Technology Enhanced Writing Therapy for People with Aphasia:

Results of a Quasi-Randomised Waitlist Controlled Study

Jane Marshall[1]

Anna Caute[1]

Katie Chadd[1]

Madeline Cruice[1]

Katie Monnelly[1]

Stephanie Wilson[2]

Celia Woolf[1]

1 Division of Language and Communication Science, City, University of London

2 Centre for Human Computer Interaction Design, City, University of London

Abstract

Background: Acquired writing impairment, or dysgraphia, is common in aphasia. It affects both handwriting and typing, and may recover less well than other aphasic symptoms. Dysgraphia is an increasing priority for intervention, particularly for those wishing to participate in online written communication. Effective dysgraphia treatment studies have been reported, but many did not target, or did not achieve, improvements in functional writing. Functional outcomes might be promoted by therapies that exploit digital technologies, such as voice recognition and word prediction software.

Aims: This study evaluated the benefits of technology enhanced writing therapy for people with acquired dysgraphia. It aimed to explore the impact of therapy on a functional writing activity, and to examine whether treatment remediated or compensated for the writing impairment. The primary question was: Does therapy improve performance on a functional assessment of writing; and, if so, do gains occur only when writing is assisted by technology? Secondary measures examined whether therapy improved unassisted written naming, functional communication, mood and quality of life.

Methods & Procedures: The study employed a quasi randomised waitlist controlled design. 21 people with dysgraphia received 12 hours of writing therapy, either immediately, or after a 6 week delay. The primary outcome measure was a functional assessment of writing, which was administered in handwriting and on a computer with assistive technology enabled. Secondary measures were: The Boston Naming Test (written version), Communication Activities of Daily Living - 2, Visual Analogue Mood Scales (*Sad* question), and the Assessment of Living with Aphasia. ANOVA analyses were used to examine change on the outcome measures over two time points, between which the immediate group had received therapy, but the delayed group had not. Pre therapy, post therapy and follow up scores on the measures were also examined for all participants.

Outcomes & Results: Time x group interactions in the ANOVA analyses showed that therapy improved performance on the functional writing assessment. Further interactions with condition showed that gains occurred only when writing was assisted by technology. There were no significant interactions in the analyses of the secondary outcome measures. A treatment effect on these measures was therefore unconfirmed.

Conclusions & Implications. This study showed that 21 people with dysgraphia improved on a functional writing measure following therapy using assistive technology. The results suggest that treatment compensated for, rather than remediated the impairment, given that unassisted writing did not change. Further studies of technology enhanced writing therapy are warranted.

What this paper adds

What is known already: Writing abilities are typically impaired in aphasia, and may recover less well than other language modalities. Many previous writing therapy studies did not achieve functional gains on everyday writing tasks.

What this study adds: This study shows that mainstream digital technologies, such as speech to text software, can be used in therapy to help compensate for writing impairments. Gains were shown on a functional task (writing emails) after 12 hours of treatment.

Clinical implications: With specific training, people with aphasia can learn to use mainstream technologies in order to support writing. Greater use of such technologies could be made in practice.

Introduction

Writing impairments (dysgraphia) are a common symptom of aphasia. Indeed, for many individuals writing is the most impaired language modality (e.g. see Panton & Marshall, 2008; Bruce et al, 2003). Yet, the rehabilitation of writing in aphasia typically receives less clinical attention than the rehabilitation of speech (Papathanasiou & Csefalvay, 2017; Beeson & Rapcsak, 2015). If this is the case, arguably, it needs to change, given the growing importance of written internet communication. Loss of writing is a barrier to many online activities, such as buying goods and services, conducting personal finances, and using email and social media to retain contact with family and friends (see arguments in Menger et al, 2015). Indeed, it could be argued that acquired dysgraphia threatens social participation more than ever before, particularly if supported internet use is not an option.

Acquired writing impairments vary both in severity and type. Thus, some individuals can barely write, while others struggle with particular word types or make frequent spelling errors (see examples in Papathanasiou & Csefalvay, 2017). Patterns of impairment are thought to reflect the nature of the underlying spelling mechanism (see arguments in Beeson & Rapcsak, 2015). This mechanism involves two spelling routes. One retrieves whole word spellings, stored in an orthographic lexicon. The other assembles spellings using regular sound to letter correspondences (Rapp & Fischer-Baum, 2015). Dysgraphic impairments can reflect failures to either (or both) of these routes. Further patterns of breakdown can arise from an inability to store spellings long enough for them to be written (Caramazza et al, 1987) or, in the case of handwriting, to access and execute letter shapes (Patterson & Wing, 1989).

Aphasic writing impairments have responded to a range of treatments, most of which have involved hand writing tasks. Many studies aimed to improve the production of single written words, for example using copying, delayed copying and anagram sorting tasks (e.g. Beeson et al, 2002; Beeson et al, 2013) or written picture naming (e.g. Robson et al, 2001; Ball et al, 2011). Treatments have also targeted sound to letter conversion (e.g. Kiran et al, 2005; Luzzatti et al, 2000). Moving beyond single words, there have been attempts to improve written sentence production (e.g. Murray et al, 2007; Salis & Edwards, 2010) and to remediate functional writing activities. For example, studies have targeted note taking skills (Panton & Marshall, 2008), and have promoted the use of trained written words in conversation (Beeson et al, 2003; Clausen & Beeson, 2003) or in messages (Robson et al, 2001).

A systematic review of the writing therapy literature identified several weaknesses in the evidence base (Thiel et al, 2015). Most of the reviewed studies used single case or case series designs, with very few group studies and no randomised controlled trials. The range of treatment outcomes was also restricted. Many studies demonstrated improved writing of single words (e.g. Ball et al, 2011, Raymer et al, 2003). Far fewer showed benefits for written discourse or functional writing activities (although see Beeson et al, 2000; Greenwald, 2004; Panton & Marshall, 2008). In part, this reflected a problem of measurement. The authors of the review make the point that there is no standard assessment of functional writing skills in aphasia. As a result, previous studies have employed informal tasks such as writing essays (Beeson et al, 2000) and notes (Panton and Marshall, 2008).

A recent addition to the therapy literature explored writing treatment outcomes with a small group of participants (N=8) and examined gains beyond single word spelling tests (Thiel et al, 2016). Each participant worked on the spelling of 80 words, using tasks that required them to copy and recall the words, and select the words from given written options. Outcome measures assessed spelling of the treated and untreated words and the number of correctly spelt words in picture description and email composition tasks. Participants were also asked to record how often they undertook everyday writing activities in a diary, and completed the Comprehensive Aphasia Test Disability Questionnaire (Swinburn et al, 2004). Results showed that all participants improved significantly in the spelling of treated words, with 6 showing generalised gains to untreated words. There were also significant improvements across the group in the picture description task. However the email task, diary records of everyday writing and the group Disability Questionnaire scores were unchanged. Thus, there were clear and encouraging therapy benefits, but little evidence that these affected functional uses of writing.

Digital technology offers exciting opportunities for the rehabilitation of writing. Such technologies can be employed to deliver writing exercises, with the aim of remediating specific skills such as access to word forms (Laganaro et al, 2006). Technology can also be used to compensate for aphasic writing disorders. For example, studies have shown that word processing and word prediction software can augment the effects of therapy, with benefits for at least some of those involved (Armstrong & MacDonald, 2000; Behrns et al 2009; Mortley et al, 2001; Murray & Karcher 2000).

Thiel, Sage and Conroy (2017) delivered a writing therapy programme to 8 individuals with dysgraphia using Co:Writer®. This assistive technology incorporates word and grammar prediction, word banks, spell checking and a text to speech facility. Following training in the use of the software, participants carried out a hierarchy of functional writing tasks using Co:Writer. An email writing evaluation measure showed significant treatment gains across the group in the number of correctly spelt words. There were also qualitative changes in email content, with longer words produced after therapy than before and, for some individuals, more content words. Gains were only evident when tasks could be performed with the software. Unassisted writing did not change.

Voice recognition software (VRS), which translates speech to text, offers a similar potential. Wade, Petherham and Cain (2001) produced early evidence that such software could recognise aphasic speech. Bruce, Edmundson and Coleman (2003) pursued this observation into therapy. They worked with MG, who had mild anomic aphasia but a much more severe writing impairment. Previous 'conventional' therapy had improved his writing of single words, but not sentences. MG was given 17 hours of training in the use of Dragon Naturally Speaking™, a voice recognition software. Training focussed on mastery of the software, editing and dictation skills. Composition skills, e.g. to help MG produce a coherent email, were also targeted. Although not an experimental study, a picture description task showed a striking improvement in the amount and quality of MG's writing following this training. This was assessed before therapy in hand writing, but after therapy with the trained technology. Perhaps more importantly, MG reported regular independent use of the software, giving examples such as cancelling standing orders and writing emails to his children.

Two further single case studies have explored applications of VRS in aphasia. Estes and Bloom (2011) were successful in training CH to use the software to improve the quality of her writing, following 10 hours of therapy. However, a subsequent episode of intervention, aiming to generalise skills to the use of email and the internet was less successful. Caute and Woolf (2016) worked with 'Stephen', who had fluent aphasia and a severe dysgraphia. Stephen was given 16 hours of therapy which covered training in the use of Dragon VRS, editing and self-monitoring skills. Homework tasks also promoted independent and functional use of the software. The study employed a single case experimental design. The primary outcome measure was a task in which Stephen was required to write three emails of varying complexity. Gains on this task were assessed by the number of words produced and quality judgements from blinded raters. Lexical scores were not statistically evaluated but were nevertheless impressive. Across all pre therapy administrations of the task Stephen never produced more than 6 words in an email; whereas in the post therapy administrations his scores ranged from 61 – 151 words. The judgement scores were statistically evaluated and showed a significant change from pre to post therapy. There was also evidence of benefit on secondary measures of social participation, on which Stephen reported more social activity after therapy, and an expansion in his social network (neither gain was statistically evaluated). Perhaps most impressively, Stephen achieved his main therapy goal, which was the resumption of voluntary work. He commented that this was made possible by his new writing skills.

To date, the evidence suggests that assistive technologies, for example offering word prediction and VRS, have the potential to improve functional writing skills in aphasia. Indeed, the results with some individuals point to substantial benefits, not only for writing but for wider social wellbeing. However, few studies have tested this hypothesis, and none has worked with more than 8 participants. There are also limited findings on the nature of therapy induced change. Caute and Woolf (2016) were interested in whether therapy using assistive technology might have an impact on Stephen's writing impairment. Before therapy, Stephen was virtually unable to write, suggesting that he had very limited access to written word forms. Although this was not extensively tested, he also seemed unable to use sound to letter conversion. The lack of change in unassisted handwriting indicated that these underlying impairments remained post therapy. Theil et al (2017) similarly found that their technology enhanced therapy only improved writing performance when the technology was available. Therefore, they also concluded that the writing impairments of their participants were untouched, and that treatment effects were compensatory. Testing with more participants might further support this hypothesis. On the other hand, evidence of improvement in unassisted writing would suggest that therapy has remediated, as well as compensated for the writing impairment.

The study reported in this paper evaluated a technology enhanced writing therapy administered to a group of 21 people with aphasia. The study extends the evidence base by employing a quasi-randomised design, comparing a group that received immediate therapy with a waitlist control group. The primary outcome measure examined benefits on a functional writing task, both with and without the technology. This enabled us to explore whether any benefits were compensatory (evident only when technology was available) or remediative (evident also in handwriting). A range of secondary measures probed for improvements in written naming, functional communication, mood and quality of life. The naming measure further explored the extent of writing gains, and whether these generalised to unassisted writing. As occurred with Stephen (Caute & Woolf, 2016), it was hoped that improved writing would give access to increased opportunities for communication and social participation. Gains in functional communication, mood and quality of life were therefore hypothesised.

The study addressed the following questions:

1. Does technology-enhanced writing therapy improve performance on a functional assessment of writing? Are gains confined to the technology assisted version of the task, or evident also in the handwritten version?
2. Does therapy improve unassisted written naming?
3. Does therapy improve functional communication, as assessed by the Communication Activities in Daily Living -2 assessment (CADL-2, Holland et al, 1999)?
4. Does therapy improve mood, as assessed by the *Sad* question on the Visual Analogue Mood Scales (Kontou et al, 2012)?
5. Does therapy improve quality of life as assessed by the Assessment of Living with Aphasia (Simmons-Mackie et al, 2014)?

The primary outcome measure used in this study was developed from that employed by Caute and Woolf (2016). As this is a novel assessment we also present inter-rater and test/re-test reliability data on the measure.

Further qualitative data were collected to explore participants' experiences of and views about the intervention, which will be reported in a future paper. These were drawn from investigations of participants' technology use and post-therapy semi-structured interviews.

Method

This study formed one wing of a wider aphasia therapy project, which offered a range of technology enhanced interventions to people with aphasia. The other wings worked on reading, discourse and conversation. The project received ethical clearance from the Bromley (London) NRES Committee

(14/LO/1531). All participants gave informed written consent, using materials designed to be accessible to people with aphasia (Rose et al, 2011).

Participants

The following selection criteria were applied: participants had aphasia following stroke; they were at least 4 months post onset and medically stable; they had no secondary cognitive diagnosis, such as dementia; they were fluent first or second language users of English pre stroke; they were not receiving speech and language therapy elsewhere during their involvement in the project. Most had to travel independently, or with support, to the University clinic. Criteria were established via language and cognitive screening, using the Comprehensive Aphasia Test (Swinburn et al, 2004) and the Cognitive Linguistic Quick Test (Helm-Estabrooks, 2001)) and via a case history interview. Participants were also asked about their prior technology use, although this did not determine inclusion. They were presented with 18 examples of technology (such as email, Facebook and Twitter) and asked if they had used these in the previous month. Participants directed to the writing wing of the wider project expressed a wish to improve their writing, and had functional writing goals. Figure 1 provides a flow diagram for the study. This shows that 25 participants were randomised, with 3 failing to complete initial assessment and 1 failing to complete intervention.

Insert Figure 1: Study Flow Diagram Here

Design

The study employed a quasi-randomised, waitlist controlled design. Following recruitment, participants were randomised to an Immediate or Delayed Therapy group (hereafter the 'Immediate' and 'Delayed' group). All undertook baseline assessment at time 1 (T1). Those in the Immediate group then received 6 weeks of technology enhanced writing therapy, while the Delayed group received no intervention. After 6 weeks, all participants were assessed again (T2). The Delayed group then received technology enhanced writing therapy for 6 weeks, while the Immediate group received no further intervention. Assessment was repeated (T3) following this period. The Delayed group received a follow up assessment 6 weeks after the end of their therapy (T4). Thus all participants underwent a pre therapy, post therapy and follow up assessment, with the delayed participants being assessed twice before therapy. In line with the recruitment criteria, participants received no other speech and language therapy during their involvement in the study. However, other forms of usual care, such as attendance at stroke support groups, continued. No therapy was received during the follow up period, six weeks after therapy ceased. However, participants retained the loaned therapy technology during this period for independent use.

Randomisation was conducted by order of recruitment, predetermined at the start of the study before any referrals were received. Thus referrals 1 - 11 and 22 - 24 were randomised to the immediate condition, while all other referrals were randomised to the delayed condition.

Therapy

Participants were given an initial 1-2 hours of technology training. This was followed by 12 one hour sessions of therapy delivered over 6 weeks (2 sessions per week), supplemented by independent homework practice. Therapy sessions took place face-to-face (1:1) and were largely delivered by students of speech and language therapy, working under the supervision of qualified therapists (AC, KM, CW). Most participants were treated in a University clinic. Two were treated in their own home and one at a community centre.

Assistive Technology

Two assistive technologies were employed, with a view to supporting individuals with a range of aphasic profiles. The first was WriteOnline^{TM}. This incorporates word prediction, vocabulary support in the form of individualised wordbars, and text to speech conversion, all of which have been shown to assist people with dysgraphia (e.g. Thiel et al, 2017). It also provides mind maps, which help to organise written work. The other technology was Dragon™, a VRS software which has been used successfully in three previous single case treatment studies for people with dysgraphia (Bruce & Edmundson, 2003; Caute & Woolf, 2016; Estes & Bloom, 2011).

Each participant used just one of these technologies in their therapy. Selection was made in discussion with their therapist, and in the light of language screening results. Those who had very impaired speech, which was likely to prohibit successful use of VRS, were directed to WriteOnline™ (N = 12). This was determined by the presence of severe dysarthria or dyspraxia, scores below 70% on the CAT word repetition sub-test, or a T score below 50 on the CAT picture description. Those who had very impaired reading and auditory comprehension (<9/15 correct on the CAT written and spoken word to picture matching tests) were discouraged from selecting WriteOnline™, and used VRS instead (N = 9). Both WriteOnline™ and Dragon™ were provided on either a laptop or an Ipad, with the choice dependent on participant preference and ability to use the touchscreen or mouse and keyboard interfaces (established via observation and discussion).

Technology Training

Therapy was preceded by 1-2 hours of initial training with the chosen technology. This introduced main aspects of technology use such as opening the software, identifying icons and carrying out key functions. For example, training with Dragon included 'waking up' the recording facility, dictating text, stopping recording, spelling unusual words, punctuating and saving files. Training was supported by an aphasia-friendly technology manual containing step-by-step instructions supported by screenshots and pictures (see samples in appendices). Following this training participants were provided with the relevant technology to take home on loan for the duration of their involvement.

Therapy Content

An initial discussion collaboratively identified individual writing goals for each participant. Individuals were facilitated during this discussion through personalised communication strategies, including the use of pictorial prompts. As a result of the goal setting discussion, each participant identified one writing activity that they aimed to accomplish by the end of the therapy period. Examples were: writing personal emails, writing a short story to share with a grandchild, and writing a biographical account of the participants' stroke and aphasia. Steps towards achieving these goals were also identified. These included aspects of the technology that needed to be mastered and gradated writing tasks leading to the end-point goal. For example, interim goals for participants who wished to write emails might include: setting up an email account, opening the relevant email software, sending and receiving emails, and creating email content.

Separate treatment manuals were created for WriteOnline™ and Dragon™. These determined the structure and content of therapy, while allowing for adaptation in response to the individual treatment goals. Therapy sessions followed a standard format, regardless of the technology employed. The first 4 sessions focused on the mastery of key features of the technology, so that these could be accomplished independently. For example, practiced features included dictating text with Dragon and using the word predictor bar with WriteOnline. Core skills relating to the personal goal were also initiated, such as navigating email software. Skill development was supported with structured tasks in the session. For example, Participants might dictate opening email greetings using Dragon or the WriteOnline word predictor bars.

Sessions 5 – 12 continued to incorporate technology practice, but at increasingly demanding levels. Participants were required to carry out tasks that included increasing numbers of steps, until they could accomplish all that were needed for their goal. Again this structure was followed regardless of the technology being employed. To exemplify, an intermediate task with Dragon would require: opening the software, recording text and pausing the recording; while a more challenging task would additionally require using punctuation, inserting a spelling and listening back to the text. Later sessions also introduced additional assistive features of the technology that supported participants towards their goal. For example, those using WriteOnline were taught how to add relevant vocabulary to the word bars.

Alongside the technology practice, sessions 5 – 12 placed an increasing emphasis on written composition, in relation to the treatment goal. For example, the participant who produced the story of her stroke and aphasia first identified the structure of what she wanted to include, then worked on separate sections of the story (events on the day of her stroke, events in hospital, life after discharge and longer term recovery). Hierarchical portions of text, of increasing length and complexity, were generated, reviewed and edited in the therapy sessions.

All participants were required to carry out at least one homework assignment each week throughout therapy. These were related to the treatment goal and were of increasing complexity. Thus, an early email homework task might be to compose an email wishing a friend Happy Birthday (without necessarily sending that email); while a later task required the participant to send an email to the treating therapist, confirming arrangements for the next treatment session. The following sessions began with a review of the homework assignment and any problems that arose carrying it out. Participants were encouraged to practice writing with the technology between sessions (over and above the homework assignments) and were asked to record this practice on a diary sheet. The sheet was reviewed and discussed each week.

Where relevant, final sessions included discussion about the presentation of the end written product (the font, binding, accompanying pictures etc). For example, this applied to individuals who produced a biographical text or a child's story. For some individuals this resulted in an output that could be shared with friends or family members and which was a source of considerable pride.

Measures

Primary Outcome Measure: Functional Assessment of Writing

Participants were asked to compose up to three emails. In the first, they were asked to arrange a meeting with a family member or friend, using a pictorial template specifying key information that should be included. In the second email they were asked to recount some news to a personal contact. In the third they were asked to write what they would do if they won the lottery. No templates were provided for the second and third emails. Participants were given 5 minutes to complete the first email, and 10 minutes for the second and third. At each time point, this task was administered under two conditions: in handwriting and technology assisted. The technology assisted task was carried out either in Microsoft Word™ on a computer or using the Ipad notes app, depending on which hardware had been selected for therapy. The relevant assistive technology (WriteOnLine™ or Dragon™) was also available. The conditions were counter balanced, to ensure that the order of administration did not favour one condition over the other, and the two conditions were not administered on the same day. Seventeen participants were tested on 3 emails in each condition at each time point, three were tested on 2 emails, and one was tested on just 1 email. The number of emails was reduced for four participants because they experienced high levels of difficulty with the task, causing potential distress and frustration. Tasks were consistent across each testing occasion; i.e. those who were tested on two emails completed the same two on each occasion.

Three scores were extracted from the email task at each time point. The first was a Lexical Quotient score. This comprised: the number of correctly spelt words across all emails, plus the number of different words produced, minus the number of spelling errors. This score had no upper limit. The second score was a Grammatical Quotient. Each email was rated for grammatical quality using criteria adapted from the CAT picture description task (Swinburn et al, 2004). Scores ranged from 0 (no syntactic structures used/no phrases well-formed) to 12 (a full range of structures used/all phrases well-formed, no phrases omitted). The Grammatical Quotient was the mean grammatical quality score across all emails. The third score was for Social Validity. Each email was rated by external raters (students of speech and language therapy), using criteria adapted from Jacobs (2001). Scores of 1 – 7 were awarded for Effectiveness (the degree to which the email conveyed the target message), Informativeness (the amount of information conveyed), Grammaticality (whether writing conformed to the rules of English) and Comfort (How comfortable you felt reading the email). The Social Validity score was the mean rating across all criteria for all emails. Separate Lexical Quotient, Grammatical Quotient and Social Validity scores were derived for the hand written and technology assisted conditions at each time point.

To assess the reliability of the Functional Assessment of Writing, 70 scripts were double scored for Lexical and Grammatical Quotient (16.2% of sample), and 46 for Social Validity (10.5% of sample). Double scored scripts were selected randomly across time points and conditions.

Secondary Outcome Measures

Written naming was assessed with the Boston Naming Test (Kaplan et al., 1983). Participants were shown each picture in turn and asked to provide its written name. Data were analysed from the handwritten version (not technology assisted) and reported the number of correctly spelt items.

Functional communication was assessed with the CADL-2 (Holland et al, 1999). This standardised assessment explores language use in everyday situations, such as going to the doctor. Although only 3 items specifically demand writing, a pencil and paper is provided throughout and many responses can be made in writing. Scoring reflects communicative success rather than formal language skills.

Mood was assessed with the Visual Analog Mood Scales Revised Version (VAMS-R: Kontou et al, 2012). This measure, which was designed for people with aphasia, collects ratings on 8 mood states (afraid, confused, sad, angry, tired, tense, happy and energetic) using pictorial visual analogue scales.

In order to reduce the number of analyses only data from the ‘sad’ question were analysed. This construct has been employed in previous aphasia trials (Thomas et al, 2013).

Aphasia related Quality of Life was assessed with the Assessment of Living with Aphasia (ALA, Simmons-Mackie et al, 2014). This self-report measure evaluates the impact of aphasia on five domains: language, participation, environment, personal and moving on with life. It yields an overall score which was analysed in this study.

The primary and secondary outcome measures were usually administered by the treating therapist or student (on occasion they were administered by other members of the research team). However, the written transcripts for the Functional Assessment of Writing and the Boston Naming Test were randomised and scored blind to time point and group allocation by scorers who had not delivered the intervention.

Analyses

For the primary outcome measure, two sets of analyses were performed. The first used mixed ANOVAs to compare the Immediate and Delayed groups across T1 and T2, with condition (technology assisted and handwritten) as a further within group factor. Here a significant treatment effect would be signaled by a time x group interaction, showing that the Immediate group (who had received therapy) improved, while the Delayed group (who had not yet received therapy) did not. A three way interaction (time x group x condition) would also show an effect of therapy, but one that was mediated by condition. This, for example, might show a compensatory pattern, whereby improvements were seen only when the task was technology assisted.

The second set of analyses was performed on combined data from all 21 participants. These employed two factor ANOVAs, with the factors being time (pre therapy, post therapy and follow up) and condition (technology assisted and handwritten). The pre therapy time point combined data from T1 for the immediate group with T2 from the delayed group. Post therapy combined data from T2 for the Immediate group with T3 for the Delayed group. Follow up combined data from T3 for the Immediate group with T4 for the Delayed group. Here a treatment effect would be supported by a main effect of time, with pairwise comparisons showing a significant difference between pre and post therapy. Maintenance of gain would be indicated by a significant difference between pre therapy and follow up. A time by condition interaction would indicate that gains were more evident in one condition than the other.

Secondary outcome measures were also subjected to two analyses. The first was a mixed ANOVA examining change over time between T1 and T2 and comparing the Immediate and Delayed group. Here a treatment effect would be signaled by a time x group interaction. The second one factor ANOVA explored change over time, from pre-therapy to post-therapy and follow up, across pooled data from all participants. Pairwise comparisons explored the locus of change if a main effect was present. On all ANOVAs the Greenhouse Geisser correction was applied when necessary.

Results

Participant details are reported in Table 1. There were no significant differences between the Immediate and Delayed groups with respect to their age, time post onset or screening scores on the CLQT and CAT. There were marginally more males in the Immediate than the Delayed group, but not significantly ($\chi^2 = 0.39$, $p = 0.53$). The technology use scores were the number of technological applications and devices (/18) that had been used by participants in the month prior to recruitment. Groups did not differ on this measure (data were unavailable for two participants in the immediate group).

T1 scores on the Functional Assessment of Writing (see table 2) were also compared using independent sample t tests. Although all scores were lower for the Delayed group, there were no significant differences between the groups (LQ tech assisted $p = .23$; LQ handwritten $p = .19$; GQ tech assisted $p = .27$; GQ handwritten $p = .12$; SV tech assisted $p = .77$; SV handwritten $p = .35$).

Insert Table 1 here: Participant Details

Reliability tests on the Functional Assessment of Writing

Data from the email task that were double scored were compared using Inter Class Correlations (absolute agreement). There was excellent inter-rater reliability for the Lexical Quotient (LQ ICC= 0.931) and moderate reliability for Grammatical Quotient (GQ ICC = .771) and Social Validity (SV ICC= .743).

The Delayed group was assessed twice before therapy began. Their T1 and T2 scores were compared to assess test/re-test reliability. Excellent reliability was evident on all scores: LQ Technology Assisted ICC = .977; LQ Handwritten ICC = .96; GQ Technology Assisted ICC = .962; GQ Handwritten ICC = .989; SV Technology Assisted ICC = .977; SV Handwritten ICC = .958.

Treatment Outcomes

Primary Outcome Measure: Functional Assessment of Writing

Scores on the Functional Assessment of Writing (T1 – T4) are reported in Table 2. The first set of ANOVAs compared the Immediate and Delayed groups across T1 and T2 on Lexical Quotient (LQ), Grammatical Quotient (GQ) and Social Validity (SV) scores. Results of these analyses are summarised in Table 3.

Taking Lexical Quotient first, the ANOVA table shows one significant main effect for time, indicating that scores increased between T1 and T2. The time by group interaction was not significant, but all other interactions were. The modality by group interaction was present because the Immediate group was particularly advantaged by the technology assisted condition. The time by condition interaction occurred because the technology assisted condition changed more over time than the handwritten condition. The three way interaction (time x modality x group) indicates a treatment effect. This shows that the Immediate group improved between T1 and T2, whereas the Delayed group did not. However, the gain occurred only in the technology assisted condition. Performance in handwriting was unchanged (see Figure 2).

The Grammatical Quotient ANOVA produced no significant main effects and only one significant interaction: time x modality x group. As with the LQ result, this shows an effect of therapy mediated by condition. Figure 3 shows that the immediate group improved between T1 and T2 when assisted by technology. In handwriting their performance declined, although not significantly.

The Social Validity ANOVA produced a highly significant main effect of time, indicating improvement between T1 and T2. One interaction was significant, between time and group. This shows an effect of therapy, as the Immediate group improved more markedly than the Delayed group. The scores reported in table 2 suggest that the improvement for the Immediate group occurred mainly in the technology assisted condition. However, the three way interaction falls just short of significance (see Figure 4).

Insert table 2 here: Scores on the Functional Assessment of Writing at each Assessment Point

Insert table 3 here: Results of ANOVA analyses on the Functional Assessment of Writing Scores

Insert Figures 2 - 4 here

Table 4 reports pre therapy, post therapy and follow up scores on the Functional Assessment of Writing. These are combined data across both groups (n = 21). Each scoring variable was examined using two factor ANOVAs (time, condition), with pairwise comparisons across the time points.

LQ results produced a significant main effect of time (F (2, 24.3) = 9.27, p = .004, η^2= .317). Pairwise comparisons showed a significant improvement between pre and post therapy (p = .001) which was well maintained (pre therapy vs follow up, p = .005). The comparison between post therapy and follow up was not significant (p = .118). There was also a significant main effect of condition (F (1,20) = 5.034, p = .036, η^2= .201), indicating that scores in the technology assisted condition were significantly higher than scores in the handwritten condition. The interaction between time and condition was significant (F (2, 24.501) = 5.8. p = .006, η^2= .225), indicating that the improvement over time occurred largely in the technology assisted condition.

GQ results produced a significant main effect of time (F (2,31.09) = 5.705, p = .007, η^2=.222). Pairwise comparisons were significant for pre-therapy vs post-therapy (p = .014) and pre-therapy vs follow up (p = .019) but not for post-therapy vs follow up (p = .916). These comparisons indicated that scores improved following therapy, with good maintenance of gain. The main effect of condition was also significant (F (1, 20) = 6.38, p = .02, η^2=.242), with technology assisted scores being highest. The interaction between time and condition was significant (F (2, 38.91) = 13.705, p <.001, η^2=.407). The descriptive statistics show that this interaction occurred because gains were entirely confined to the technology assisted condition.

The SV analysis produced a significant main effect of time (F (2,40) = 21.52, p <.001, η^2=.518). Pairwise comparisons showed an improvement between pre- and post-therapy (p <.001) which was well maintained (pre-therapy vs follow up, p <.001). The difference between post-therapy and follow up was not significant (p = .087). The main effect of condition was not significant (p = .08), but there was a significant interaction between time and condition (F (2,40) = 6.87, $p = .003$, $\eta^2=.256$). This indicated that the gain was more marked in the technology assisted condition.

Insert Table 4 here: Scores on the Functional Assessment of Writing across all participants (n = 21), pre-therapy, post therapy, and follow up.

Secondary Outcome Measures

Table 5 shows the scores (T1 – T4) on all secondary outcome measures. The first set of analyses employed mixed ANOVAs with the factors of time (T1 and T2) and group (Immediate and Delayed). Four participants in the Immediate group, and one participant in the Delayed group did not complete the written Boston Naming Test (BNT). Data for this measure are therefore incomplete and the analysis underpowered. The ANOVA for the BNT scores produced a main effect of time (F (1, 14) = 5.504, $p = .034$, $\eta^2=.282$), indicating that scores across all participants improved. There was no effect of group ($p=.091$) and, crucially, no interaction ($p=.844$). Thus the improvement could not be attributed to treatment. There was a full data set for all remaining outcome measures (CADL-2, VAMS, ALA), but results were all non-significant.

Insert table 5 here: Scores on the Secondary Outcome Measures at each Assessment Point

Table 6 shows pre-therapy, post-therapy and follow up scores on the secondary measures, combining data across the Immediate and Delayed groups. These data were analysed with one factor ANOVAs (time), with pairwise comparisons across the time points.

There was no effect of time on the analysis of the Boston Naming Test ($p = .153$). The analysis of the CADL-2 data produced a main effect of time (F (2, 38) = 10.927, $p < .001$, $\eta^2=.365$). Pairwise comparisons were significant for pre-therapy vs post-therapy ($p = .002$) and pre-therapy vs follow up ($p < .001$) but not for post-therapy vs follow up. There were no significant findings from the analysis of the VAMS data. The analysis for the ALA produced a main effect of time ($F (2, 38) = 4.47$, $p = .018$, $\eta^2=.19$). In the pairwise comparisons only pre- vs post-therapy was significant ($p = .002$).

Insert table 6 here: Pre Therapy, Post Therapy and Follow up Scores on the Secondary Outcome Measures

Discussion

This study evaluated outcomes from technology enhanced writing therapy delivered to 21 people with aphasia and dysgraphia. Treatment aimed to improve functional writing skills, which have proved difficult to remediate in previous treatment studies (Thiel et al, 2015). Accordingly, the primary outcome measure reflected an authentic writing task, that of email composition. We first discuss the reliability of this measure. We then consider each of the research questions, before discussing limitations of the study and future implications.

Blind double scoring of a proportion of data from the primary outcome measure (the Functional Assessment of Writing) was conducted to assess inter-rater reliability. Agreement for the Lexical Quotient scores was excellent, but less strong for Grammatical Quotient and Social Validity. This probably reflects the nature of the scoring, in that LQ scores were based on word counts, whereas the GQ and SV scores were rating judgements. Despite this, the latter scores still demonstrated moderate reliability. Test/re-test reliability was assessed by comparing T1 and T2 scores from the Delayed group. Here reliability was excellent, showing that performance on the measure did not change over a 6 week no treatment period. The Functional Assessment of Writing is a novel assessment, and its psychometric properties have not been fully assessed. However, these reliability data allow for confidence in the scoring methods, and show that scores were stable before therapy began.

The first research questions asked whether technology enhanced writing therapy would improve performance on the Functional Assessment of Writing, and whether gains would be evident in both the technology assisted and handwritten conditions. Results pointed to a very clear treatment effect. At T2 the Immediate group improved on the assessment, whereas the Delayed group (who had yet to receive therapy) did not. The descriptive data at T3 indicated that the Delayed group also profited from therapy, once this was instigated. Indeed, when scores were combined across the whole group, there was a significant improvement from pre- to post-therapy. This was also maintained at 6 weeks follow up, suggesting some durability of treatment gains. Encouragingly, all scores in the technology assisted condition improved. Thus, therapy elicited more correctly spelt words (LQ), improved the grammatical status of participants' writing (GQ) and enhanced the perceived quality of their output (SV).

Turning to the second part of the question, treatment gains were clearly mediated by condition, since, in almost all analyses, improvements occurred only when writing was assisted by technology. As in previous studies of technology enhanced writing treatments (Thiel et al, 2016; Caute & Woolf, 2016) this points to a compensatory outcome. Participants learnt to use the technology to improve the quality of their writing. The dysgraphia was not remediated and became evident when unassisted handwriting was assessed.

The lack of remediation may raise concerns about the clinical utility of therapy. It indicates, for example, that long term therapy benefits depend on the accessibility of the technology and its continuing use by participants. Our choice of widely available mainstream technologies ensured these were relatively affordable, particularly compared to bespoke Alternative & Augmentative Communication systems. The assistive features that we explored, such as VRS and predictive spelling, are also increasingly included within the standard operating systems of new hardware. We would therefore hope that some potential barriers to long term use, relating to availability and cost, were mitigated. However, continuing use of the trained technologies will, likely, depend on a range of personal and environmental factors (see Menger et al, 2016 for similar arguments). Our participant interviews and observations of their technology use further explored these factors and will be reported in future papers.

The clinical benefits from therapy are further illustrated by the extent of change on the Functional Assessment of Writing, where scores were substantially, as well as significantly improved by therapy. For example, the average gain on the lexical quotient was over 30 for both groups. As illustrated by Figure 5, this gain was often marked for the individuals involved. Writing shifted from being unusable, to a viable and frequently employed medium of communication.

Insert Figure 5 about here

Two compensatory software packages were employed in treatment. The study did not, however, explore whether these led to differing outcomes. Descriptive statistics suggested that gains were comparable, but numbers were too low to permit finer grained analyses. The relevance of any software comparison could also be questioned, given that the technologies were not randomly assigned, but were selected in the light of the participants' aphasic profiles. The study findings show that two different technologies can be employed by people with aphasia to assist with writing, rather than assessing the relative merits of those technologies.

The second research question asked whether therapy would improve written naming when not assisted by technology. Results on the written version of the Boston Naming Test indicated that it did not. Owing to missing data, the first between group analysis was underpowered, making possible effects difficult to detect. However, even when results for the whole group were analysed (n = 16) there was no evidence of change from pre therapy to post therapy and follow up. This measure, therefore, strengthens the conclusion that therapy did not remediate the writing impairment.

The third research question was whether therapy would enhance functional communication as assessed by the CADL-2. Results did not point to a treatment effect, as there was no interaction in the first analysis which compared the Immediate and Delayed groups across T1 and T2. The combined, whole group analysis did show improvement over time. However, this analysis lacked experimental control, making it difficult to attribute change to therapy. We hypothesised that improved access to writing might generally enhance functional communication; for example, by enabling participants to convey information in the written modality when they were not able to access the spoken word. The fact that writing remediation did not occur, and that few subtests of the CADL-2 explicitly call on writing, might account for the null result on this test.

The final questions asked whether therapy would improve mood, as assessed by the VAMS, and Quality of Life, as assessed by the ALA. Findings for mood in both analyses were negative. There was also little evidence of change on the ALA, given the lack of any interaction in the first analysis. The whole group analysis showed a main effect of time, with a significant improvement from pre to post therapy. However, as with the CADL-2, this evidence alone is inconclusive. Stroke and aphasia often have profoundly negative consequences for emotional state and quality of life (Kauhanen et al, 1999; Hilari, 2011). Expecting changes in these areas was possibly over optimistic, given that therapy was only 6 weeks in duration and not specifically focussed on psychological or quality of life goals. There may also have been a problem of measurement. The VAMS manual acknowledges that test-retest reliability is affected by the fluctuating nature of mood states (Stern, 1996). Those who administered the VAMS in the current study also observed that scores were influenced by many external factors, such as problems with finances or housing. It is difficult to demonstrate a treatment effect when there are many other sources of variation on a measure, particularly when the sample size is low. Responses to the post therapy interviews might offer more insights into whether therapy has brought about emotional change.

Before discussing the implications of these findings, some limitations need to be acknowledged. The first relate to our primary outcome measure, which has not been subject to extensive psychometric testing. Further checking of reliability with a larger sample would strengthen confidence in this measure. Comparison with a standard assessment would also be desirable, although difficult to achieve, given the dearth of measures in this domain. The burden of the Functional Assessment of Writing is high, both for participants and scorers, since it involves up to 3 email tasks that are subjected to 3 scoring methods. Further testing could usefully explore whether a pruned version of the test can achieve sensitivity, particularly if the test is to be used in routine clinical practice. Data from healthy controls would also support interpretation of aphasic data.

The primary outcome measure was administered in technology assisted and handwritten conditions. It might have been informative additionally to assess participants in typing alone. For example, in comparison to handwriting, this condition might have been more sensitive to any therapy induced change in unassisted writing. This option was not pursued for reasons of participant burden.

The wider project, which hosted this study, aimed to provide a therapy service to people with aphasia together with clinical training for students of speech & language therapy. Achieving stringent research standards within this context was difficult. So, test administration was not blind (although scoring of the primary outcome measure was) and randomisation was conducted on a pragmatic, order of referral, basis. The context also limited the follow up period to 6 weeks. Findings would be augmented by a much longer term follow up after the technology had been withdrawn. In addition to retained skills, such follow up could also investigate whether participants were still, independently, accessing the technology.

We would also acknowledge limitations related to the sample. Although larger than many previous studies in this area, the number of participants was small. Findings are therefore vulnerable to type 1 and type 2 errors. The typicality of the sample could also be questioned. With an average age of

56, the sample is younger than most stroke survivors (Engelter et al, 2006). The University setting, and the need for independent travel by most participants may have introduced further skews. For example, the sample may have had fewer co-morbidities, and may have been more pre-disposed to technology use than the wider stroke population. Our sample was also not subjected to extensive background testing, to determine the precise nature of their dysgraphia. We cannot, therefore, comment on whether some dysgraphic impairments are more readily compensated for by assistive technology than others.

Despite these limitations, this study showed that 21 people with aphasia were able to employ a mainstream, assistive technology in order to improve the quality of their writing. Unlike many previous therapy investigations (see Thiel et al 2015, for review) the improvement was not confined to tests of written naming, but was evident on a functional task involving email composition. Many participants also generated personally meaningful outputs during their therapy, such as autobiographical texts, which provided them with a satisfying legacy of the treatment. The gains were achieved from a modest therapy dose (12 sessions), which is important given the constraints operating over most health services (Code & Petherham, 2011). Although participants were younger than many stroke survivors they had very chronic aphasia, showing that at least compensatory change can be achieved well beyond the acute period. The study also shows that positive outcomes can accrue from a delegated model of therapy, given that delivery was largely by students of speech and language therapy.

Further research is needed to explore outcomes with a larger and possibly more typical sample, e.g. with respect to age. Findings would be augmented with long term follow up, exploring barriers and facilitators to long-term use of the technologies and a cost benefit analysis. Fine grained investigations of factors that affect candidacy, and comparing different models of delivery, would help to clarify which people with aphasia are most likely to benefit from this type of therapy and how best to meet their needs.

Acknowledgements

This study was funded by The Barts Charity. We thank all our participants, and all our students of speech and language therapy who supported this work.

References

ARMSTRONG, L. & MACDONALD, A. (2000) Aiding chronic written language expression difficulties: a case study. Aphasiology, 14(1), 93–108

BALL, A. L., DE RIESTHAL, M., BREEDING, V. E., & MENDOZA, D. E. (2011). Modified ACT and CART in severe aphasia. Aphasiology, 25, 836–848. doi:10.1080/02687038.2010.544320.

BEESON, P., REWEGA, M., VAIL, S. & RAPCSAK, S. (2000) Problem-solving approach to agraphia treatment: Interactive use of lexical and sublexical spelling routes, Aphasiology, 14:5-6, 551-565, DOI: 10.1080/026870300401315.

BEESON, P. M., HIRSCH, F. M., & REWEGA, M. A. (2002). Successful single-word writing treatment: Experimental analyses of four cases. Aphasiology, 16, 473–491. doi:10.1080/02687030244000167

BEESON, P., HIGGINSON, K., & RISING, K. (2013). Writing treatment for aphasia: A texting approach. Journal of Speech, Language, and Hearing Research, 56, 945–955. doi:10.1044/1092-4388(2012/11-0360)

BEESON, P. & RAPCSAK, S. (2015) Clinical diagnosis and treatment of spelling disorders. . In A. Hillis (ed) The Handbook of Adult Language Disorders, Second Edition, New York: Psychology Press.

BEHRNS, I., HARTELIUS, L. & WENGELIN, A. (2009) Aphasia and computerised writing aid supported treatment. Aphasiology, 23(10), 1276–1294

BRUCE, C., EDMUNDSON, A. & COLEMAN, M. (2003), Writing with voice: an investigation of the use of a voice recognition system as a writing aid for a man with aphasia. International Journal of Language and Communication Disorders, 38(2), 131–148.

CARAMAZZA, A., MICELI, G., VILLA, G. & ROMANI, C. (1987) The role of the graphemic buffer in spelling: Evidence from a case of acquired dysgraphia. Cognition, 26, 59 – 85.

CAUTE, C. & WOOLF, C. (2016) Using voice recognition software to improve communicative writing and social participation in an individual with severe acquired dysgraphia: an experimental single-case therapy study, Aphasiology, 30:2-3, 245-268, DOI: 10.1080/02687038.2015.1041095

CLAUSEN, N. S., & BEESON, P. M. (2003). Conversational use of writing in severe aphasia: A group treatment approach. Aphasiology, 17, 625–644. doi:10.1080/02687030344000003

CODE, C. & PETHERAM, B. (2011) Delivering for aphasia, International Journal of Speech-Language Pathology, 13:1, 3-10, DOI: 10.3109/17549507.2010.520090

ESTES, C., & BLOOM, R. L. (2011). Using voice recognition software to treat dysgraphia in a patient with conduction aphasia. Aphasiology, 25, 366–385. doi:10.1080/02687038.2010.493294

ENGELTER S., GOSTYNSKI, M., PAPA, S., FREI, M., BORN, C., AJDACIC-GROSS, V., GUTZWILLER, F., & LYRER, P. (2006) Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. Stroke, 37(6), 1379-84.

GREENWALD, M. (2004). “Blocking” lexical competitors in severe global agraphia: A treatment of reading and spelling. Neurocase, 10, 156–174.

HELM-ESTABROOKS, N. (2001). Cognitive-Linguistic Quick Test (CLQT). Pearson, UK.

HILARI, K. (2011). The impact of stroke: are people with aphasia different to those without? Disability and Rehabilitation, 33(3), pp. 211–218.

HOLLAND A., FRATTALI C., FROMM D. (1999) Communication Activities of Daily Living-2. Austen TX: Pro-Ed.

JACOBS, B. J. Social validity of changes in informativeness and efficiency of aphasic discourse following Linguistic Specific Treatment (LST). Brain and Language, Vol 78(1), Jul, 2001. pp. 115-127.

KAPLAN, E., GOODGLASS, H. & WEINTRAUB, S (1983). Boston Naming Test. Philadelphia: Lea & Febiger.

KAUHANEN, M. L., KORPELAINEN J.T., HILTUNEN, P., BRUSIN, E., MONONEN, H., MÄÄTTÄ, R., NIEMINEN, P., SOTANIEMI, K. A., & MYLLYLÄ, V. V. (1999). Poststroke depression correlates with cognitive impairment and neurological deficits. Stroke, 30(9), 1875-1880.

KIRAN, S. (2005). Training phoneme to grapheme conversion for patients with written and oral production deficits: A model-based approach. Aphasiology, 19, 53–76. doi:10.1080/02687030444000633

KONTOU, E., THOMAS, S. & LINCOLN, N. (2012) Psychometric properties of a revised version of the Visual Analog Mood Scales, Clinical Rehabilitation, 26, 12, 1133 – 1140.

LAGANARO, M., DI PIETRO, M., AND SCHNIDER, A. (2006). Computerised treatment of anomia in acute aphasia: treatment intensity and training size. Neuropsychol. Rehabil. 16, 630–640. doi: 10.1080/09602010543000064

LUZZATTI, C., COLOMBO, C., FRUSTACI, M., & VITOLO, F. (2000). Rehabilitation of spelling along the sub-word-level routine. Neuropsychological Rehabilitation, 10, 249–278. doi:10.1080/096020100389156

MENGER, F., MORRIS, J. & SALIS, C. (2016) Aphasia in an Internet age: wider perspectives on digital inclusion, Aphasiology, 30, 2-3, 112-132, DOI: 10.1080/02687038.2015.1109050

MORTLEY,J., ENDERBY, P. & PETHERAM, B. (2001) Using a computer to improve functional writing in a patient with severe dysgraphia, Aphasiology, 15, 5, 443-461, DOI: 10.1080/02687040042000188

MURRAY, L. & KARCHER, L. (2000), A treatment for written verb retrieval and sentence construction skills. Aphasiology, 14(5–6), 585–602.

MURRAY, L. L., TIMBERLAKE, A., & EBERLE, R. (2007). Treatment of underlying forms in a discourse context. Aphasiology, 21, 139–163. doi:10.1080/02687030601026530

PANTON, A., & MARSHALL, J. (2008). Improving spelling and everyday writing after a CVA: A single case therapy study. Aphasiology, 22, 164–183. doi:10.1080/02687030701262605

PAPATHANASIOU, I. & CSEFALVAY, Z. (2017) Written language and its impairments. In I. Papathanasiou & P. Coppens (eds) Aphasia and related neurogenic communication disorders (second edition). (Burlington MA: Jones and Bartlett Learning) pp. 219 – 244.

PATTERSON, K. 7 WING, A. (1989) Processes in handwriting: A case for case. Cognitive Neurospychology, 6, 1 – 23.

RAPP, B. & FISCHER-BAUM, S. (2015) Uncovering the cognitive architecture of spelling. In A. Hillis (ed) The Handbook of Adult Language Disorders, Second Edition, New York: Psychology Press.

RAYMER, A. M., CUDWORTH, C., & HALEY, M. A. (2003). Spelling treatment for an individual with dysgraphia: Analysis of generalisation to untrained words. Aphasiology, 17, 607–624. doi:10.1080/02687030344000058

ROBSON, J., MARSHALL, J., CHIAT, S., & PRING, T. (2001). Enhancing communication in jargon aphasia: A small group study of writing therapy. International Journal of Language & Communication Disorders, 36, 471–488. doi:10.1080/13682820110089371

ROSE, T., WORRALL, L., HICKSON, L. & HOFFMANN, T. (2011) Aphasia friendly written health information: Content and design characteristics. International Journal of Speech-Language Pathology, 13 4: 335-347. doi:10.3109/17549507.2011.560396

SALIS, C., & EDWARDS, S. (2010). Treatment of written verb and written sentence production in an individual with aphasia: A clinical study. Aphasiology, 24, 1051–1063. doi:10.1080/02687030903269648

SIMMONS-MACKIE, N., KAGAN, A., VICTOR, C., CARLING-ROWLAND, C., MOK, A., HOCH, J., HUIJBREGTS, M., STREINER, D. (2014) The assessment for living with aphasia: Reliability and construct validity. International Journal of Speech-Language Pathology, 16(1), 82–94.

STERN, R. (1996) Visual Analog Mood Scales, Professional Manual. Odessa FL: Psychological Assessment Resources.

SWINBURN, K., PORTER, G., & HOWARD, D. (2004). *Comprehensive Aphasia Test.* Psychology Press, UK

THIEL, L., SAGE, K. & CONROY, P. (2015), Retraining writing for functional purposes: a review of the writing therapy literature. Aphasiology, 29(4) 423–441.

THIEL, L., SAGE, K. & CONROY, P. (2016) The role of learning in improving functional writing in stroke aphasia. Disability and Rehabilitation, 38, 21, 2122 – 2134.

THIEL, L., SAGE, K. & CONROY, P. (2017) Promoting linguistic complexity, greater message length and ease of engagement in email writing in people with aphasia: initial evidence from a study utilizing assistive writing software. International Journal of Language and Communication Disorders, 52, 1, 106–124.

THOMAS, S.A., WALKER, M.F., MACNIVEN J.A., HAWORTH, H. & LINCOLN, N.B. (2013). Communication and Low Mood (CALM): a randomized controlled trial of behavioural therapy for stroke patients with aphasia Clinical Rehabilitation. 27(5), 398-408

WADE, J., PETHERAM, B. & CAIN, R. (2001), 'Voice recognition and aphasia: can computers understand aphasic speech?' Disability and Rehabilitation, 23, 14, 604-613.

Table 1: Participant Details

	N	Male n (%)	Female n (%)	Age at recruitment Mean Range (SD)	Months post stroke Mean Range (SD)	CLQT composite score Mean Range (SD)	CAT summative score Mean Range (SD)	Technology use score Mean Range (SD)
Whole group	21	12 (57%)	9 (43%)	56.00 44-76 (8.01)	57.21 12-146 (37.99)	2.71 1-4 (0.80)	128.17 24-169 (38.53)	11.26* 4-17 (3.43)
Immediate	11	7 (64%)	4 (36%)	55.44 49-60 (4.07)	65.00 15-146 (41.12)	2.46 1-3.6 (0.77)	129.00 65-164 (32.32)	11.11# 5-15 (3.14)
Delayed	10	5 (50%)	5 (50%)	56.50 44-76 (10.64)	50.20 12-132 (35.59)	2.98 2-4 (0.79)	127.25 24-169 (46.23)	11.40 4-17 (3.83)
			Sig. (2 tailed) *p*	0.783	0.412	0.146	0.92	.861
			T	-0.279	0.841	-1.518	0.101	-.178

* n = 19; #n = 9

Table 2: Scores on the Functional Assessment of Writing (T1 – T4)

	Group	Tech-Assisted Mean (SD)				Handwritten Mean (SD)			
		T1	T2	T3	T4	T1	T2	T3	T4
LQ	Whole group	42.17 (42.52)	70.13 (70.56)	100.00 (102.52)		47.78 (37.31)	52.97 (37.75)	56.75 (36.68)	
	Immediate	53.00 (49.34)	101.06 (81.97)	126.15 (132.30)		58.03 (33.29)	58.79 (35.13)	63.03 (28.90)	
	Delayed	30.25 (31.80)	36.10 (33.80)	71.23 (46.41)	77.60 (49.16)	36.50 (39.92)	46.57 (41.34)	49.83 (44.27)	57.63 (52.43)
GQ	Whole group	4.97 (2.87)	5.64 (2.65)	6.70 (1.67)		5.15 (2.93)	5.11 (3.02)	5.12 (2.87)	
	Immediate	5.64 (2.22)	6.94 (1.32)	6.70 (1.89)		6.11 (2.33)	5.64 (2.68)	5.95 (2.52)	
	Delayed	4.23 (3.05)	4.20 (3.05)	6.70 (1.48)	6.23 (2.02)	4.10 (3.26)	4.53 (3.40)	4.20 (3.08)	4.67 (3.30)
SV	Whole group	4.34 (2.87)	5.31 (2.73)	5.97 (2.03)		4.47 (2.95)	4.95 (2.91)	5.04 (2.95)	
	Immediate	4.52 (2.25)	6.36 (1.69)	5.69 (1.66)		5.06 (2.26)	5.50 (2.55)	5.31 (2.64)	
	Delayed	4.14 (3.55)	4.15 (3.24)	6.2850 (2.43)	6.03 (1.89)	3.83 (3.57)	4.35 (3.29)	4.75 (3.37)	4.53 (3.48)

Key: LQ = Lexical Quotient (no maximum score); GQ = Grammatical Quotient (maximum score 12); SV = Social Validity (Maximum score 7)

Table 3: Results of the Analyses of Functional Assessment of Writing Scores (T1 vs T2; Immediate vs Delayed)

Variable	Main Effects												Interaction Effects															
	Time				**Modality**				**Group** (*between subjects variable*)				**time*group**				**modality*group**				**time*modality**				**time* modality*group**			
	Df	F	*p*	η 2	Df	F	*p*	η 2	df	F	*p*	η 2	df	F	*p*	η 2	df	F	*p*	η 2	df	F	*p*	η 2	df	F	*p*	η 2
LQ	1	11.239	0.003**	0.372	1	0.658	0.427	0.033	1	2.838	0.108	0.130	1	2.903	0.105	0.133	1	4.546	0.046*	0.193	1	4.857	0.040*	0.204	1	6.944	0.016*	0.268
GQ	1	1.901	0.184	0.091	1	0.317	0.580	0.016	1	2.578	0.125	0.119	1	0.235	0.633	0.012	1	0.843	0.370	0.042	1	1.898	0.184	0.091	1	5.579	0.029*	0.227
SV	1	17.872	< 0.0001**	0.485	1	0.178	0.677	0.009	1	1.112	0.305	0.055	1	6.958	0.016*	0.268	1	0.042	0.839	0.002	1	0.865	0.364	0.044	1	4.094	0.057	0.177

Table 4: Pre Therapy, Post Therapy and Follow up Scores on the Functional Assessment of Writing

Variable	Modality	Mean (SD)		
		Pre-Therapy	**Post- Therapy**	**Follow- Up**
LQ	**Tech-assisted**	44.95 (42.50)	86.86 (67.54)	103.03 (102.26)
	Handwritten	52.57 (36.84)	54.52 (38.99)	60.46 (40.77)
GQ	**Tech-assisted**	4.95 (2.68)	6.83 (1.37)	6.48 (1.92)
	Handwritten	5.36 (2.93)	4.95 (2.90)	5.34 (2.92)
SV	**Tech-assisted**	4.34 (2.70)	6.32 (2.02)	5.85 (1.74)
	Handwritten	4.72 (2.75)	5.14 (2.92)	4.94 (3.02)

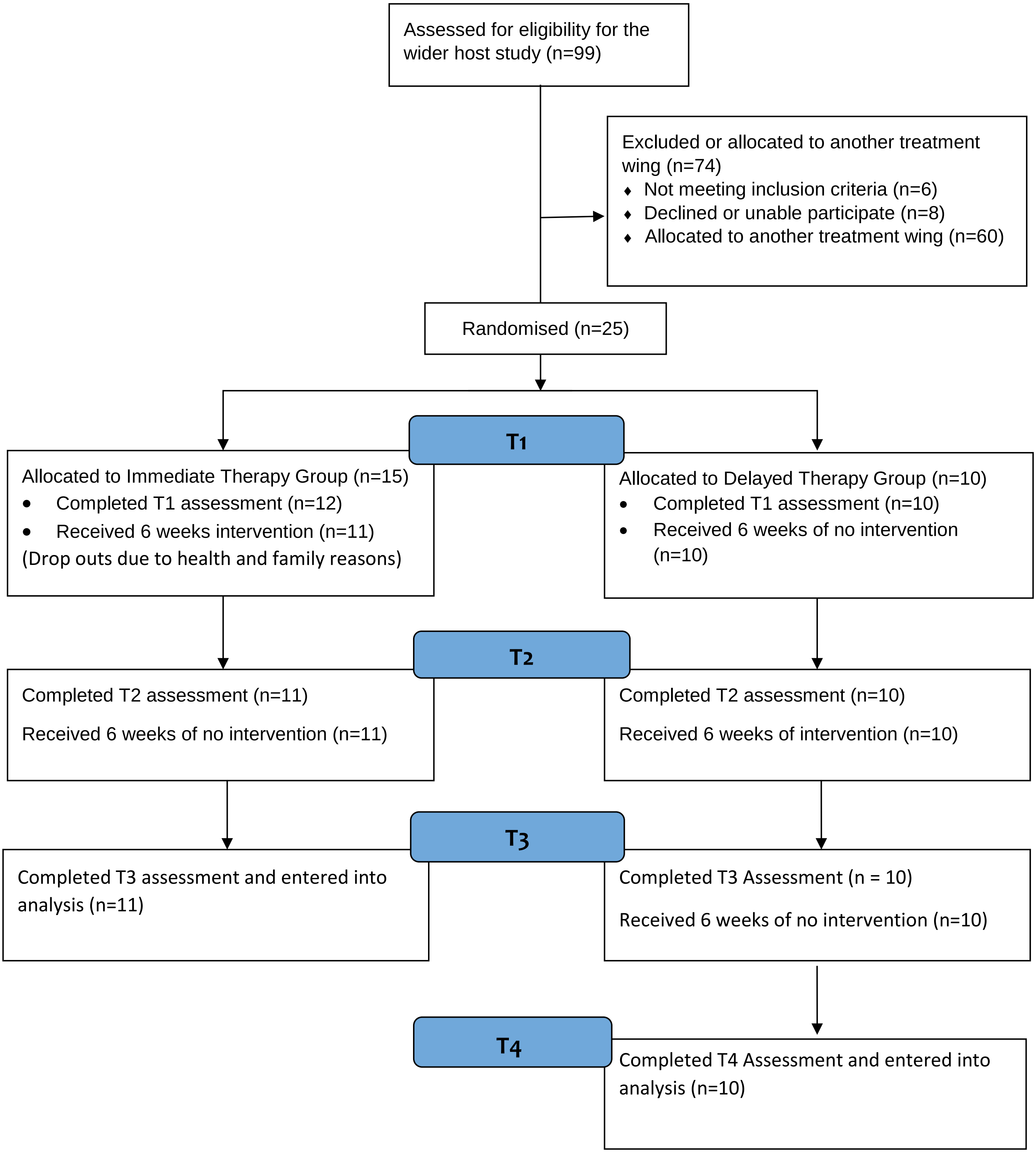
Table 5: Scores on the Secondary Outcome Measures at each Assessment Point

Measure	Group	Mean (SD)			
		T1	T2	T3	T4
BNT (written)	Whole group N = 16	15.44 (16.16)	18.44 (17.49)	18.06 (16.46)	
	Immediate N = 7	23.57 (19.79)	26.28 (21.44)	25.43 (19.58)	
	Delayed N = 9	9.11 (9.62)	12.33 (11.55)	12.33 (11.67)	17.7 (18.75)
CADL-2	Whole group	82.86 (9.15)	84.57 (8.42)	87.43 (6.54)	
	Immediate	84.45 (3.75)	88.18 (4.58)	89.27 (5.59)	
	Delayed	81.10 (12.81)	80.60 (10.03)	85.40 (7.18)	85.78 (10.32)
VAMS (Sad)	Whole group	50.86 (10.06)	50.05 (10.27)	53.24 (14.59)	
	Immediate	53.09 (11.43)	49.18 (11.73)	50.82 (12.56)	
	Delayed	48.40 (8.18)	51.00 (8.92)	55.90 (16.82)	48.10 (12.81)
ALA	Whole group	2.48 (.52)	2.62 (.68)	2.67 (.72)	
	Immediate	2.41 (.61)	2.63 (.81)	2.54 (.93)	
	Delayed	2.56 (.41)	2.60 (.53)	2.82 (.38)	2.76 (.37)

Table 6: Pre Therapy, Post Therapy and Follow up Scores on the Secondary Outcome Measures

Measure	Mean (SD)		
	Pre-Therapy	Post- Therapy	Follow- Up
BNT (written) N = 16	18.94 (17.13)	20.29 (18.63)	20.88 (18.89)
CADL-2	82.20 (7.44)	86.55 (5.96)	87.70 (8.03)
VAMS (Sad)	52.09 (10.11)	52.38 (14.42)	49.52 (10.14)
ALA	2.48 (.57)	2.72 (.65)	2.64 (.72)

Figure 1: Study Flow Diagram



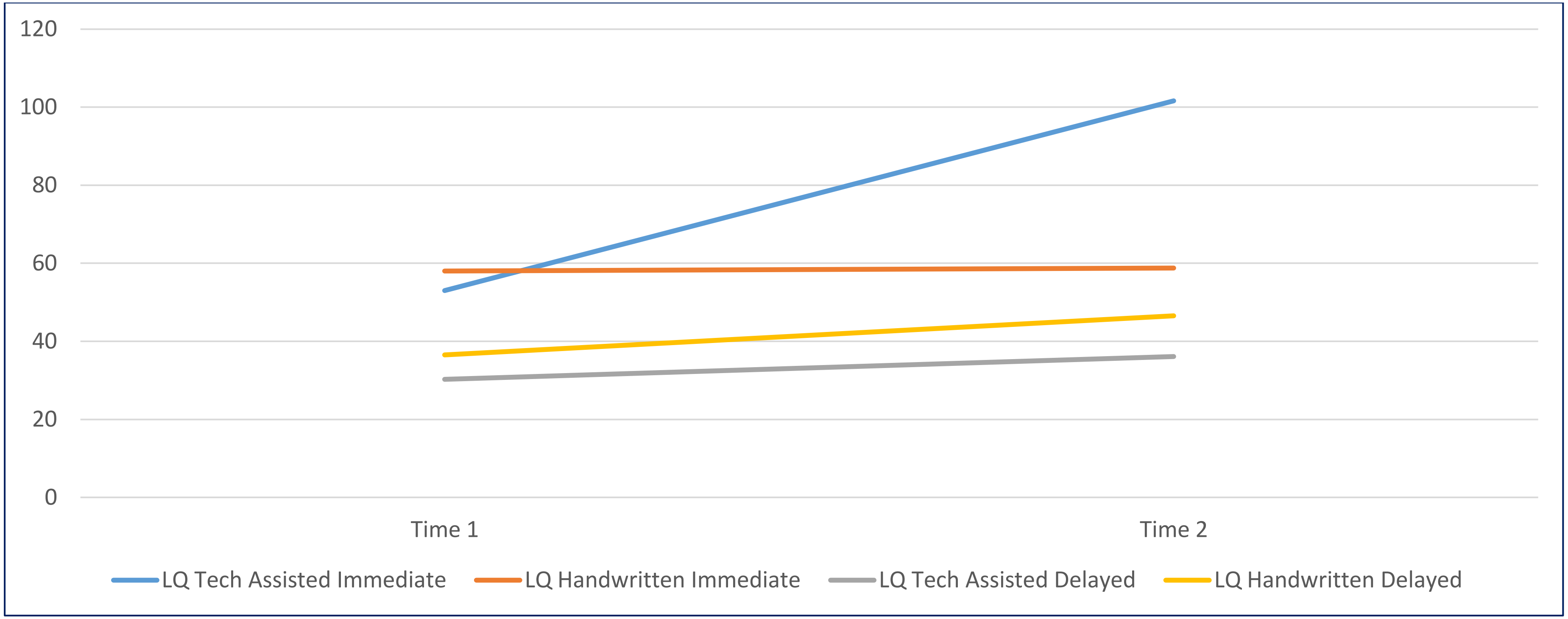


Figure 2: Mean Lexical Quotient Scores at Time 1 and Time 2: Immediate and Delayed Groups

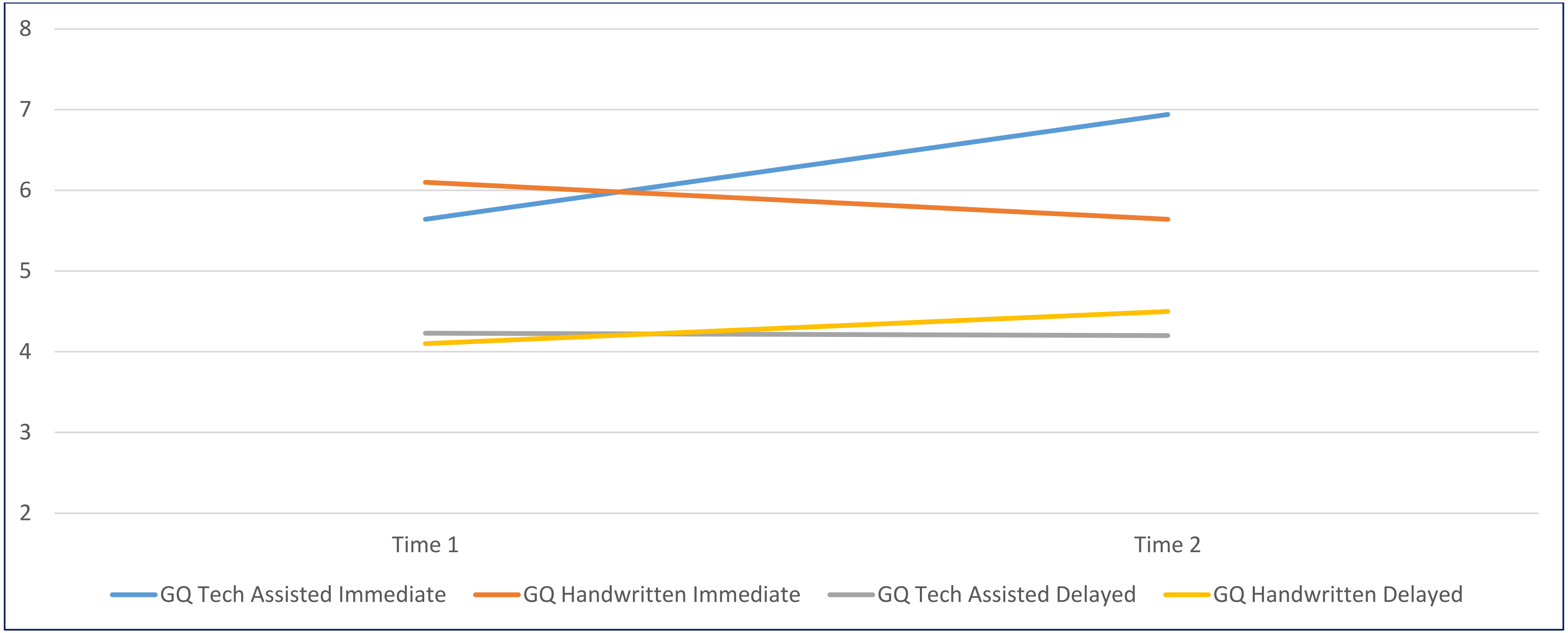


Figure 3: Mean Grammatical Quotient Scores at Time 1 and Time 2: Immediate and Delayed Groups

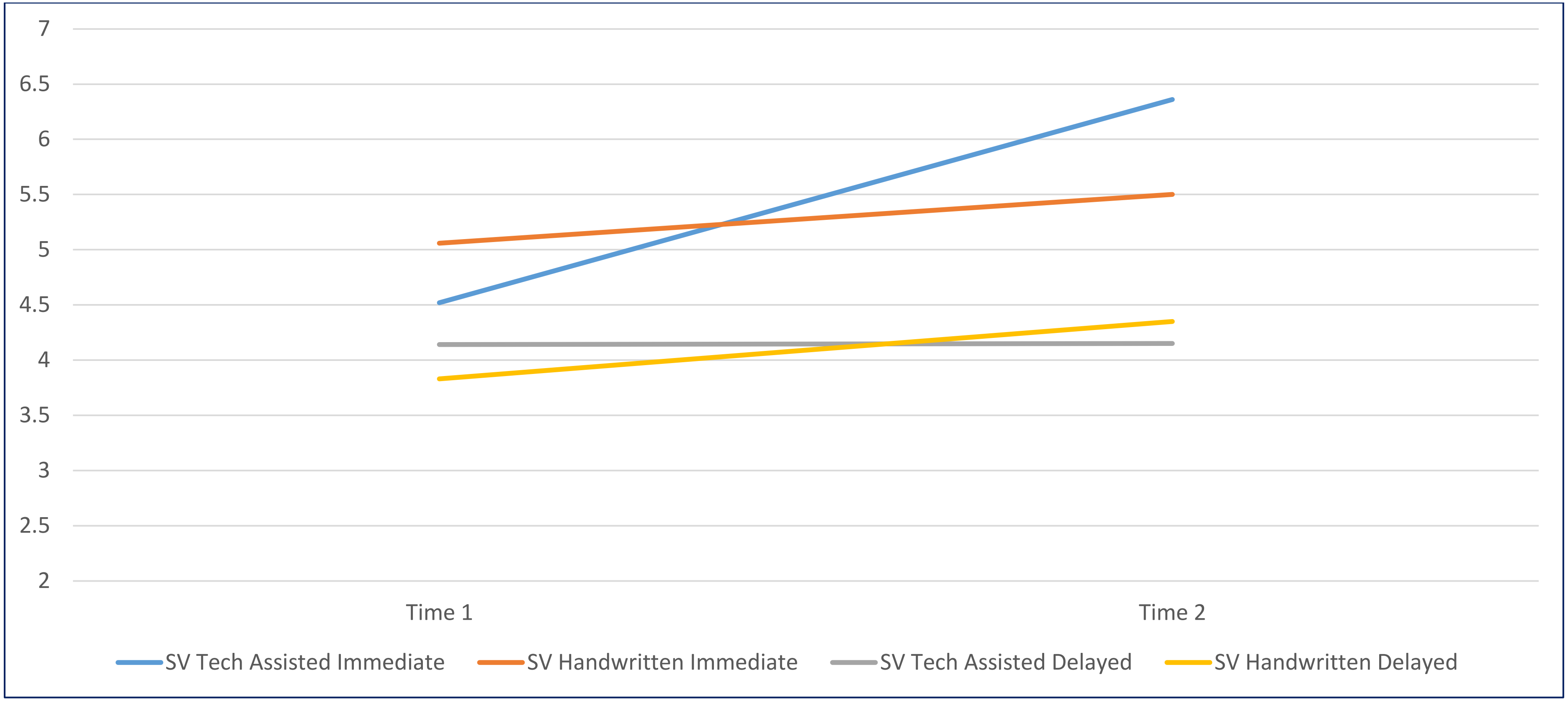


Figure 4: Mean Social Validity Scores at Time 1 and Time 2: Immediate and Delayed Groups

Figure 5: Pre and Post Therapy Examples from one participant on the Functional Assessment Writing

Time 1 (pre therapy) news email:

The name

The

Time 2 (post therapy) news email:

Dear C (name written),

It sounds like you had a lovely time. I am feeling good would be good for you,

Tell me your news,

What have you been doing,

S (name written)

😍😍😍

Time 3 (follow up) news email:

Dad,

Come on you Reds!

Good result. What did you think?

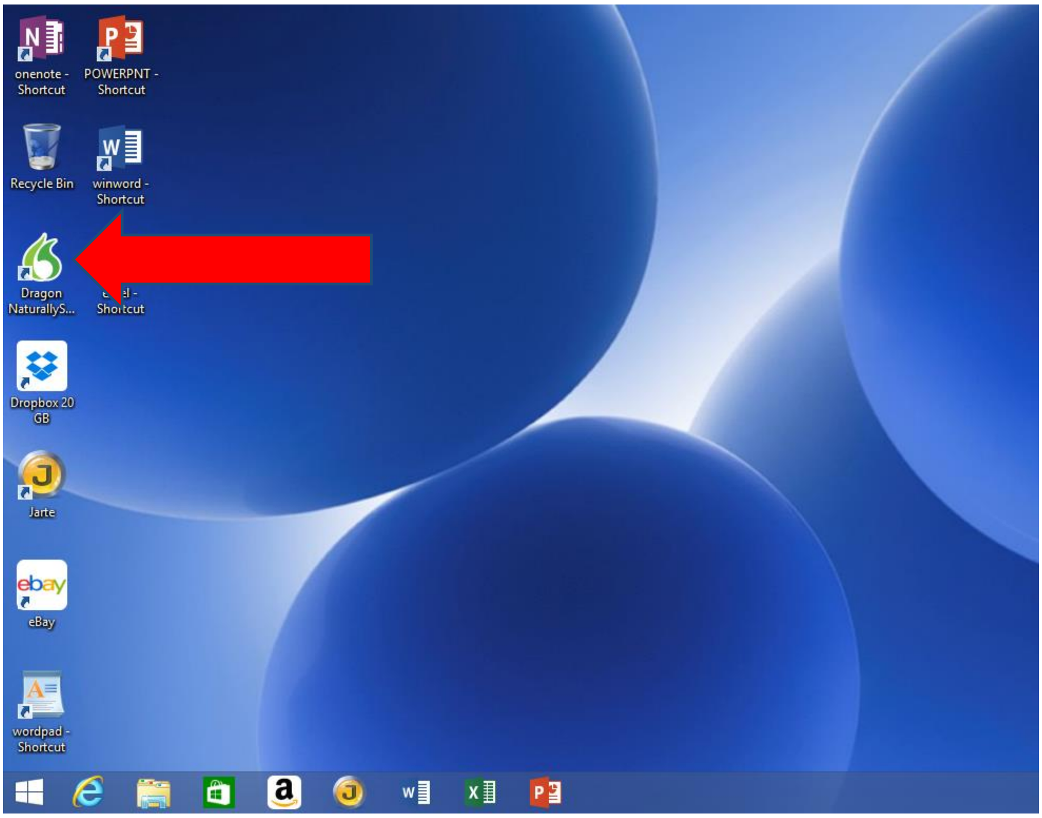
Team played really well. Liverpool 0 v 0 Southampton.

Wednesday 9/5/17

Appendix 1: Sample of Dragon Manual

To Open Dragon:

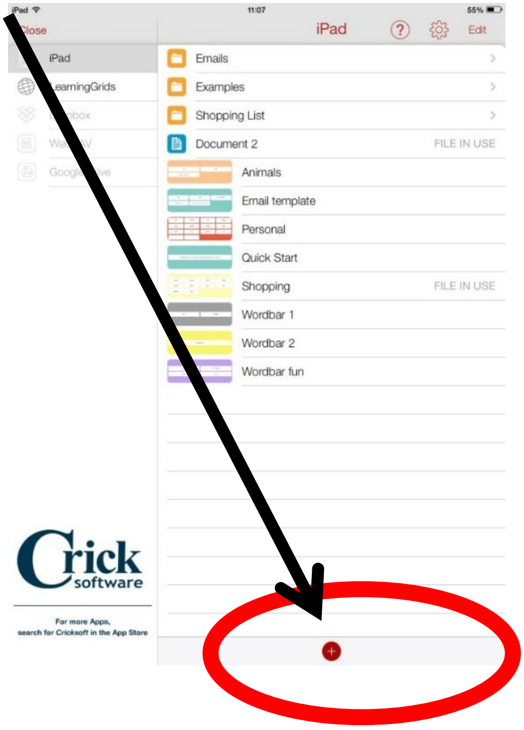
1.) **Double** Click on the Dragonicon on the desktop



Then the Recording Bar will appear:

Creating a **document**

- Press  at the **bottom of the page**



- Press