expressive CAT, showing a component of expressive aphasia impairment. Refer to Table 1.

As this study was interested in the functional measurement of language, patients were also scored on their content unit (CU) production and rate of speech on the Cookie Theft Picture adapted from the Boston Diagnostic Aphasia Examination 3, created by Harold Goodglass and Edith Kaplan (Goodglass et al., 2000). As a reference, included patients showed CUs and rate of speech similar to those shown by a large population of chronic severe-moderate to mild aphasics (Yorkston & Beukelman, 1980) and similar to those shown in a cohort of acute patients (Hillis et al, in press). These scores are shown in Table 2.

12 patients met inclusion parameters and were included in the study. The four who did not meet inclusion parameters showed poor comprehension on the CAT (n=1) or scored above aphasia cut offs on all subtests of the expressive aphasia component of the CAT (n=3).

Materials

An iPad was provided to all patients with the therapy app and mind-game pre-downloaded. A short 15-minute introduction to the iPad, including maintenance, charging and how to open the apps, was done by the experimenter with the patient.

An app by Tactus Therapy Solutions©, called Language Therapy, served as the therapy component within the crossover design. The company donated all copies of the app used in this study. Beyond that donation, there was no affiliation to disclose with the company. The app was created by a speech and language therapist and provided four categories for study: Reading, Naming, Comprehension and Writing. Within these four categories were several tasks that patients chose to complete. The type of training provided was both phonological and semantic in nature. For example, in the 'Naming: Describe' exercise, hierarchical cues provided both semantic and phonological cueing for successful picture naming. 'Reading: Fill in the Blanks' offered both semantically relevant and irrelevant and phonetically similar and dissimilar options for filling in words within given sentences. There were 700+ core nouns, verbs and adjectives throughout the app, and a choice of UK or US English.

The app provided feedback to the patient (correct, incorrect) and adapted difficulty as the user attained more correct answers. The app sent the users' data, including type of exercise used and number of errors within the exercise, via email to the experimenter every time the patient used the app, allowing for remote monitoring of compliance. The app was also customizable for the patient—they could enter their own pictures or words that they wanted to master, though this was not required.

Bejeweled© by PopCap, a spatial awareness and decision mind-game, was used as the control mind-game app to directly compare to Language Therapy©. The app interface is an 8 x 8 grid of gems of varying shapes. The goal is to swap gems with adjacent gems to make lines of three or more of the same gem. When this is done, the matched gems disappear, allowing more gems to fall into the board from above. The game ends when there are no more moves or time. The game progresses through difficulty levels when levels are completed. This mind-game was chosen because of its lack of language component, thus providing an active, attention-necessary comparison task not touching upon the cognitive component of language.

Study Design

This study, called "CATCHeS: Computerized Aphasia Therapy, Investigating Inner Speech" was approved by the NRES Committee East of England – Essex, study reference

13/EE/0382, from 16 December 2013. Addenbrookes Hospital was the sponsor. This study has been adopted by the NIHR Stroke Research Network portfolio.

The outcome measurements of this study were the expressive portion of the CAT and content units and rate of speech produced during the CTPD, often used as a functional measurement of speech (Prins & Bastiaanse, 2004; Williams et al., 2010; Yorkston & Beukelman, 1980). These are discussed in more detail in the results section.

10 patients completed all portions of the study; the two patients who only completed baseline measurements are not discussed. Three patients, age range 75-87 years (1 female, 2 males), time since stroke 12-19 months, could not complete the scanning portion of the study and took part only in before-and-after outcome measurements with the therapy.

A crossover design was conducted to directly compare the non-language mind-game and the therapy app. Seven patients, age range 54-71 years (3 females, 6 males) were allocated to use Bejeweled or Language Therapy for four weeks for 20 minutes per day, every day. Patients were recruited on a rolling basis over the course of a year and a half; as patients did not all enroll at the same time, and therefore the assignment to conditions could not be matched, they were pseudo-randomly assigned to conditions as they enrolled. The difference in groups is discussed in the first section of the results.

All patients were shown how to use the programs and given a ‘how to’ sheet to take home with them. They were told to self-select their therapy regime and were not instructed to use a specific subset of the Language Therapy app, as this study was interested in the self-delivery and self-chosen method. Upon completion of this first condition, patients returned to Addenbrookes Hospital for assessments. Patients then completed the contrasting condition with the same dosage parameters. Following this condition, patients returned to Addenbrookes Hospital for final assessments.

In the following results section, patients who completed Bejeweled as the first condition are called “Group 1” and those who completed therapy as the first condition are called “Group 2.” Refer to Table 2 for included patients and their group assignments and Figure 1 for visualization of this design.

Results

Final recruitment statistics for pilot and crossover designs (n=10) showed age range 54-87 years (3 females, 7 males) with time since stroke 12-67 months. These patient details can be found in Table 2.

Analysis

As the therapy training was self-selected, the primary outcome measurements were chosen as global measures of language quality. Therefore, the primary outcome measurements were the expressive portion of the CAT and measurements of CUs and rate of speech during the CTPD (as used by Katz 1997) (Richard C. Katz & Wertz, 1997). These measurements were collected at baseline, post-therapy and in the case of the crossover design, post-Bejeweled assessment.

Each subtest of the expressive CAT battery had different amount of subsections. In order to directly compare these subtests across time-points, and to give the data comparisons clinical significance, the raw scores were transformed as a proportion of the aphasia cut-off score for each subsection. This aphasia cut-off was a validated clinical score as specified by the CAT. Therefore, each patient received a proportional clinically-relevant score. For example, 80% score at baseline for the expressive battery indicated their score was 80% of the aphasia cut-off for the battery. Raw subtest scores could have been transformed into T-scores, which were provided in the CAT manual, but transformation in the proportion way provided standardization as well as a clinical understanding of the data.

The standardized scores for all subtests were averaged for baseline, post-therapy and post-Bejeweled conditions for each subject. The CUs and rate of speech at each time-point were also collected. 40% of CTPD samples were independently, blindly scored for CUs by three speech therapists and the main investigator of the study ($\alpha=0.912$) indicating a strong agreement rating.

These scores are shown in Table 2. The raw, pre-standardized scores at baseline of all patients can be found in the Table 1.

Shapiro-Wilk scores indicated CAT measures at baseline ($W=0.93$), post-therapy ($W=0.98$) and post-Bejeweled ($W=0.88$) showed normal distribution ($p>0.05$), and estimates of sphericity were normal (Mauchly's W was calculated, and resulting Chi-square test showed p value <0.01). Therefore, parametric tests were used for analysis of this outcome measurement. Nonparametric tests were used for all other comparisons and are specified in the text.

Comparison of Treatment Groups.

The members of each group (Group 1 $n=3$, Group 2 $n=4$) were pseudo-randomly assigned due to rolling recruitment, age (in years, at recruitment), time since stroke (in months, at recruitment) and years of formal education (in years, at recruitment). Mann Whitney U tests showed that age ($U=2$, $Z=1.24$, $p>0.05$), time since stroke ($U=1$, $Z=0.87$, $p>0.05$) and years of formal education ($U=2.5$, $Z=0$, $p>0.05$) were not significantly different between groups. Group 1 scored significantly more severe on the CAT than Group 2 ($U=0$, $Z=1.94$, $p=0.03$), though the two groups did not score significantly different on content units ($U=4.5$, $Z=0.36$, $p>0.05$) or rate of speech ($U=5$, $Z=0.18$, $p>0.05$) at baseline.

Effect of Therapy.

Figure 2 shows the language improvement proportional scores across subtests post-therapy for all study patients ($n=10$). A paired t-test was computed across all patients ($n=10$)

on baseline and post-therapy CAT scores (*t*=6.58, *p*=0.0001), showing a significant improvement on expressive CAT score total after therapy as compared to baseline condition. A strong negative correlation was shown between baseline standardized and proportion improvement on the CAT post-therapy ((baseline score-total possible score)/post-therapy score) (r=-0.92, p<0.01), showed that those scoring more severely on the expressive CAT at baseline made greater proportional improvement post-therapy.

There was a significant difference between post-therapy and baseline for content units (W=-55, Z=-2.78, p<0.01), where content units were greater post-therapy (Figure 8). Rate of speech approached significance (W=-37, Z=-1.86, p=0.06), showing quicker rate of speech post-therapy. Patients who scored most severe at baseline on either CUs or rate of speech did not show a strong correlation for making more improvement post-therapy.

Patients made improvements across four of the six subtests on the expressive CAT post-therapy. The total exercises used within the app did not best describe proportional post-therapy improvement (r=0.15). Patients tended to use portions of the app equally; usage data can be found in Table 3.

Effect of Session Timing.

An effect size measurement best for single-subject research studies, specifically in the aphasia literature, was used (Beeson & Robey, 2006). This effect size, based on the Cohen's d statistic, subtracted the mean of time 1 from the mean of time 2 and divided this by the standard deviation of time 1. This effect size calculation was then applied to all comparisons of interest within the crossover study. Cohen's d benchmarks are often cited (0.2=small, 0.5=medium, 0.8=large), though single-subject research studies have provided new benchmarks, from Robey's 1999 review of 12 studies (2.6=small, 3.9=medium, 5.8=large) (Robey, Schultz, Crawford, & Sinner, 1999). However, these benchmark effect sizes were provided for treated outcome measurements; untreated measurement benchmarks are still not verified (Beeson & Robey, 2006). For this reason, the original benchmarks of Cohen's d are cited below.

For the CAT score outcome, a large effect size post-therapy compared to baseline (d=1.07) and medium effect size post-Bejeweled compared to baseline (d=0.608) were shown. Post-therapy compared to post-Bejeweled showed a small effect size (d=0.258), indicating a clear effect of therapy on CAT scores for the group.

Group 1 showed a large effect size of post-therapy compared to baseline (d=1.155). A Wilcoxon Signed Rank test showed no significant different between post-Bejeweled scores and baseline scores ($Z=1.34$, $p>0.05$). Post-therapy compared to post-Bejeweled for this group produced a very large effect size (d=3.491), confirming the impact of language therapy on the CAT outcome measurement.

Group 2, who received language therapy as the first condition, showed a very large effect size for therapy compared to baseline (d=2.730) and for post-Bejeweled compared to baseline (d=2.120). Post-therapy compared to post-Bejeweled showed small effect size (d=0.228). These effect sizes further indicate that Bejeweled may be an adequate maintenance mechanism when presented after the therapy condition.

Content unit analysis showed a clear interaction of intervention (Therapy/Bejeweled) and course of study. All patients showed that post-therapy as compared to baseline produced a small-medium effect size (d=0.448) while post-Bejeweled as compared to baseline (d=0.365) and post-therapy compared to post-Bejeweled (d=0.073) showed small effect sizes. The course of study showed a similar effect size to that of therapy, showing that, by the end of the study, patients made gains in content units (d=0.464).

Compliance.

The app automatically captured usage information related to how many tasks were completed, what type of task was completed, and the percentage correct on the task.

However, the app did not automatically report the amount of time used. To investigate this area of compliance, an informal interview was conducted post-therapy condition to assess dosage compliance and to get feedback on the app.

This compliance data is reported in Table 3. 70% of patients had not used a tablet before, but only 1, patient AB, stated he was very comfortable using the device. All patients indicated that they used the app by themselves and this was confirmed by carer interviews (not included in Table 3). All patients said they used the app for at least the recommended dosage: 20 minutes per day, every day, for four weeks, but this could not be verified by the remote data sent by the app.

Alongside these self-report measures, compliance was measured remotely via automatic emails sent to the researcher by the therapy app. If the patient was connected to WiFi, every time they used the therapy app, an email comprising all completed tasks during that session would be delivered. 70% of patients had WiFi in their homes and were able to remotely send the data. The remote data suggested an average usage of 85.71 exercises over 4 weeks, equating to roughly 3 exercises per day. The exercise length ranged from a set of 30 stimuli through to a set of 100 stimuli on the most difficult setting, ranging anywhere from 5 – 15 minutes to complete. The length of time to complete a task varied, so it was not possible to measure the amount of time per exercise, or the total amount of time used during the four weeks.

As patients self-selected their training regime, it was of interest to determine their preference for types of training within the app. The usage of Reading (m=18.14), Writing (m=19.14), Naming (m=27.29) and Comprehension (m=19.71) showed no difference ($F(3,20)=0.06$, $p=0.98$), indicating an equal use of all parts of the app. There was no relationship between total exercises used and severity on CAT at baseline (r=-0.17) or with proportion change post-therapy (r=0.15).

Maintenance.

CTPD data from 5 patients (AD, PF, AB, GD and PB) at 6 months post-study was collected. To test maintenance from completion of the study, Time 3 (the last time patients were tested in the crossover study) and maintenance data were compared and to test maintenance since the therapy intervention, maintenance data was also compared to the post-therapy time-point for patients. For both content units ($W=-4$, $p>0.05$) and rate of speech ($W=-5$, $p>0.05$), there was no significant change from Time 3 to maintenance. For content units ($W=-5$, $p>0.05$) and rate of speech ($W=-5$, $p>0.05$), maintenance data was not significantly different than the post-therapy time-point. CU and rate of speech improvements made post-therapy and by the end of the crossover study were maintained at 6 months follow-up.

Discussion

The self-delivery method of speech therapy has great potential as a means to supplement the long-term rehabilitation pathway in chronic aphasia. The aim of this study was to investigate the effectiveness and feasibility of self-delivered iPad-based speech therapy in patients with chronic aphasia. There was significant improvement following therapy on measures of the CAT and an increase in content units and rate of speech during spontaneous speech acquired from the CTPD. Patients showing most severe CAT scores at baseline made the most improvement on the CAT post-therapy. The feasibility of the method was also validated. Though patients were older and 70% had no previous experience using a tablet, all patients self-reported that they utilized the therapy to at least prescribed dose, some for longer periods than 20 minutes. However, it was not possible to remotely monitor time used by the app data alone. Data provided remotely by the app suggested an average use of 3 exercises per day, arguably not an ‘intensive’ dose as previously supported by the research (Bhogal et al., 2003). However, as noted in the introduction, dose may be inextricably tied to individualized factors such as cognition, environment and motivation (Baker, 2012; Cherney, 2013), and this study provides insight that the self-delivery method may support the factors important in this broader, individualized concept of dose.

Effectiveness was measured by performance on a standardized test, the CAT, as well as on measures of spontaneous speech via the CTPD. Language Therapy ©, which trained semantic and phonological language with over 700 different stimuli, did not specifically train those items or subtests evaluated by the CAT or the spontaneous speech task. It is arguable that the CAT did indeed touch upon similar items or categories that were trained in the therapy (such as naming), but there was no explicit training to the CAT stimuli. Other single-case treatment studies have used components of and whole standardized batteries to validate intervention-based improvement (Doesborgh et al., 2004; Richard C. Katz & Wertz, 1997; Wertz et al., 1986) though improvements for untrained items are relatively rare and the evidence provided has shown vast variability between patients (Best et al., 2013; Mortley et al., 2004; Salter, Teasell, Foley, & Allen, 2013; Varley et al., 2005). This study provides evidence for a self-selected training program producing improvements across several subsets of language. This improvement across subtests did not appear directly related to the amount of training or the types of exercises used during the language therapy.

It is conceivable that the outcome measure improvements may be influenced by repeated exposure to the same outcome measurement. The specific crossover design was implemented for this purpose, to be able to confidently conclude that, for instance, patients who received Bejeweled as the first intervention did not make significant gains on the CAT. A caveat is that the usage of Bejeweled as a condition could not be monitored to the extent that Language Therapy usage was monitored; therefore, it is impossible to say whether patients used Bejeweled every day except on subjective assurance.

It is important to note that improvements to spontaneous speech (increased CUs and rate of speech) were due to both session (therapy/Bejeweled) and course of the study (time 3), and that the effect size of the therapy intervention and course of study were equal. Significantly, there was a maintenance effect shown for spontaneous speech, where improvements made post-therapy and by the end of the study on content units and rate of speech were maintained at 6 months follow-up.

This study shows that an entirely self-delivered semantic and phonological language training app shows improvement in expressive chronic aphasia, supporting its potential role in the long-term rehabilitation pathway. Further, an inverse relationship between severity and proportion improvement was shown. This result could mean that the therapy is best for more severe patients and not difficult enough to elicit improvement in mild patients; or, that the CAT or CTPD was not sufficiently comprehensive or difficult enough to pick up any changes in the mild patients. It is perhaps the case that an inclusive app (like the app used in this study), using semantic and phonological training across several subsets of language, where patients self-select their exercises, is more beneficial for achieving individualized dose in severe expressive aphasia, while a more tailored and challenging language-based app experience will be more beneficial for moderate-mild expressive aphasia. This idea of user-specific/tailored experience on a mostly self-delivered platform is currently being explored in the literature (Kiran et al., 2014) and future research with larger patient numbers must verify whether tablet-delivered therapy should or should not be individually tailored based on the phenotype and severity at baseline.

In conclusion, this proof-of-concept study provides evidence for improvement following entirely self-delivered iPad-based speech therapy in the chronic population and support for further investigation of individualized factors in the understanding of dose effectiveness. This study drives further research with larger number of patients and with functional language outcome measurements to explore the potential for widespread use of tablet-based speech therapy in all classifications of chronic aphasia and for understanding the therapy type and specificity necessary for patients with varying degrees of aphasia severity. With over sixty apps in a recently compiled online aphasia resource still lacking clinical evidence, this is an exciting area of study with great potential to supplement gaps in the long-term rehabilitation pipeline and provide superior, supportive technology for thousands (Snape & Maiolo, 2013; The Tavistock Trust for Aphasia, 2015).

Baker, E. (2012). Optimal intervention intensity. *International Journal of Speech-Language Pathology*, *14*(5), 401–409. doi:10.3109/17549507.2012.700323

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychological Review*, *16*(4), 161–169. doi:10.1007/s11065-006-9013-7.Evaluating

Best, W., Greenwood, A., Grassly, J., Herbert, R., Hickin, J., & Howard, D. (2013). Aphasia rehabilitation: does generalisation from anomia therapy occur and is it predictable? A case series study. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(9), 2345–57. doi:10.1016/j.cortex.2013.01.005

Bhogal, S. K., Teasell, R., & Speechley, M. (2003). Intensity of aphasia therapy, impact on recovery. *Stroke; a Journal of Cerebral Circulation*, *34*(4), 987–93. doi:10.1161/01.STR.0000062343.64383.D0

Cherney, L. R. (2010). Oral reading for language in aphasia (ORLA): evaluating the efficacy of computer-delivered therapy in chronic nonfluent aphasia. *Topics in Stroke Rehabilitation*, *17*(6), 423–431. doi:10.1310/tsr1706-423

Cherney, L. R. (2013). Aphasia treatment: Intensity, dose parameters and script training. *Journal of Speech and Language Pathology*, *14*(5), 424–431.

doi:10.3109/17549507.2012.686629.Aphasia

Cherney, L. R., Halper, A. S., Holland, A. L., & Cole, R. (2008). Computerized script training for aphasia: preliminary results. *American Journal of Speechlanguage Pathology American SpeechLanguageHearing Association*, *17*(1), 19–34. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2896889&tool=pmcentrez&rendertype=abstract

Cherney, L. R., Halper, A. S., & Kaye, R. C. (2011). Computer-based script training for aphasia: Emerging themes from post-treatment interviews. *Journal of Communication Disorders*, *44*(4), 493–501. doi:10.1016/j.jcomdis.2011.04.002

Cherney, L. R., & Robey, R. (2001). Aphasia treatment: recovery, prognosis, and clinical effectiveness. In *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders*. Philadelphia, Pa: Lippincott Williams & Wilkins.

Code, C., & Heron, C. (2003). Services for aphasia, other acquired adult neurogenic communication and swallowing disorders in the United Kingdom, 2000. *Disability and Rehabilitation*, *25*(21), 1231–7. doi:10.1080/09638280310001599961

Code, C., & Petheram, B. (2011). Delivering for aphasia. *International Journal of Speechlanguage Pathology*, *13*(1), 3–10. doi:10.3109/17549507.2010.520090

Doesborgh, S., van de Sandt-Koenderman, M., Dippel, D., van Harskamp, F., Koudstaal, P., & Visch-Brink, E. (2004). Cues on request: The efficacy of Multicue, a computer program for wordfinding therapy. *Aphasiology*, *18*(3), 213–222. doi:10.1080/02687030344000580

Goodglass, H., Kaplan, E., & Barresi, B. (2000). *Boston Diagnostic Aphasia Exmination* (3rd ed.). Pearson.

Katz, R. C., Hallowell, B., Code, C., Armstrong, E., Roberts, P., Pound, C., & Katz, L.

(2000). Notes and Discussion A multinational comparison of aphasia management practices. *International Journal of Language and Communication Disorders*, *35*(2), 303–314.

Katz, R. C., & Nagy, V. T. (1983). A Computerized Approach for Improving Word Recognition in Chronic Aphasic Patients. *Aphasiology*, 65–72.

Katz, R. C., & Nagy, V. T. (1985). A Self-Modifying Computerized Reading Program for Severely-Impaired Aphasic Adults. *Aphasiology*, 184–188.

Katz, R. C., & Wertz, R. T. (1992). Computerized hierarchical reading treatment in aphasia. *Aphasiology*, *6*(2), 165–177. doi:10.1080/02687039208248588

Katz, R. C., & Wertz, R. T. (1997). The Efficacy of Computer-Provided Reading Treatment for Chronic Aphasic Adults. *Journal of Speech, Language and Hearing Research*, *40*, 493–507.

Katz, R. C., & Wertz, R. T. (1997). The efficacy of computer-provided reading treatment for chronic aphasic adults. *Journal of Speech, Language, and Hearing Research : JSLHR*, *40*(3), 493–507. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9210109

Kelly, H., Brady, M., & Enderby, P. (2010). Speech and language therapy for aphasia following stroke (Review). *The Cochrane Collaboration*, (7), 166.

Kiran, S., Roches, C. Des, Balachandran, I., & Ascenso, E. (2014). Development of an impairment-based individualized treatment workflow using an iPad-based software platform. *Seminars in Speech and Language*, *35*(212), 38–50. doi:10.1055/s-0033-1362995

Koenig-Bruhin, M., Kolonko, B., At, A., Annoni, J.-M., & Hunziker, E. (2013). Aphasia following a stroke: recovery and recommendations for rehabilitation. *Swiss Archives of Neurology and Psychiatry*, *164*(8), 292–8.

Lee, J., Fowler, R., Rodney, D., Cherney, L., & Small, S. L. (2010). IMITATE: An intensive computer-based treatment for aphasia based on action observation and imitation. *Aphasiology*, *24*(4), 449–465. doi:10.1080/02687030802714157

Manheim, L. M., Halper, A. S., & Cherney, L. (2009). Patient-reported changes in communication after computer-based script training for aphasia. *Archives of Physical Medicine and Rehabilitation*, *90*(4), 623–7. doi:10.1016/j.apmr.2008.10.022

Marshall, J., Roper, A., Galliers, J., Wilson, S., Cocks, N., Muscroft, S., & Pring, T. (2013). Computer delivery of gesture therapy for people with severe aphasia. *Aphasiology*, (May), 1–19. doi:10.1080/02687038.2013.786803

Mortley, J., Davies, A., & Enderby, P. (2003). Investigating the feasibility of telemedicine in aphasia therapy. *Advances in Speech Language Pathology*, *5*(1), 27–36.

Mortley, J., Wade, J., Enderby, P., & Hughes, A. (2004). Effectiveness of computerised rehabilitation for long-term aphasia: a case series study. *The British Journal of General Practice the Journal of the Royal College of General Practitioners*, *54*(508), 856–857. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1324921&tool=pmcentrez&rendertype=abstract

Neuroscience Research Australia (NeuRA). (2012). *Addenbrookes Cognitive Test III*.

Oldfield, R. (1971). The Assessment and Analysis of Handedness: The Edinburgh Inventory. *Neuropsyychologia*, *9*, 97–113.

Palmer, R., Enderby, P., Cooper, C., Latimer, N., Julious, S., Paterson, G., … Hughes, H. (2012). Computer therapy compared with usual care for people with long-standing aphasia poststroke: a pilot randomized controlled trial. *Stroke; a Journal of Cerebral Circulation*, *43*(7), 1904–11. doi:10.1161/STROKEAHA.112.650671

Palmer, R., Enderby, P., & Paterson, G. (2013). Using computers to enable self-management of aphasia therapy exercises for word finding: the patient and carer perspective.

International Journal of Language & Communication Disorders, n/a–n/a. doi:10.1111/1460-6984.12024

PEW Research Center. (2014). *PEW Research Center's Internet & American Life Project: E-reading and Gadgets*. Retrieved from http://www.pewinternet.org/datasets/january-2014-e-reading-and-gadgets-omnibus/

Pollock, A., St George, B., Fenton, M., & Firkins, L. (2012). Top ten research priorities relating to life after stroke. *Lancet Neurology*, *11*(3), 209. doi:10.1016/S1474-4422(12)70029-7

Prins, R., & Bastiaanse, R. (2004). Review: Analysing the spontaneous speech of aphasic speakers. *Aphasiology*, *18*(12), 1075–1091. doi:10.1080/02687030444000534

RCSLT. (2009). *RCSLT Resource Manual For Commissioning And Planning Services For SLCN. Royal College of Speech & Language Therapist* (Vol. 2009).

Robey, R., Schultz, M., Crawford, A., & Sinner, C. (1999). Single-subject clinical-outcome research: designs, data, effect sizes, and analyses. *Aphasiology*, *13*, 445–473.

Royal College of Physicians. (2012). *National clinical guideline for stroke*.

Salter, K., Teasell, R., Foley, N., & Allen, L. (2013). Aphasia. In *Evidence-Based Review of Stroke Rehabilitation* (pp. 1–83). Retrieved from www.ebrsr.com

Sarno, M. (1991). Recovery and rehabilitation in aphasia. In *Acquired Aphasia* (pp. 595–631). San Diego, CA: Academic Press.

Sidock, J. (2011). Critical Review : Is the integration of mobile device apps ' into speech and language therapy effective clinical practice ? This critical review examines literature currently available on the implications of integrating specific applications (Apps ') for.

Snape, J., & Maiolo, B. (2013). Using iPads in Speech Pathology, (April), 1–135.

Swinburn, K., Porter, G., & Al, E. (2004). *Comprehensive Aphasia Test*. East Sussex: Psychology Press.

The Tavistock Trust for Aphasia. (2015). *Aphasia Software Finder*.

van Heugten, C., & Geusgens, C. (2006). Apraxia Test.

Varley, R. (2011). Rethinking aphasia therapy: a neuroscience perspective. *International Journal of Speechlanguage Pathology*, *13*(1), 11–20. doi:10.3109/17549507.2010.497561

Varley, R., Windsor, F., & Whiteside, S. (2005). Whole word therapy for acquired apraxia of speech. In *Clinical Aphasiology Conference* (pp. 1–9). Sanibel Island, FL.

Verna, A., Davidson, B., & Rose, T. (2009). Speech-language pathology services for people with aphasia: A survey of current practice in Australia. *International Journal of Speech-Language Pathology*, *11*(3), 191–205.

Wade, J., Mortley, J., & Enderby, P. (2003). "Talk about IT": Views of people with aphasia and their partners on receiving remotely monitored computer-based word finding therapy. *Aphasiology*, *17*(11), 1031–1056.

Wenke, R., Lawrie, M., Hobson, T., Romano, M., Comben, W., Ward, E., & Cardell, E. (2014). Feasibility and Cost Analysis of Implementing High Intensity Aphasia Clinics Within a Subacute Setting. *High Intensity Aphasia Clinics*, 1–33.

Wertz, R., Weiss, W., Aten, J., Brookshire, R., Garcia-Bunuel, L., Holland, A., & Al, E. (1986). Comparison of clinic, home and deferred language treatment. *Archives of Neurology*, *43*, 653–658.

Williams, C., Thwaites, A., Buttery, P., Geertzen, J., Randall, B., Shafto, M., … Tyler, L. (2010). The Cambridge Cookie-Theft Corpus : A Corpus of Directed and Spontaneous Speech of Brain-Damaged Patients and Healthy Individuals.

Yorkston, K. M., & Beukelman, D. R. (1980). An Analysis of Connected Speech Samples of Aphasic and Normal Speakers. *Journal of Speech and Hearing Disorders*, *45*, 27–36. doi:10.1044/jshd.4501.27

Zheng, C., Lynch, L., & Taylor, N. (2015). Effect of computer therapy in aphasia: a systematic review. *Aphasiology*, *7038*(November), 1–34. doi:10.1080/02687038.2014.996521

T:\Online2PDF\Figure1.docx

T:\Online2PDF\Figure2.docx

T:\Online2PDF\Table1.docx

T:\Online2PDF\Table2.docx

T:\Online2PDF\Table3.docx

Tables

Table 1: Raw scores on the expressive portion of the CAT at baseline (post-therapy in parentheses)

CAT Sub-Tests	SG	DH	DE	AD	NP	PF	AB	BM	GD	PB	Cut Off
Repetition	31 (57)	17 (21)	44 (63)	44 (69)	69 (72)	42 (54)	65 (72)	66 (72)	52 (67)	60 (66)	67
Naming	19 (57)	20 (30)	41 (51)	69 (76)	76 (71)	31 (60)	69 (72)	71 (86)	71 (84)	74 (84)	69
Reading	0 (31)	6 (28)	26 (43)	39 (62)	66 (70)	17 (29)	57 (64)	64 (66)	66 (66)	67 (67)	58
Spoken Pic Descrip	10 (24)	11.5 (39)	1 (4)	16 (47)	27.5 (47)	3 (9)	9 (26)	53 (90)	22 (33)	38 (50)	33
Writing	69 (75)	46 (52)	47 (50)	75 (73)	76 (76)	59 (60)	57 (65)	74 (76)	74 (74)	76 (76)	66
Written Pic Descrip	6 (18)	0 (0)	-2 (0)	15 (51)	15 (20)	9 (6)	3 (4)	44 (106)	22 (43)	30 (76)	19

Table 2: Included participants who undertook entire study, baseline scores; CAT scores expressed as a proportion of aphasia cut-off score

Patient	Age	Sex	Time Since Stroke (months)	CAT Compre-hension	CAT Expres-sive	Content Units at Baseline	Rate of Speech(CUs/min)	Study Portion
SG	87	F	13	1.06	0.43	6	6	Pilot
DH	78	M	12	1.01	0.32	1	0.5	Pilot
DE	75	M	19	1.00	0.50	10	20	Pilot
AD	71	M	79	1.08	0.83	13	30	Crossover
PF	55	M	49	1.08	0.84	3	1.6	Crossover
NP	54	M	48	1.03	1.06	10	15.6	Crossover
AB	61	M	38	1.00	0.52	6	3.6	Crossover
BM	63	F	72	1.13	1.20	32	24.6	Crossover
GD	47	M	12	1.06	0.99	12	4.8	Crossover
PB	45	F	20	1.14	1.11	13	17.4	Crossover

Table 3: All Patients Baseline and Post-Therapy Proportional Scores

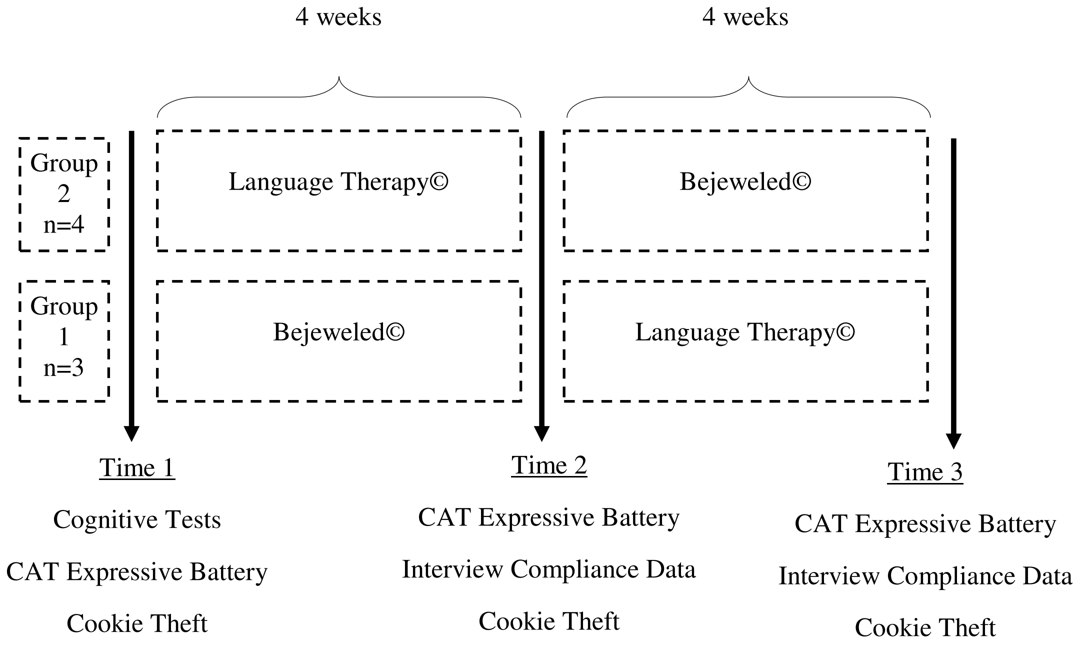
		Post-Therapy			Post-Bejeweled		
	CAT Baseline	Expressive CAT	CUs	Rate of Speech	Expressive CAT	CUs	Rate of Speech
SG	0.43	0.84	9	9	x	x	x
DH	0.32	0.54	4	2.67	x	x	x
DE	0.50	0.68	12	12	x	x	x
AD	0.83	1.21	15	30	0.83	12	24
AB	0.84	0.97	4	2.2	0.85	6	2.43
NP	1.06	1.14	15	24.6	1.07	16	30
PF	0.52	0.70	11	7.3	0.88	7	4.42

BM	1.20	1.59	38	24	1.58	32	20.87
GD	0.99	1.18	14	13.8	1.11	14	9.55
PB	1.11	1.34	19	19.2	1.35	24	26.67

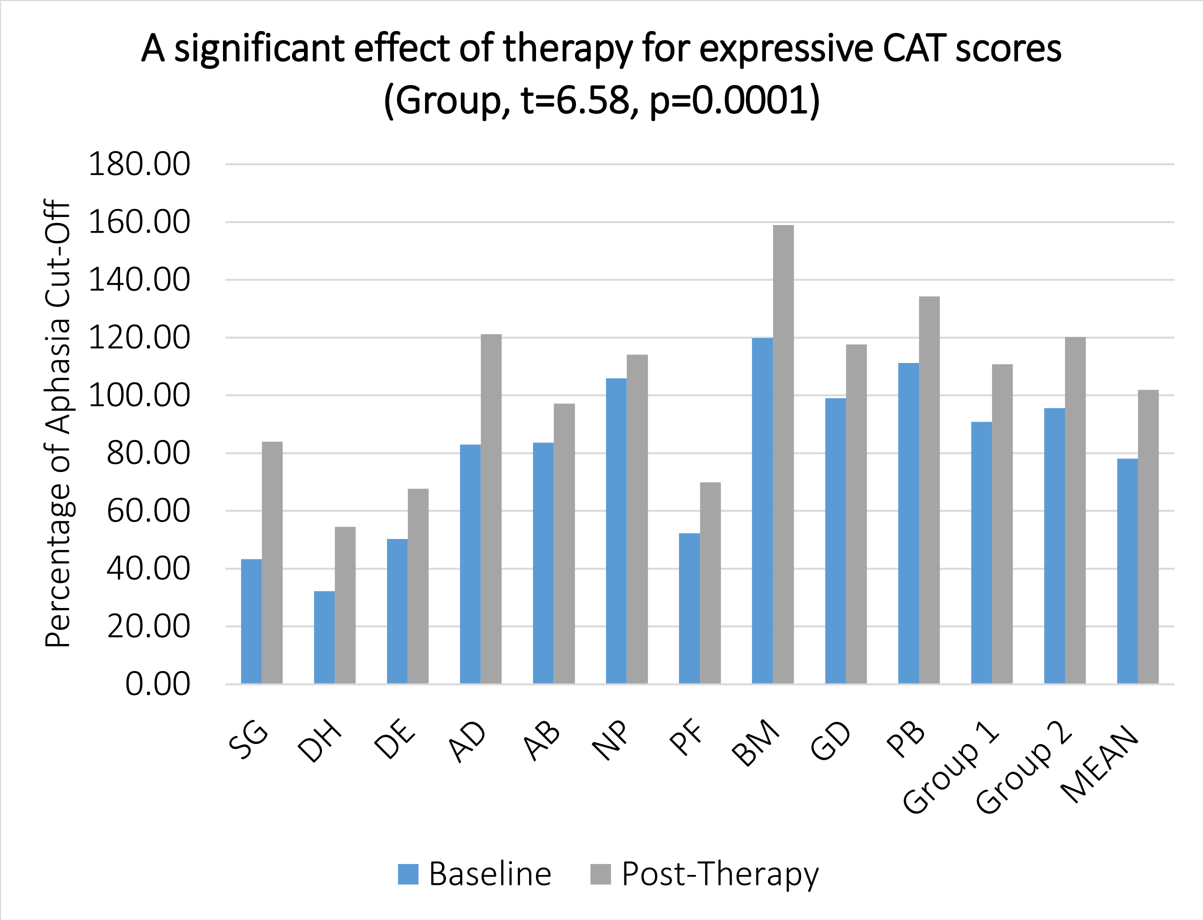
Table 4: Compliance data

		Interview Questions		**Remote Feedback Data – Number of Exercises in 4 Weeks**				
		Had used iPad	*Completed Dosage Without Help*	*Reading*	*Writing*	*Naming*	*Compre-hension*	*Total*
Pilot Study	SG	Yes	Yes	No Data				
	DH	No	Yes	No Data				
	DE	No	Yes; some help	17	18	76	7	**128**
Crossover Study	AD	No	Yes	23	19	10	20	**72**
	AB	Yes	Yes	0	0	38	0	**38**
	NP	No	Yes	14	6	9	6	**35**
	PF	Yes; windows	Yes	20	31	7	16	**74**
	BM	No	Yes	No Data				
	GD	No	Yes	26	33	33	64	**156**
	PB	No	Yes	27	27	18	25	**97**

Figures



Visualization of crossover design



Proportional Subtest Scores by Patient Baseline and Post-Therapy