A Comparison of Lower Intensity Constraint-Induced and Multimodal Therapies in Chronic Aphasia

John Evan Pierce

Bachelor of Speech Pathology

Post Graduate Diploma in Health Research Methodology

A thesis submitted by published work in total fulfilment of the requirements for the

degree of Doctor of Philosophy

Discipline of Speech Pathology,

School of Allied Health, Human Service and Sport

College of Science, Health, and Engineering

La Trobe University, Victoria, 3086 Australia

April 2021

Abstract

Background: Systematic reviews with meta-analyses have shown that speech-language therapy interventions for aphasia after stroke are broadly effective compared to no treatment. However, the comparative effectiveness of different types of intervention has not been adequately established, and the optimal intensity at which to provide intervention is unclear.

Research aims: The aim of this thesis was to compare the efficacy of two established treatment approaches – constraint-induced and multimodal therapy – in people with chronic aphasia. Previously, researchers have typically provided such treatments at high intensity; most often 30 hours over two weeks. This is uncommon in clinical practice where most people with chronic aphasia receive very low intensity treatment. An additional aim of this thesis was therefore to determine the efficacy of lower intensity intervention.

Method: This thesis comprises four investigations, including two systematic reviews, one scoping review and a pilot randomised controlled trial. A systematic review of high-quality studies of constraint-induced and multimodal treatments was conducted, followed by a scoping review of the use of the term "multimodal" in aphasia literature. To explore the importance of scheduling, a second systematic review examined studies which compared different schedules of the same dose and type of therapy. Finally, a pilot randomised controlled trial was conducted with two arms: Constraint-induced Aphasia Therapy Plus and Multi-Modality Aphasia Therapy, both delivered at six hours per week over five weeks.

Results: Significant treatment effects were demonstrated on naming but not on aphasia severity, functional communication or quality of life measures. Both interventions, provided at low-moderate intensity, were equally efficacious.

Conclusions: This thesis provides evidence that neither constraint-induced nor multimodal treatment is superior for naming outcomes and that a low-moderate intensity (i.e., 6 hours per week) of treatment is efficacious. Findings also highlight multiple issues that need to be considered in future aphasia research.

Table of Contents

List of figures

List of Tables

Acknowledgements

To Miranda, thank you for your foolishness in taking on a student who plans to continue working and raising a family and might not finish for 6-8 years; thank you for your wisdom in guiding me through such a long journey without losing my family, job or sanity, and for never pushing me to do more, only less when I needed to hear it.

To Rob, thank you for instantly accepting and believing in me, for your unwavering support and for always coming through with a new angle on our work that none of us saw. You have so much of my respect.

To Leanne, thank you for taking me on without knowing me, despite your considerable workload, and for your consistent warmth and insightfulness over all these years. Your equanimity in the midst of everything has been so reassuring.

Thank you to the COMPARE trial monitors, the chief investigators, the hundreds of trial participants and their families, the assessors and therapists – each of whose labour and time allowed me the opportunity to be part of research on a grand scale.

To Jo, thank you for your long, long patience with me, your support and unquestioning loyalty. There is no doubt at all that I wouldn't have been able to do this without you. I consider it another one of *our* achievements.

Statement of Authorship

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis accepted for the award of any other degree or diploma. No other person's work has been used without due acknowledgment in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

The contributions of the candidate and others to each published work were as follows:

1. Chapter Two, published as *Constraint and multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis*

   − The study was designed by the candidate in collaboration with his supervisors and the candidate conducted the search

   − The candidate and Mrs Menahemi-Falkov jointly reviewed the search results and categorised and rated included studies

   − The candidate was responsible for analysing the data and writing the complete first draft of the manuscript.

   − The candidate's supervisors critically appraised the manuscript and subsequent revisions

2. Chapter Three, published as *What is meant by "Multimodal Therapy" for aphasia?*

- The study was designed by the candidate in collaboration with his supervisors

- The candidate was solely responsible for reviewing the literature and extracting data

- The candidate and his supervisors jointly developed the categorisation system and proposed model

- The candidate was responsible for the complete first draft of the manuscript.

- The candidate's supervisors critically appraised the manuscript and subsequent revisions

3. Chapter Four, published as *Comparing higher and lower weekly treatment intensity for chronic aphasia: A systematic review and meta-analysis*

- The study was designed by the candidate in collaboration with his supervisors and the candidate conducted the search

- The candidate and Mrs Menahemi-Falkov jointly reviewed the search results for inclusion and categorised the included studies

- Professor Rose, Mrs Menahemi-Falkov and the candidate jointly rated included studies

- The candidate was responsible for analysing the data and writing the complete first draft of the manuscript.

- The candidate's supervisors critically appraised the manuscript and subsequent revisions

The contributions of the candidate to the trial outlined in chapter five were as follows:

- The candidate and his supervisors jointly designed elements of the protocol where they differed from the primary trial

- During the trial, the candidate conducted:

  o Two full therapy cohorts (60 hours of therapy)

  o 18 assessments

  o 48 assessment fidelity checks

  o 74 hours of therapy fidelity checks

- Analysis was designed by the candidate in collaboration with his primary supervisor and the trial statistician

- The candidate conducted statistical analyses

All research procedures reported in the thesis were approved by the La Trobe HREC and relevant site-specific HRECs where data was collected.

John Evan Pierce

21st April 2021

Notes on Formatting

This thesis has been formatted as per the APA 7th Edition guidelines. However, chapters 2-4 are presented as Accepted Manuscripts which were formatted in APA 6th, thus there are minor discrepancies in citations and references. To improve readability, these three published chapters were inserted from the word processor rather than as PDF pages and this has resulted in a revised layout. Figures and Tables were also renumbered to clearly indicate the chapter, e.g., Table 1 became Table 2-1. In accordance with the La Trobe Schedule for Presentation of Theses, the content of these chapters was not modified.

Publications and presentations arising during candidature

Peer-reviewed publications

Pierce, J. E., O'Halloran, R., Menahemi-Falkov, M., Togher, L., & Rose, M. L. (2020). Comparing higher and lower weekly treatment intensity for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation, 34*(2), 1–25. https://doi.org/10.1080/09602011.2020.1768127

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2019). What is meant by "Multimodal Therapy" for aphasia? American Journal of Speech-Language Pathology. *American Journal of Speech-Language Pathology, 28*(2), 706-716. http://dx.doi.org/10.1044/2018_AJSLP-18-0157

Rose, M. L., Ali, M., Elders, A., Godwin, J., Sandri, A. K., Williams, L. J., Williams, L. R., VandenBerg, K., Abel, S., Abo, M., Becker, F., Bowen, A., Brandenburg, C., Breitenstein, C., Copland, D., Cranfill, T., Pietro-Bachmann, M. D., Enderby, P., Fillingham, J., … Pierce, J. E., ... Brady, M. C. (2018). Tidier descriptions of speech and language therapy interventions for people with aphasia; consensus from the RELEASE collaboration. *Aphasiology*, *32*(S1), 183–186. https://doi.org/10.1080/02687038.2018.1487021

Pierce, J. E., Menahemi-Falkov, M., O'Halloran, R., Togher, L., & Rose, M. L. (2017). Constraint and multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation, 29*(7), 1005-1041. https://doi.org/10.1080/09602011.2017.1365730

Presentations

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2019, June). *Constraint-Induced and Multimodal approaches to aphasia therapy – what are they and which one leads to better outcomes?* [Paper presentation]. Speech Pathology Subacute Seminar, St Vincent's Hospital, Melbourne, Australia.

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2019, June). *A Systematic Review of Intensity in Chronic Aphasia* [Paper presentation]. Speech Pathology Australia Conference, Brisbane, Australia.

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2018, October). *What do speech pathologists mean by 'multimodal therapy' for aphasia?* [Paper presentation] Aphasiology Symposium of Australia, Sunshine Coast, Queensland, Australia.

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2018, June). *Update on intensity and dose in aphasia rehabilitation* [Paper presentation]. Speech Pathologists in Adult Rehabilitation, Bundoora, Australia.

Pierce, J. E. (2017, June). *Words fail me* [Presentation]. Three Minute Thesis, La Trobe University, Australia.

Pierce, J. E., Nickels, L., Togher, L., Meinzer, M., Rai, T., Godecke, E., Kim, J., Cadilhac, D., Foster, A., Hurley, M., Copland, D., O'Halloran, R., and Rose, M. (2017, August). *Treatment for people with chronic aphasia – investigation of high and low intensity, constraint and multimodal treatments* [Paper presentation]. Smart Strokes Conference, Gold Coast, Australia.

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2017, October). *Revisiting the case for single case research* [Paper presentation]. Cabrini Research week, Melbourne, Australia.

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2016, October). *What is constraint therapy? What is multimodal therapy? What is the difference?* [Paper presentation]. Australian Aphasia Association National Conference, Adelaide, Australia.

Conference posters

Rose, M. L., Carragher, M., Nickels, L., Togher, L., Meinzer, M., Rai, T., Godecke, E., Kim, J., Cadilhac, D. A., Foster, A., Pierce, J. E., Hurley, M., & Copland, D. (2018, May). *Methods of monitoring fidelity in trials of complex behavioural interventions: The COMPARE fidelity protocol* [Poster presentation]. European Stroke Organisation Conference, Gothenburg, Sweden.

Rose, M., Nickels, L., Togher, L., Copland, D., Meinzer, M., Rai, T., Godecke, E., Kim, J., Cadhilac, D., Foster, A., Hurley, M., O'Halloran, R., & Pierce, J. E. (2017, October). *The COMPARE Trial - a national RCT for individuals with chronic post-stroke aphasia* [Poster presentation]. La Trobe Research Week, Melbourne, Australia.

Rose, M., Foster, A., Copland, D., Nickels, L., Togher, L., Meinzer, M., Rai, T., Hurley, M., Pierce, J. E., Godecke, E., Cadilhac, D., Kim, J., & Godecke, E. (2017, May). *Embedding Measures of Therapy Fidelity in "Compare": A Randomised Controlled Trial of Complex Behavioural Interventions for Post-Stroke Aphasia* [Poster presentation]. European Stroke Organisation Conference, Prague, Czech Republic.

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2016, September). *Intensity of aphasia therapy - How intensive does therapy need to be in chronic aphasia?* [Poster presentation]. La Trobe Research Showcase, Melbourne, Australia.

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. (2016, September). *Are multi-modal methods as effective as constraint in aphasia treatment?* [Poster presentation]. Cabrini Research Week, Melbourne, Australia.

Preface

Of all the conditions that a speech pathologist in the medical field works with, stroke is one that results in one of the most instantaneous changes to a person's life. The person has often no warning prior to a stroke and is living an ordinary day. When the stroke occurs, they are teleported from a place of seemingly good health into a lottery, where impairments will largely depend on which artery an embolus bounces into, or in which artery a haemorrhage occurs.

In my opinion, aphasia can be a most uniquely devastating condition. While I have worked with patients with anarthria from motor neurone disease, with aphagia from cancer, with cognitive impairment from dementia and with anatomical alterations of the face, mouth and neck after surgery, none of these seemed to challenge a person's identity in quite the same manner as aphasia. A study in Canada of 66,000 quality of life measures from people in long term care found poorer quality of life in people with aphasia compared to every other medical condition, including cancer, dementia and quadriplegia[1]. Why might this be? Communication allows us to connect with others. Sometimes there are no good treatment options for a condition and the only recourse is to be able to share the experience with others — talk to family, join a support group, write a book. In aphasia, the ability to do these things is often reduced or even lost, and perhaps this makes a health condition more difficult to bear. I am not suggesting that there is any scale on which experiences of illness can truly be compared,

---

[1] Lam, J. M. C., & Wodchis, W. P. (2010). The Relationship of 60 Disease Diagnoses and 15 Conditions to Preference-Based Health-Related Quality of Life in Ontario Hospital-Based Long-Term Care Residents. *Medical Care*, *48*(4), 380–387. https://doi.org/10.1097/mlr.0b013e3181ca2647

but instead explaining why my drive to do more and know more was strongest for my patients with aphasia.

In my position in community-based rehabilitation, some people with aphasia would present annually for a block of treatment. Some had clear functional goals, but others wanted impairment-based therapy for general word retrieval, sentence production, or reading and spelling, as well as direction for future self-practice. They believed in lifelong rehabilitation with sufficient practice. However, I could not find clear research guidance on which treatment approach to use, how much therapy to provide and how often.

My attempts to describe the impact of aphasia has painted a dark picture, yet in parallel I cannot express how strong these individuals are. They persevere, they do what they can, accept what they can't, and adjust. It is for them that I started this research journey and I hope that, even in a small way, the work can help discover the treatments that best improve communication and quality of life.

Chapter 1 – Introduction and overview

Stroke is the third most common cause of death in Australia (Australian Institute of Health and Welfare, 2019) and accounts for approximately 4.5% of national disease burden, in terms of healthy years and total years lost (Australian Institute of Health and Welfare, 2013). The direct economic cost of stroke in Australia in 2012 was estimated at $5 billion with $49 billion in additional disease burden, with over 420,000 people living with effects from stroke and two thirds of these impacted in their ability to complete normal activities (Deloitte Access Economics, 2013).

Aphasia is a language impairment that occurs following brain damage to the left cerebral hemisphere, occurring in over one third of stroke survivors (Dickey et al., 2010). The effects of aphasia on individuals can be profound and far-reaching, with poor outcomes in return to employment (Flowers et al., 2016), negative changes to family relationships and loss of friendships (Dalemans et al., 2010), reduced quality of life and high rates of depression (Hilari et al., 2003; Hilari, 2011). Family members and carers of people with aphasia also experience significant negative effects such as loneliness, anxiety, increased responsibilities and need for assistance (Patrício et al., 2013), lost income and employment and limited time for social activity (Grawburg et al., 2019). The cost of healthcare to stroke patients with aphasia is significantly elevated in comparison to stroke without aphasia (Ellis et al., 2012). In addition, despite some natural recovery of symptoms in the early stages, the majority of people who develop aphasia after stroke do not have total resolution of their symptoms—in one study, only one quarter of cases fully recovered after 18 months (Flowers et al., 2016). For this reason, most people with aphasia are unable to return to their previous occupation, or to work of any kind (Dalemans et al., 2010; Graham et al., 2011).

Such broad negative impacts mean that treatments to reduce the effects of aphasia can be highly beneficial to people with aphasia, family members and society at large. Recently, a Cochrane systematic review found sufficient high-quality evidence to show superiority of aphasia treatment over no treatment in meta-analysis (Brady et al., 2016). While this finding was encouraging, it lacked specificity as a result of synthesising papers with heterogeneous methods. Across the included papers, participants were at different phases of recovery, given different treatments at a range of schedules and doses, and tested with an array of different outcome measures. The authors thus called for further research in more specific aspects of aphasia intervention, including two aspects that are the focus of this thesis: (a) comparative effectiveness of different treatment approaches and (b) intensity of treatment.

Topic A – Comparative effectiveness

Two key treatment approaches for aphasia are investigated in this thesis: multimodal and constraint-induced. Both are prominent approaches in aphasia research and clinical practice and distinct from one another in terms of underlying theoretical rationales (Pierce et al., 2017). A recent collaborative project reviewed a large dataset of aphasia research and classified the reported treatment approaches (RELEASE Collaboration, 2020). The final categorisation structure was comprised of nine overarching theoretical approaches including multimodal and Constraint. Each is outlined below.

Multimodal

"We should apply the principle that speech is a total bodily response

and accompanies motor activity. It is a known fact that some

spontaneous speech often occurs in aphasics when they are doing

things." (Backus, 1945)

In this work, multimodal treatment refers to treatments that employ one or more relatively less impaired, non-verbal modalities in order to facilitate verbal language production. The non-verbal modality (or modalities) may be incorporated into any part of the treatment, whether as stimuli, therapist cueing, patient self-cueing or patient response. Multimodal approaches to therapy have a long history of use in aphasia, with one case report from 1879 describing facilitation of speech via rhythm and melody (Sharpey, as cited in Howard & Hatfield, 1987). However, the earliest published *experimental* research on multimodal treatment was published nearly a century later (Sparks et al., 1974). It described a now widely known treatment, Melodic Intonation Therapy (MIT), where the intact modality of singing is used with the aim of improving propositional speech. MIT is just one manifestation of a range of singing treatments in aphasia. Gesture is also commonly employed to promote recovery of verbal skills (Rose et al., 2013). Drawing is less often reported but has been explored (Hung & Ostergren, 2019), while reading and written expression are employed as cues in aphasia intervention almost ubiquitously (Lorenz & Nickels, 2007). Treatments may also employ multiple non-verbal modalities to facilitate verbal output and/or support language relearning. Multi-Modality Aphasia Therapy (M-MAT) uses gesture, drawing, reading and writing in sequence to cue word retrieval and phrase and sentence production (Attard et al., 2013).

These and other manifestations of multimodal approaches invariably describe some form of links or interconnections between systems as their underlying mechanism. For example, Luria (1970) described *intersystemic reorganisation* where

impaired systems can be supported and facilitated by related, more intact systems and this concept is frequently cited as a possible mechanism in multimodal treatments. In the 1960s, Weigl described deblocking, a process in which producing a word in a relatively spared modality (e.g. writing) appears to facilitate oral production of the same word (Basso, 2003). More recently, embodied language has been proposed and supported experimentally with strong action-language connections in functional imaging studies (Fischer & Zwaan, 2008). For example, upon processing of a spoken verb there is near-simultaneous activation of the analogous representation in the motor cortex, e.g., hearing the word "kick" → lower limb activation (Pulvermüller & Berthier, 2008). Given this coactivation, pairing of action with language in the alternate direction (i.e., action followed by language) is proposed as one method to facilitate recovery in aphasia. The above explanations describe weaving of apparently distinct brain processes into one another (Pulvermüller & Berthier, 2008) and contrast with models where language is viewed as a modular, standalone system that receives inputs and produces outputs in a serial fashion (Fischer & Zwaan, 2008; Pulvermüller & Berthier, 2008).

Orthographic and graphemic cues are employed in the majority of aphasia treatments (Lorenz & Nickels, 2007; Thomson, 2012), demonstrating the implicit acceptance of multimodal cueing in the field of speech language pathology, at least for the modalities of reading and writing. Nonetheless, the research base and comparative effectiveness of the multimodal approach requires examination.

Constraint

"Stroke recovery: he can but does he?" (Andrews & Stewart, 1979)

Understanding constraint-induced approaches to aphasia rehabilitation requires a brief review of its origins and history, beginning with the concept of learned non-use and Constraint-induced Movement Therapy.

In the 1970s, experiments on monkeys with unilaterally deafferented limbs repeatedly demonstrated that they would not use the impaired limb despite the physical ability to do so (Taub, 1976) . This disparity between the capacity for movement and the performance in everyday function was termed *learned non-use* to signal the behavioural component of the phenomenon (Taub et al., 2006). Learned non-use, in monkeys and humans, is thought to arise from both positive reinforcement of avoidance (e.g., higher success and efficiency using the intact limb) and negative reinforcement of any use of the impaired limb, such as increased effort of movement and dropping of handled items. A compounding negative feedback loop is then commenced where the cortical area available for the impaired limb shrinks as a result of low use, further negatively reinforcing any impaired limb use, and so on (Taub et al., 2006). However, further experiments with monkeys demonstrated that use of the impaired limbs could be retaught. Various methods of forcing or 'constraining' the monkeys into using the limb were employed, including straitjackets to restrain the intact limb and electric shock (Andrews & Stewart, 1979).

Although the ends do not justify the means in terms of animal cruelty, the above research did result in a beneficial treatment for humans. Constraint-Induced Movement

Therapy (CIMT) was built on the findings of the above research. CIMT is designed around three key principles (Andrews & Stewart, 1979):

i) High intensity training

ii) Real-world task simulations to encourage generalisation

iii) Restriction of the unimpaired limb (constraint).

CIMT now has a strong evidence base for upper limb rehabilitation in stroke (Peurala et al., 2012) as well as other aetiologies and motor functions (Dong et al., 2013; Taub et al., 2006).

"In motor therapy, the arm can be constrained by a sling, but how can [speech] articulation be induced by constraints?" (Pulvermüller et al., 2001)

Constraint-Induced Aphasia Therapy (CIAT) was designed as a language analogue to CIMT with the above three principles interpreted accordingly (Pulvermüller et al., 2001). In the seminal CIAT paper, (i) high intensity training was applied, at three hours per day, five days per week for two weeks; and (ii) treatment focused on actions proposed to be relevant to real life, with small groups of participants requesting common objects from one another through card games. The third principle (iii), that of constraint to the impaired limb, is not as obviously applicable to language. However, the authors (Pulvermüller et al., 2001) drew parallels between learned non-use in motor function and language. While the brain does not simply have a fully functional contralateral language area to fall back on as it does in limb function, they proposed that learned non-use occurs in aphasia with the avoidance of challenging verbal

expressions in favour of simpler and overlearned ones, or with favouring non-verbal communication over verbal expression (Pulvermüller & Berthier, 2008). Constraint was therefore described as those parts of therapy that forced patients to employ language actions that they would normally avoid (Pulvermüller et al., 2001). Constraint in CIAT included:

- Limiting communication to verbal communication only, through use of visual barriers between patients that negated the communicative utility of any gestures or other nonverbal communication produced.

- Requiring specific grammatical structures in participant utterances.

- Using target stimuli that require specific language to accurately request.

- Shaping responses from simple to more complex utterances over time.

Despite its recent conception (2001), CIAT has become widely known and frequently researched (Pierce et al., 2017). Various authors have reinterpreted the principle of constraint within CIAT and given the treatment new labels. "Constraint-Induced *Language* Therapy" was favoured by some researchers and is now largely synonymous with CIAT (Kurland et al., 2010; Maher et al., 2006). In 2008, some of the original CIAT authors proposed renaming CIAT Intensive Language Action Therapy to avoid negative connotations with concepts of restraint and forced use (Pulvermüller & Berthier, 2008). "CIAT Plus" added written language to the original protocol as well as a home program to encourage transfer of treated targets (Meinzer et al., 2005). CIAT II was designed to more closely align with CIMT (Johnson et al., 2014) by a research group that also included an original author of both CIAT and CIMT, Dr Edward Taub. Modifications included a larger range of games and tasks, including role play, a higher

number of utterances required per hour, the training of relatives in communication support behaviours, and a home transfer package. In addition to these protocolised variants of CIAT, there are variations in therapy protocols between studies with the *same* label and these are explored in Chapter Two.

Why compare multimodal and constraint?

When compared side by side, multimodal and constraint approaches may superficially appear to have only minor practical differences in the therapy room. However, they represent major differences in the conceptualisation of language representations and processing in the brain. Constraint treatment, at least in many studies, holds that use of non-verbal modalities interfere with the opportunity for massed practice of impaired speech and language, and thus neuroplastic change (Pulvermüller et al., 2001). In direct contrast to this 'competing systems' view, multimodal treatments are based on language being one network situated within broader brain networks, and assert that different circuits of different communication functions are nonetheless interconnected and might have synergistic effects that can be harnessed in therapy (Rose, 2013).

Studying the comparative effectiveness of these approaches is therefore important not only to inform clinical practice—which treatments are more effective and for whom—but also to provide information about language processing networks in the brain.

Topic B – Intensity of treatment

The second component covered in this thesis is Intensity of Treatment. How much therapy to provide, and how frequently, in order to have the best outcomes for people with aphasia, is recognised as a knowledge gap that needs to be addressed (Brady et al., 2016). Scheduling is "essential to design and implementation of any treatment program for aphasia" (Cherney et al., 2011, p. 560) and is important information for clinicians, patients, family and service providers. For example, in a consensus study in Canada, stroke survivors, their caregivers and family, and health professionals were asked to prioritise research areas across stroke rehabilitation. Ideal timing and intensity of aphasia intervention was ranked third out of all areas, behind only community reintegration and severe stroke recovery (Bayley et al., 2007).

There are multiple components to scheduling which have not been well separated in previous research, and there are no widely agreed definitions or terminology across rehabilitation or even aphasia. More recently, elements have been described and labelled with reference to pharmaceutical aspects of dose (Warren et al., 2007) and Baker (2012) proposed that the terminology and system can be applied to describe speech pathology treatments. These elements include the duration of a therapy session, the weekly frequency of sessions and the total duration of the treatment in weeks or months, but the author also called for more fine-grained description of the treatment by defining the 'active ingredient' and measuring its occurrences within a session. This allows a count of the total number of active ingredients across the treatment course. In this work we focus on the frequency/distribution of treatment rather than the dose provided, and we refer to the frequency as Intensity. 'Intensity' is a term which is used loosely in rehabilitation and

aphasia (Baker, 2012): often used to describe the frequency of treatment (e.g., Sage et al., 2011); often for the overarching concept of scheduling (Baker, 2012); and by some, the amount of effort required of the patient (Harnish et al., 2008). However, we feel that the first definition of intensity is most intuitively understood by readers—both clinicians and researchers.

Intensity in this thesis is therefore specified, using the first definition described above, as the quotient of the dose of treatment over the duration, as depicted in Figure 1-1, where *dose* is the amount of treatment and *duration* is the time over which it is provided. This formula can apply to the duration of treatment per day or per week (e.g. 5 hours per week), or the total number of hours over the total treatment duration (e.g. 20 hours over 4 weeks). It can be seen that intensity is inextricably linked with therapy dose and duration. Any change in the dose of treatment, with duration held constant, would result in higher or lower intensity accordingly. However, it is the inverse alteration—the effect of varying the duration of treatment whilst keeping the dose constant—that is the key focus of this thesis. There are contrasting theoretical arguments for the effectiveness of both increasing and decreasing the intensity in this manner that are briefly introduced below.

Figure 1-1 – Definition of "intensity" used within this thesis
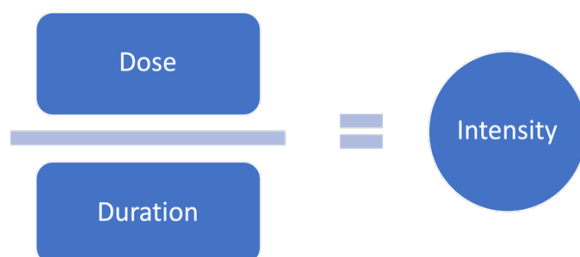
Neuroplasticity → higher intensity

> *"What is the point of such short lessons? Could one believe that a child*
>
> *would make progress, if one only made him speak for a few minutes*
>
> *each day?" (Paul Broca,1865 as cited in Howard & Hatfield, 1987)*

Neuroplasticity—the ability of the brain to adapt and remodel neuronal cells and connective tracts in response to information—was traditionally considered only a developmental phenomenon but was shown, from the latter 20th century onward, to occur across the lifespan (Kolb et al., 1999). Neuroplasticity occurs in normal learning as well as rehabilitation following brain damage, wherever sensory, behavioural and cognitive experiences repeatedly occur (Kleim & Jones, 2008). Brain reorganisation takes place spontaneously following damage, even without any rehabilitation input, but basic research within neuroscience has determined principles important in promoting additional neuronal change (Kleim & Jones, 2008). One accepted principle of neuroplasticity is that "Intensity Matters" (Kleim & Jones, 2008). This is suggested by experiments where high numbers of repetition of motor tasks or electrical pulses were more effective in producing neuron-level changes than lower numbers (Kleim & Jones, 2008). However, the majority of these experiments were undertaken on motor tasks and mainly with animals (Raymer et al., 2008). While brain activation changes following aphasia treatment (Breier et al., 2006, 2009; Marcotte et al., 2018), it does not necessarily follow that the neuroplasticity principle of intensity applies to language recovery in a simple fashion. For example, it is not yet clear that a) animal models are adequately predictive of human responses to intensity, or b) motor tasks are sufficiently comparable to more complex tasks such as language (Raymer et al., 2008). Finally, while

it is difficult to argue against the existence of a minimum intensity threshold to produce brain change, it does not follow that increased intensity will result in superior treatment effects in a linear fashion. In fact, evidence from cognitive psychology may provide an argument for the reverse effect (Cepeda et al., 2006), as discussed below.

Cognitive Psychology → lower intensity

There is ample evidence that overall, distributing learning events over a longer period (viz., lower intensity) results in greater recall for new learning and this is particularly effective for long term recall (Dignam, Rodriguez, et al., 2016). This has been shown in motor and cognitive tasks, but crucially, also in verbal learning (Dignam, Rodriguez, et al., 2016). In a meta-analysis of verbal learning research with neurotypical participants, 271 experiments were reviewed which compared different distributions of practice for verbal recall (Cepeda et al., 2006). Overall, spaced practice resulted in significantly better recall compared to no spacing (t = 6.6, p < .001) and this was true whether recall testing occurred after less than one minute or after 31 days. When the relationship between different spacing between learning sessions and recall was analysed, the authors concluded that there was a positive influence of increased spacing on recall, but this was not a simple monotonic relationship and that there was a drop-off in effectiveness after particular intervals. It is important to note that only 31 of the 271 experiments used spacing of one day or greater, and so the translation of these findings to the commonly employed weekly or monthly distribution of aphasia treatment sessions cannot be assumed. In addition, the review only included studies of neurotypical participants.

Rehabilitation is typically believed to be a form of learning (Dignam, Rodriguez, et al., 2016), and one recent study did find a correlation between participants' performance on new word learning and treatment outcomes for anomia (Dignam, Copland, et al., 2016). Nonetheless, the applicability of principles for learning new verbal information—including scheduling principles—to rehabilitating language is not yet clear (Middleton et al., 2020). The neurotypical brain and the impaired brain may differ in the mechanisms of learning, and post stroke cognitive impairment may preclude the effectiveness of some principles (Middleton et al., 2020).

Comparing higher and lower intensity

"Thus, the research direction is clear—to systematically investigate the

effects of treatment intensity independently and in combination with

influencing factors" (Cherney et al., 2011, p. 565)

Currently, very little is known about ideal scheduling and dose for aphasia therapy, even on a macro scale (Brady et al., 2016). Clinical Practice Guidelines frequently recommend intensive treatment for aphasia without stipulating any numbers or thresholds (Dignam, Rodriguez, et al., 2016), but even these non-specific recommendations are based on systematic reviews that compare different durations and doses of treatment (e.g. Bhogal et al., 2003; Brady et al., 2016), and thus the effect of amount of therapy and frequency of therapy are confounded.

As stated earlier, the Cochrane Review by Brady et al. (2016) called for more research into optimum intensity for aphasia treatment and this is just one of many appeals—a great number of papers have made the same call for more investigation in

intensity for aphasia (Cherney et al., 2011; Dignam et al., 2015; Dignam, Rodriguez, et al., 2016; Middleton et al., 2020; Mozeiko et al., 2015; Sage et al., 2011).

This thesis examines questions which interrogate the issue of intensity in the setting of chronic aphasia. Evidence supporting higher or lower intensity of treatment is important in informing research schedules and for clinical practice, which are currently provided at very low intensity according to a number of surveys (Code & Heron, 2003; Palmer et al., 2018; Verna et al., 2009).

People with chronic aphasia as the target population of this thesis

There are indications from neuroimaging studies that neural changes in stroke recovery, and language recovery specifically, differ depending on the stage of recovery (Raymer et al., 2008). Effect sizes may differ between acute, subacute and chronic phases for various aphasia treatment types and other variables including scheduling. Consequently, it is important to separate or at least control for recovery phase in research. This thesis focuses on people with chronic aphasia because (a) these individuals have a need for interventions to improve language function, and (b) the original motivation for the doctorate was based on patients with chronic aphasia (see Preface).

Purpose of this research

The overall aims of this thesis are:

- To understand and expand upon the evidence for (1) constraint-induced aphasia interventions and (2) Multimodal aphasia interventions on

outcomes of impairment, activity/ participation and quality of life, in chronic aphasia

- To compare the outcomes of constraint-induced and multimodal aphasia interventions

- To explore the influence of intervention intensity on treatment outcomes in chronic aphasia.

Breakdown of thesis structure

There are six chapters in this thesis, including this introduction, three of which are presented in their published form (with minor formatting changes only). Figure 1-2 illustrates the structure visually.

Figure 1-2 – Thesis structure

*Published in peer-reviewed journal
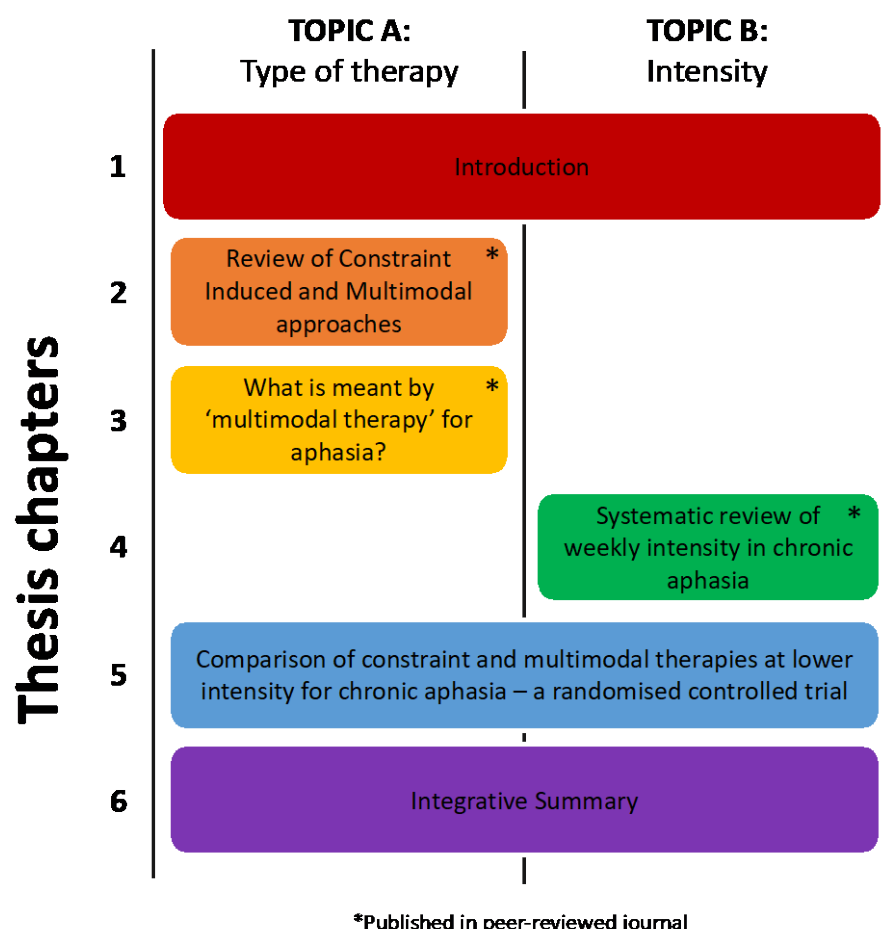
Chapter 2 presents a published systematic review and meta-analysis which examined and compared constraint-induced and multimodal approaches to therapy. This review was published in Neuropsychological Rehabilitation (Pierce et al., 2017). Using a relatively novel method of meta-analysis for Single Case Experimental Designs, Tau-U, this work was able to compare effect sizes for each approach. The gaps in the literature in terms of methods, rigour and outcome types were also outlined, forming a rationale for the experiment conducted.

Chapter 3 introduces the emergence of a paper from the production of the systematic review in Chapter 2 which was not originally anticipated. During the search and screening process for the Chapter 2 systematic review, the authors discovered that the term 'multimodal therapy' was used for a range of disparate treatments, seemingly without any common features. To address this problem, a scoping review was undertaken to describe current use of the term within the field of aphasia and propose inclusion and exclusion criteria for future use. The resulting work was published in the American Journal of Speech Language Pathology (Pierce et al., 2019) and is reproduced as part of Chapter 3.

In Chapter 4, the focus shifts to the second key question of the thesis, that of intensity. A published systematic review and meta-analysis is presented which examined articles directly comparing higher and lower weekly intensity (Pierce et al., 2020). The findings illustrate the progress made on the question within the chronic phase specifically and the need for more investigation into this factor.

An experimental study is outlined in Chapter 5. This pilot Randomised Controlled Trial (RCT) was a substudy of a larger RCT and its relationship with the 'parent' trial is explained. CIAT Plus and M-MAT were chosen as the specific treatments within each

approach. As will be outlined, they are operationally symmetrical in many ways, allowing comparison of the therapy types while controlling for other extraneous variables; for example, both treatments have syntactically identical target utterances, use the same language games and are provided in a group setting. The study's method and results are presented and then discussed.

Finally, the sixth chapter of this thesis considers the outcomes of the work within this thesis in terms of the two main topics - the relative efficacy of constraint and multimodal approaches, and the exploration of treatment intensity. Each is discussed in terms of implications for research and clinical practice. A number of additional issues raised by the results of this work are also considered. The study strengths and limitations are reviewed as well as recommendations for ongoing research on these questions.

References

Andrews, K., & Stewart, J. (1979). Stroke recovery: He can but does he? *Rheumatology and Rehabilitation*, *18*(1), 43–48. https://doi.org/10.1093/rheumatology/18.1.43

Australian Institute of Health and Welfare. (2013). Stroke and its management in Australia: An update. *Cardiovascular Disease Series*, *37*(Cat. no. CVD 61). https://www.aihw.gov.au/getmedia/3d56c949-68a4-46f3-bc7c-c40c89904d38/13994.pdf.aspx?inline=true

Australian Institute of Health and Welfare. (2019). *Deaths in Australia* [Web Report]. Deaths in Australia. https://www.aihw.gov.au/reports/life-expectancy-death/deaths-in-australia

Attard, M. C., Rose, M. L., & Lanyon, L. E. (2013). The comparative effects of Multi-Modality Aphasia Therapy and Constraint-Induced Aphasia Therapy-Plus for severe chronic Broca's aphasia: An in-depth pilot study. *Aphasiology, 27*(1), 80–111. https://doi.org/10.1080/02687038.2012.725242

Backus, O. L. (1945). The rehabilitation of aphasic Veterans. *Journal of Speech Disorders*, *10*(2), 149–153. https://doi.org/10.1044/jshd.1002.149

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

Basso, A. (2003). *Aphasia and Its Therapy*. Oxford University Press.

Bayley, M. T., Hurdowar, A., Teasell, R., Wood-Dauphinee, S., Korner-Bitensky, N., Richards, C. L., Harrison, M., & Jutai, J. W. (2007). Priorities for stroke rehabilitation and research: results of a 2003 Canadian Stroke Network consensus conference. *Archives of Physical Medicine and Rehabilitation*, *88*(4), 526–528. https://doi.org/10.1016/j.apmr.2007.01.005

Bhogal, S. K., Teasell, R., & Speechley, M. (2003). Intensity of aphasia therapy, impact on recovery. *Stroke; a Journal of Cerebral Circulation*, *34*(4), 987–993. https://doi.org/10.1161/01.str.0000062343.64383.d0

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *The Cochrane Database of Systematic Reviews*, *6*(6), CD000425–CD000425. https://doi.org/10.1002/14651858.cd000425.pub4

Breier, J. I., Juranek, J., Maher, L. M., Schmadeke, S., Men, D., & Papanicolaou, A. C. (2009). Behavioral and neurophysiologic response to therapy for chronic aphasia. *Archives of Physical Medicine and Rehabilitation*, *90*(12), 2026–2033. https://doi.org/10.1016/j.apmr.2009.08.144

Breier, J. I., Maher, L. M., Novak, B., & Papanicolaou, A. C. (2006). Functional imaging before and after constraint-induced language therapy for aphasia using magnetoencephalography. *Neurocase*, *12*(6), 322–331. https://doi.org/10.1080/13554790601126054

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Cherney, L. R., Patterson, J. P., & Raymer, A. M. (2011). Intensity of aphasia therapy: evidence and efficacy. *Current Neurology and Neuroscience Reports*, *11*(6), 560–569. https://doi.org/10.1007/s11910-011-0227-6

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), 1231–1237. https://doi.org/10.1080/09638280310001599961

Dalemans, R. J. P., P., D. W., L., Wade, D. T., & Van den Heuvel, W. J. A. (2010). A description of social participation in working-age persons with aphasia: A review of the literature. *Aphasiology*, *22*(10), 1071–1091. https://doi.org/10.1080/02687030701632179

Deloitte Access Economics. (2013). *The economic impact of stroke in Australia*. Stroke Foundation.

Dickey, L., Kagan, A., Lindsay, M. P., Fang, J., Rowland, A., & Black, S. (2010). Incidence and profile of inpatient stroke-induced aphasia in Ontario, Canada. *Archives of Physical Medicine and Rehabilitation*, *91*(2), 196–202. https://doi.org/10.1016/j.apmr.2009.09.020

Dignam, J. K., Copland, D., McKinnon, E., Burfein, P., O'Brien, K., Farrell, A., & Rodriguez, A. (2015). Intensive Versus Distributed Aphasia Therapy A Nonrandomized, Parallel-Group, Dosage-Controlled Study. *Stroke*, *46*(8), 2206–2211. https://doi.org/10.1161/strokeaha.115.009522/-/dc1

Dignam, J. K., Copland, D., Rawlings, A., O'Brien, K., Burfein, P., & Rodriguez, A. D. (2016). The relationship between novel word learning and anomia treatment success in adults with chronic aphasia. *Neuropsychologia*, *81*, 186–197. https://doi.org/10.1016/j.neuropsychologia.2015.12.026

Dignam, J. K., Rodriguez, A. D., & Copland, D. (2016). Evidence for Intensive Aphasia Therapy: Consideration of Theories from Neuroscience and Cognitive Psychology. *American Academy of Physical Medicine and Rehabilitation*, *8*(3), 254–267. https://doi.org/10.1016/j.pmrj.2015.06.010

Dong, V. A.-Q., Tung, I. H.-H., Siu, H. W.-Y., & Fong, K. N.-K. (2013). Studies comparing the efficacy of constraint-induced movement therapy and bimanual training in children with unilateral cerebral palsy: a systematic review. *Developmental Neurorehabilitation*, *16*(2), 133–143. https://doi.org/10.3109/17518423.2012.702136

Ellis, C., Simpson, A. N., Bonilha, H., Mauldin, P. D., & Simpson, K. N. (2012). The one-year attributable cost of poststroke aphasia. *Stroke; a Journal of Cerebral Circulation*, *43*(5), 1429–1431. https://doi.org/10.1161/strokeaha.111.647339

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of*

Experimental Psychology, *61*(6), 825–850.
https://doi.org/10.1080/17470210701623605

Flowers, H. L., Skoretz, S. A., Silver, F. L., Rochon, E., Fang, J., Flamand-Roze, C., & Martino, R. (2016). Poststroke Aphasia Frequency, Recovery, and Outcomes: A Systematic Review and Meta-Analysis. *Archives of Physical Medicine and Rehabilitation*, *97*(12), 2188-2201.e8.
https://doi.org/10.1016/j.apmr.2016.03.006

Graham, J. R., Pereira, S., & Teasell, R. (2011). Aphasia and return to work in younger stroke survivors. *Aphasiology*, *25*(8), 952–960.
https://doi.org/10.1080/02687038.2011.563861

Grawburg, M., Howe, T., Worrall, L. E., & Scarinci, N. (2019). Family-Centered Care in Aphasia: Assessment of Third-Party Disability in Family Members With the Family Aphasia Measure of Life Impact. *Topics in Language Disorders*, *39*(1), 29–54. https://doi.org/10.1097/tld.0000000000000176

Harnish, S. M., Neils-Strunjas, J., Lamy, M., & Eliassen, J. C. (2008). Use of fMRI in the study of chronic aphasia recovery after therapy: a case study. *Topics in Stroke Rehabilitation*, *15*(5), 468–483. https://doi.org/10.1310/tsr1505-468

Hilari, K. (2011). The impact of stroke: are people with aphasia different to those without? *Disability and Rehabilitation*, *33*(3), 211–218.
https://doi.org/10.3109/09638288.2010.508829

Hilari, K., Byng, S., Lamping, D. L., & Smith, S. C. (2003). Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of Acceptability, Reliability, and Validity. *Stroke*, *34*(8), 1944–1950. https://doi.org/10.1161/01.str.0000081987.46660.ed

Howard, D., & Hatfield, F. M. (1987). *Aphasia therapy: Historical and contemporary issues*. Lawrence Erlbaum Associates Ltd.

Hung, P.-F., & Ostergren, J. (2019). A comparison of drawing and writing on facilitating word retrieval in individuals with aphasia. *Aphasiology*, *33*(12), 1462–1481.
https://doi.org/10.1080/02687038.2019.1602861

Johnson, M. L., Taub, E., Harper, L. H., Wade, J. T., Bowman, M. H., Bishop-McKay, S., Haddad, M. M., Mark, V. W., & Uswatte, G. (2014). An enhanced protocol for constraint-induced aphasia therapy II: A case series. *American Journal of Speech-Language Pathology*, *23*(1), 60–72. https://doi.org/10.1044/1058-0360(2013/12-0168)

Kleim, J. A., & Jones, T. A. (2008). Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research*, *51*(1), S225-39. https://doi.org/10.1044/1092-4388(2008/018)

Kolb, B., Gibb, R., & Robertson, I. (1999). Neuroplasticity and recovery of function after brain injury. In G. Stuss, D. Winocur, & I. Robertson (Eds.), *Cognitive Neurorehabilitation* (pp. 9–25). Cambridge University Press.

Kurland, J., Baldwin, K., & Tauer, C. (2010). Treatment-induced neuroplasticity following intensive naming therapy in a case of chronic Wernicke's aphasia. *Aphasiology*, *24*(6–8), 737–751. https://doi.org/10.1080/02687030903524711

Lorenz, A., & Nickels, L. A. (2007). Orthographic cueing in anomic aphasia: How does it work? *Aphasiology*, *21*(6–8), 670–686. https://doi.org/10.1080/02687030701192182

Luria, A. R. (1970). *Traumatic aphasia: Its syndromes, psychology and treatment* (D. Bowden, Trans.). Mouton & Co.

Maher, L. M., Kendall, D. L., Swearengin, J. A., Rodriguez, A., Leon, S. A., Pingel, K., Holland, A. L., & J, G.-R., Leslie. (2006). A pilot study of use-dependent learning in the context of Constraint Induced Language Therapy. *Journal of the International Neuropsychological Society*, *12*(6), 843–852. https://doi.org/10.1017/s1355617706061029

Marcotte, K., Laird, L., Bitan, T., Meltzer, J. A., Graham, S. J., Leonard, C., & Rochon, E. (2018). Therapy-Induced Neuroplasticity in Chronic Aphasia After Phonological Component Analysis: A Matter of Intensity. *Frontiers in Neurology*, *9*, 225–227. https://doi.org/10.3389/fneur.2018.00225

Meinzer, M., Djundja, D., Barthel, G., Elbert, T., & Rockstroh, B. (2005). Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy. *Stroke*, *36*(7), 1462–1466. https://doi.org/10.1161/01.str.0000169941.29831.2a

Middleton, E. L., Schuchard, J., & Rawson, K. A. (2020). A Review of the Application of Distributed Practice Principles to Naming Treatment in Aphasia. *Topics in Language Disorders*, *40*(1), 36–53. https://doi.org/10.1097/tld.0000000000000202

Mozeiko, J., Coelho, C. A., & Myers, E. B. (2015). The role of intensity in constraint-induced language therapy for people with chronic aphasia. *Aphasiology*, *30*(4), 339–363. https://doi.org/10.1080/02687038.2015.1070949

Palmer, R., Witts, H., & Chater, T. (2018). What speech and language therapy do community dwelling stroke survivors with aphasia receive in the UK? *PLOS ONE*, *13*(7), e0200096. https://doi.org/10.1371/journal.pone.0200096

Patrício, B., Jesus, L., & Cruice, M. (2013). *Quality of life of the caregivers of people with aphasia. A systematic review. [Master's Thesis]*. Citeseer. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.926.7095&rep=rep1&type=pdf

Peurala, S. H., Kantanen, M. P., Sjögren, T., Paltamaa, J., Karhula, M., & Heinonen, A. (2012). Effectiveness of constraint-induced movement therapy on activity and participation after stroke: a systematic review and meta-analysis of randomized controlled trials. *Clinical Rehabilitation*, *26*(3), 209–223. https://doi.org/10.1177/0269215511420306

Pierce, J. E., Menahemi-Falkov, M., O'Halloran, R., Togher, L., & Rose, M. L. (2017). Constraint and multimodal approaches to therapy for chronic aphasia: A

systematic review and meta-analysis. *Neuropsychological Rehabilitation*, *29*(7), 1005–1041. https://doi.org/10.1080/09602011.2017.1365730

Pierce, J. E., O'Halloran, R., Menahemi-Falkov, M., Togher, L., & Rose, M. L. (2020). Comparing higher and lower weekly treatment intensity for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation*, *34*(2), 1–25. https://doi.org/10.1080/09602011.2020.1768127

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. L. (2019). What Is Meant by "Multimodal Therapy" for Aphasia? *American Journal of Speech-Language Pathology*, *28*(2), 706–716. https://doi.org/10.23641/asha.7646717

Pulvermüller, F., & Berthier, M. L. (2008). Aphasia therapy on a neuroscience basis. *Aphasiology*, *22*(6), 563–599. https://doi.org/10.1080/02687030701612213

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, *32*(7), 1621–1626. https://doi.org/10.1161/01.str.32.7.1621

Raymer, A. M., Beeson, P., Holland, A. L., Kendall, D. L., Maher, L. M., Martin, N., Murray, L., Rose, M. L., Thompson, C. K., & Turkstra, L. (2008). Translational research in aphasia: From neuroscience to neurorehabilitation. *Journal of Speech, Language, and Hearing Research*, *51*, S259–S275. https://doi.org/10.1044/1092-4388(2008/020)

RELEASE Collaboration. (2020). Communicating simply, but not too simply: Reporting of participants and speech and language interventions for aphasia after stroke. *International Journal of Speech-Language Pathology*, *22*(3), 302–312. https://doi.org/10.1080/17549507.2020.1762000

Rose, M. L. (2013). Releasing the constraints on aphasia therapy: the positive impact of gesture and multimodality treatments. *American Journal of Speech-Language Pathology*, *22*(2), S227-39. https://doi.org/10.1044/1058-0360(2012/12-0091)

Rose, M. L., Raymer, A. M., Lanyon, L. E., & Attard, M. C. (2013). A systematic review of gesture treatments for post-stroke aphasia. *Aphasiology*, *27*(9), 1090–1127. https://doi.org/10.1080/02687038.2013.805726

Sage, K., Snell, C., & Ralph, M. A. L. (2011). How intensive does anomia therapy for people with aphasia need to be? *Neuropsychological Rehabilitation*, *21*(1), 26–41. https://doi.org/10.1080/09602011.2010.528966

Sparks, R., Helm, N., & Albert, M. (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex*, *10*(4), 303–316.

Taub, E. (1976). Movement in nonhuman primates deprived of somatosensory feedback. *Exercise and Sport Sciences Reviews*, *4*, 335–374.

Taub, E., Uswatte, G., Mark, V. W., & Morris, D. M. (2006). The learned nonuse phenomenon: implications for rehabilitation. *Europa Medicophysica*, *42*(3), 241–255.

Thomson, J. (2012). *Assessing the benefits of multimodal rehabilitation therapy for aphasia [Masters Dissertation]* (pp. 1–143). https://www.escholar.manchester.ac.uk/uk-ac-man-scw:188802

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. https://doi.org/10.1080/17549500902726059

Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: a missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(1), 70–77. https://doi.org/10.1002/mrdd.20139

Chapter 2 – What is known about the effectiveness of constraint and multimodal treatments?

Publication

Pierce, J. E., Menahemi-Falkov, M., O'Halloran, R., Togher, L., & Rose, M. L. (2017). Constraint and multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation*, 1005-1041.

https://doi.org/10.1080/09602011.2017.1365730

Constraint and Multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis

John E. Pierce[a,b], Maya Menahemi-Falkov[a], Robyn O'Halloran[a], Leanne Togher[c] and Miranda L. Rose[a]*

[a]School of Allied Health, La Trobe University, Melbourne, Australia; [b]Speech Pathology, Cabrini Health, Melbourne, Australia; [c]Speech Pathology, Faculty of Health Sciences, The University of Sydney, Lidcombe, Australia

John E Pierce. 494 Glen Huntly Road, Elsternwick 3185, Australia. +61430 130 893. ORCID identifier: 0000-0001-5164-5106 Twitter: @johnpierce85 pierce.john.e@gmail.com

Maya Menahemi-Falkov. Health Sciences 1, La Trobe University, Plenty Road, Bundoora 3083, Australia. mayam10@gmail.com

Robyn O'Halloran. Health Sciences 1, La Trobe University, Plenty Road, Bundoora 3083, Australia. ORCID identifier: 0000-0002-2772-2164 R.OHalloran@latrobe.edu.au

Leanne Togher. Faculty of Health Sciences, The University of Sydney, 75 East St, Lidcombe 2141, Australia. ORCID identifier: 0000-0002-4518-6748 Leanne.togher@sydney.edu.au

*Miranda L. Rose. Health Sciences 1, La Trobe University, Plenty Road, Bundoora 3083, Australia. ORCID identifier: 0000-0002-8892-0965 M.rose@latrobe.edu.au

Constraint and Multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis

Aphasia is a significant cause of disability and reduced quality of life. Two speech pathology treatment approaches appear efficacious; Multimodal and Constraint Induced aphasia therapies. In Constraint Induced therapies, nonverbal actions (e.g., gesture, drawing) are believed to interfere with treatment and patients are therefore constrained to speech. In contrast, Multimodal therapies employ non-verbal modalities to *cue* word retrieval. Given the clinical and theoretical implications, a comparison of these two divergent treatments was sought. This systematic review investigated both approaches in chronic aphasia at the levels of impairment, participation and quality of life. After a systematic search, the level of evidence and methodological quality were rated. Meta-analysis was conducted on 14 single case experimental designs using Tau-U, while heterogeneity in the four group designs precluded meta-analysis. Results showed that high-quality research was limited; however, findings were broadly positive for both approaches with neither being judged as clearly superior. Most studies examined impairment-based outcomes without considering participation or quality of life. The application and definition of constraint varied significantly between studies. Both constraint and multimodal therapies are promising for chronic post-stroke aphasia, but there is a need for larger, more rigorously conducted studies. The interpretation of "constraint" also requires clearer reporting.

Keywords: systematic review; aphasia; constraint; multimodal; therapy

Word count: 8050 word (inc. citations)

Introduction

The presence of aphasia after stroke results in significantly poorer quality of life than stroke alone (Hilari, 2011). Aphasia is perceived as more detrimental for quality of life than any other illness including cancer, Alzheimer's Disease and quadriplegia (Lam & Wodchis, 2010). There are also significant financial and carer burdens (Flowers, Silver, Fang, Rochon, & Martino, 2013; Patrício, Jesus, & Cruice, 2013) and thus, effective treatments for aphasia are highly sought after.

Constraint and multimodal are two treatment approaches in aphasia. Multimodal treatments have a long history of use in aphasia research and treatment, while constraint principles were first introduced in 2001. However, these two approaches have theoretically distinct rationales. Authors of constraint therapies such as Constraint Induced Aphasia Therapy (CIAT) posit that the use of other communication modalities distract from, and therefore weaken, recovery of verbal output, while authors of multimodal therapies suggest that these additional modalities can facilitate word retrieval and learning. This results in a key difference between treatments regarding the cueing of patient responses. These two therapy approaches and their rationales will be described before their evidence is compared.

Multimodal

The concept of multimodal cueing in aphasia rehabilitation dates back at least to the 1940s, when Luria put forward the principle of *Intersystemic Reorganisation* (Luria, 1970). This principle proposes that a defective system can be supported and supplemented by another, less damaged, system. Luria gave the example of recruiting the visual system to compensate for impaired proprioception or balance during walking. In language, the

impaired system (spoken word retrieval) might be assisted by gesture production, for example. It is important to note that in multimodality therapies, other modalities are used to *facilitate* spoken output and not to replace it, though compensation may be a secondary goal if improvement of spoken output is not successful. The most common communicative modalities used to promote word retrieval and their proposed mechanisms are described below.

Reading.

Orthographic cues, typically manifested as first letter cues, are widely used in aphasia speech pathology practice (Lorenz & Nickels, 2007). Research has found orthographic cues to be generally successful for spoken naming (Nickels, 2002). The orthographic cue is thought to provide complete or partial activation of the phonological form of the target word using grapheme-phoneme conversion (Lorenz & Nickels, 2007); in other words, the letter/s provide an internal phonemic cue to facilitate word production.

Writing.

As with reading, writing takes advantage of the deep links between graphemes and phonemes. When a patient writes all or part of a target word, this might provide an alternative route to the phonological output lexicon (a store of all phonological forms of known words) via the reading route(s) described above (Nickels, 1992). Alternatively, the connection between the orthographic and phonological output lexicons may be bidirectional, meaning that phonological information is active whenever written word forms are accessed, as well as vice versa (Kiran, 2005). This connection is proposed based on the fact that written naming alone can improve spoken naming (DeDe, Parris, & Waters, 2003; Wright, Marshall, Wilson, & Page, 2008).

Gesture.

The *lexical retrieval hypothesis* holds that, even in non-aphasic speakers, gestures used in conversation are more for the speaker's benefit than the listener's in that they aid in word retrieval (Beattie & Shovelton, 2006). Evidence for this hypothesis comes from findings that a) more gestures are produced during word finding difficulties in the speech of both normal and aphasic individuals, and b) restricting gestures during speech increases the frequency of dysfluencies (Rose, 2006). There is "intense theorising" about exactly how gestures assist with word finding with language (Hadar & Rumiati, 2006, p. 141), but two possible mechanisms are described here.

Language is traditionally viewed as a discrete system in the brain. *Embodied language*, a subset of *embodied cognition*, is the theory that language is connected to action and sensory systems and there is experimental support for this theory (Fischer & Zwaan, 2008). For example, one study showed that reading of "action" words was immediately followed by activation of a relevant motor cortex area, such as the word "kick" activating the leg motor areas (Pulvermüller & Berthier, 2008). It follows from such close motor-language connections that where a word cannot be produced, its corresponding gesture might aid retrieval. The other explanation is that gesture aids the speaker in preverbal message planning. That is, gesture helps the speaker mentally arrange the spatial and visual thoughts behind a message, and this stimulates processing of semantic features which assist word retrieval (Feyereisen, 2006).

Drawing.

In aphasia, drawing has typically been used to augment or compensate for speech loss (Sacchett, 2002). However, there are case reports of individuals using drawing to self-

cue verbal word retrieval, or individuals whose naming has improved after treatment targeting drawing accuracy (Cubelli, 1995). Farias, Davis and Harrington (2006) explored drawing as a facilitator of word finding in 22 people with aphasia. While participants were drawing the target, confrontation naming improved compared to both baseline and attempts at written naming. Interestingly, this effect was not influenced by drawing quality (as measured by recognisability ratings), suggesting that it could be a suitable treatment even for those with limb apraxia or significant hemiplegia.

How would drawings assist word retrieval? Drawings, like gesture, have the advantage of being free from linguistic symbolism. Drawing quality in aphasia correlates with the integrity of the semantic system (Farias et al., 2006), which suggests that the process of drawing requires access to semantic features. For naming, drawing is thought to stimulate the semantic aspects of objects by placing attention on visual features (Farias et al., 2006) and may do so for longer than naming attempts alone (Makuuchi, Kaminaga, & Sugishita, 2003). For example, in drawing a truck, a person needs to focus more deeply on features such as its large size, square shape and additional wheels, and perhaps its function, than when merely naming it from sight. This process might suppress competing concepts whose features do not match the target while activating sufficient semantic features to assist word retrieval.

Music.

Melodic Intonation Therapy (MIT; Sparks, Helm, & Albert, 1974) is probably the most widely used and recognised music-related therapy for aphasia. MIT utilises "intoned speech", a song-like prosody, for phrases and utterances. Notes can be either high or low depending on syllable stress, and syllables are rhythmicised. Simultaneous left hand tapping

is also employed during word or phrase production (Zumbansen, Peretz, & Hébert, 2014a). There are other variations of music and rhythm treatments, including Modified MIT, Singen Intonation Prosodie Atmung Rhytmusübungen Improvationen (SIPARI) and Speech-Music Therapy in Aphasia (Hurkmans et al., 2012). As with other multimodal treatments, all these treatments aim to encourage verbal output. Generalisation to untreated phrases in conversational speech is the ultimate goal and patients are not expected to sing in everyday life (Zumbansen, Peretz, & Hébert, 2014a).

There are several suggested mechanisms for melodic therapies. Sparks et al. (1974) were initially uncertain as to how to explain the positive results of MIT, but did not believe that the right hemisphere was learning to take over language production. Instead, they proposed that the right hemisphere was assisting the left hemisphere. Tapping in MIT is left handed for this reason — to encourage activation of the right hemisphere. Neuroimaging evidence is mixed for the theory of increased right hemisphere activation in MIT, with some showing increased perilesional left hemisphere activation and others, increased right hemisphere activation (Zumbansen, Peretz, & Hébert, 2014a). Even if present, right hemisphere activity might only increase due to the high intensity of "singing" that occurs during such treatments, with the improvements in speech due to the repetition of phrases. That is, right hemisphere activity and verbal improvement could be independent events (Stahl, Kotz, Henseler, Turner, & Geyer, 2011).

A more recent hypothesis is that the rhythmic component of these treatments, particularly the tapping of the left hand, is the true underlying mechanism (Stahl et al 2011). This has some experimental support (Zumbansen, Peretz, & Hébert, 2014a), while other studies found superiority of combined rhythm and pitch (Zumbansen, Peretz, & Hébert, 2014b).

Combined Multimodal

Multiple non-verbal modalities may be combined within a treatment to maximise cueing of verbal output. M-MAT (Multi-modality Aphasia Therapy) is a high intensity combined multimodal treatment (Attard, Rose, & Lanyon, 2013) which utilises a structured cueing hierarchy of gesture, drawing and writing to cue word retrieval. A structured and detailed protocol delineates that each time a patient is unable to produce a target, they are asked to gesture, draw and write or copy the word while repeating it verbally. M-MAT has shown improvements across receptive and expressive language measures, at both impairment and activity/participation levels (Rose, Attard, Mok, Lanyon, & Foster, 2013).

Constraint

At the heart of constraint therapy lies the concept of *learned nonuse*. Taub and colleagues developed the term based on observation of monkeys with deafferented upper limbs (Taub, 1976). They proposed that the deficit from an injury is not wholly due to the physiological impairment, but also a subconscious preference not to use the affected body part (Taub, Uswatte, Mark, & Morris, 2006). They hypothesised that nonuse is learned through (a) punishment when trying to use the affected limb (e.g., issues with incoordination or dropping), and (b) positive reinforcement when using the alternative limb. A "vicious spiral" (Taub et al., 2006, p. 245) of nonuse then commences as reduced use leads to shrinkage of the relevant cortical area, which results in less use of the limb, and so on. However, by constraining function to the damaged limb, through either restraining the unaffected limb or presenting tasks which necessitate the use of both limbs, improvements in the function of the deafferented upper limbs were seen.

Constraint Induced Movement Therapy (CIMT) was subsequently developed to address learned nonuse in adult stroke patients. The key elements of CIMT are intensive training, use of functional transfer tasks, and constraint applied to the affected limb (Taub et al., 2006). The efficacy of CIMT has been demonstrated extensively in upper and lower limbs in stroke as well as other conditions (Smania, 2006; Taub et al., 2006).

CIAT was developed through combining the principles of CIMT with an existing aphasia treatment, Communicative Aphasia Therapy (Pulvermüller & Berthier, 2008), which employs language games that rely on the correct verbal response of patients. First reported in 2001 (Pulvermüller et al., 2001), CIAT has three primary principles: 1. Massed practice over a short period, 2. Action-embedded, relevant language, and 3. Constraint to possible, but avoided, verbal output. It is the third principle that addresses learned nonuse. In the context of language, learned nonuse is proposed to occur where patients avoid problematic words or phrases, reduce their attempts at verbal communication overall, or use alternative modalities to compensate, such as gesture or writing. In CIAT, constraint is applied through progressive difficulty of stimuli, gradual shaping of responses into more complex utterances, and prohibiting nonverbal modes of communication.

More recently, the potentially detrimental effects of inhibiting multimodal self-cueing have been recognised in CIAT and its use has now been permitted (Difrancesco, Pulvermüller, & Mohr, 2012). Even so, any multimodal cues used within CIAT remain incidental and patient-generated, in contrast to the systematic, clinician-prompted cueing of multimodal treatments. A number of variations on the original CIAT protocol have been described, including CIAT II (Johnson et al., 2014) and CIAT Plus (Attard et al., 2013; Meinzer, Djundja, Barthel, Elbert, & Rockstroh, 2005). CIAT Plus builds upon CIAT by assigning home tasks in order to improve carryover of language skills into real life and using written stimuli

as well as photographs (Meinzer et al., 2005). CIAT II uses a wider variety of language activities than CIAT, including a role-playing task, picture description and repetition drills (Johnson et al., 2014).

Summary

As outlined, constraint aphasia therapies seek to avoid learned nonuse of language partly through constraining participants to the verbal modality, whereas multimodal treatments seek to leverage intact modalities to aid verbal output.

The most recent Cochrane review of aphasia therapy called for further data comparing different therapies in order to identify the most effective treatments (Brady, Kelly, Godwin, Enderby, & Campbell, 2016). A comparison of constraint and multimodal approaches will improve outcomes for people with chronic aphasia if one is found to be more effective (Rose et al., 2013). If patients clearly respond better to being constrained to the verbal modality, then use of multimodal cues should be re-examined. Conversely, if patients improve more with multimodal cues, the contribution of verbal constraint in the CIAT protocol may be questioned.

CIAT and M-MAT are two treatments that are operationally similar despite being constraint and multimodal treatments, respectively. Both use group language games which provide a social imperative for successful communication, have a high intensity of treatment (30 hours over two weeks), and use shaping of responses to gradually increase the complexity of utterances. To date, CIAT and M-MAT have been directly compared in two studies. In a pilot study (Attard et al., 2013), a single case crossover design was used to provide CIAT Plus and M-MAT to two participants. Confrontation naming of treated items

was marginally superior for M-MAT compared to CIAT Plus. In a phase 1 trial (Rose et al.,

2013), 11 participants underwent both treatments in a multiple baseline crossover design.

Both produced strong positive effects for the primary outcome measure (confrontation

naming), with comparable mean effect sizes (M-MAT = 8.00, CIAT Plus = 8.58) according to

Busk and Serlin's *d* (Busk & Serlin, 1992). These similar results in direct comparisons are

puzzling given the contrasting nature of cues provided.

In a narrative review of constraint and multimodal treatments, Rose (2013)

examined the theoretical explanations for constraint and multimodal treatments in detail as

well as critically examining the literature to date. The review concluded that there was

limited theoretical support for constraining people with aphasia to the verbal modality and

that doing so was potentially counterproductive for word retrieval. The research examined

within the review did not favour either approach. However, as a narrative review, the

literature reviewed was not exhaustive and was not subjected to meta-analysis. Therefore,

the purpose of this review was to systematically examine and compare the efficacy of

constraint and multimodal therapies more broadly than CIAT Plus and M-MAT. There were

three key questions. Two examined outcomes based on the World Health Organisation's

International Classification of Functioning, Disability and Health (ICF). Quality of life was also

considered, frequently described as a missing component of the ICF and under

consideration for inclusion in future versions of the ICF (Ravenek, Skarakis-Doyle, Spaulding,

Jenkins, & Doyle, 2015). Finally, the outcome of carer burden was included due to the

significant negative impacts of aphasia on carers (Patrício et al., 2013).

For stroke-induced chronic aphasia, what is the influence of constraint and multimodal treatments on measures of *(1) language impairment, (2) communication activity/participation*, and *(3) quality of life and carer burden*?

Definitions

Multimodal therapy or *multimodal training* has many variations in speech pathology and even within the field of aphasiology, so it is important to describe our definition for this systematic review. In this paper we use *multimodal therapy* to mean treatments which expect the patient to produce output in different modalities alongside speech production in order to cue speech. Our own operationalised definition of *multimodal therapy* also formed the basis for our inclusion criteria:

1. The patient carries out a communication task (particularly writing, gesture, singing, or drawing) at the same time as or immediately before an attempt at speech.

2. Speech targets are practised repeatedly in a therapy task with the intention of improving spontaneous speech in the long term rather than as a short-term facilitation effect.

Based on this definition, studies looking at a simultaneous action that is not communicative are not eligible. There is a series of research looking at the effects of "intention gesture", for example, where participants make a nonsymbolic, circular gesture with their left hand while repeating the target word. These are not eligible as the nonsymbolic gesture is designed to activate the right hemisphere and is not related to the target word in a linguistic sense. Outside of this restriction we included any gesture, as

classified in McNeill's (1992) model which differentiates gesticulation, pantomime, emblems and sign languages.

MIT and other melodic treatments were included as per the above definition as they combine the modalities of singing and speech in a song-like prosody. Studies investigating the use of rhythm on language without a melodic component do not meet our definition as rhythm is not a communicative act in isolation.

Studies including orthographic cueing (i.e., reading the target word or part of it) as their only multimodal cue were not included as they are so commonly used as to include the majority of the aphasia literature. We also limited the writing studies to those focussing on writing or typing the whole target word and not those that taught phoneme-grapheme correspondences as an explicit strategy.

Method

Search terms relating to aphasia, constraint and multimodal therapies (including drawing, writing, gesture and music) were used. An example of the search strategy employed is displayed in Table 2-1. Three major databases were searched in September 2015: Medline (OVID, 1946-present), CINAHL (Ebscohost) and Psycinfo (OVID, 1987-present). No limits for date of publication were applied as we wanted to find as many relevant articles as possible.

Duplicate results were excluded using citation software and manual checking. Resulting titles and abstracts were then screened by each of the first two authors as per the following inclusion criteria:

- Original data

- Experimental design

- Article from peer reviewed journal, text in English

- Adults (18 years +) with stroke-induced, chronic aphasia (≥6 months)

- Uses constraint to verbal output or uses multimodal therapy (see earlier definition)

- Purposefully investigates effects on verbal output

If a study included participants who did not meet the demographic criteria, it was excluded unless results for eligible participants could be separated.

Table 2-1 – Search strategy example: Medline (OVID, 1946-present)

Keywords:	aphasia, dysphasia, anomia, NOT "primary progressive"	AND	Therapy, Intervention, Treatment	AND	CIAT*, CILT*, "Constraint induced language", "constraint induced aphasia", "Constraint language", "Constraint aphasia", "Forced language treatment", ILAT, "intensive language action"	OR	M-MAT, MMAT, "Multi-modality aphasia therapy", multimodal*, cross modal*, Draw*, Amerind, Amer-ind, Pantomim*, Sign, signs, signing, Makaton, Gestur*, gestic*, iconic, Writ*, graphem*, orthograph*, Read*, music*, melod*, sing, singing, rhythm
	OR						
Subject headings:	*Aphasia, Anomia,* NOT *Aphasia, primary progressive*						

Note. **All subject headings were exploded, and all subheadings were included.**

After the screening process, the first two authors met to achieve consensus on discrepancies regarding inclusion of studies. Full texts of the included articles were obtained and checked again according to the inclusion criteria. Any discrepancies were discussed until consensus was reached. The screening process is visualised in Figure 2-1. This selection process resulted in 60 papers.

Figure 2-1 – Flow diagram of search results
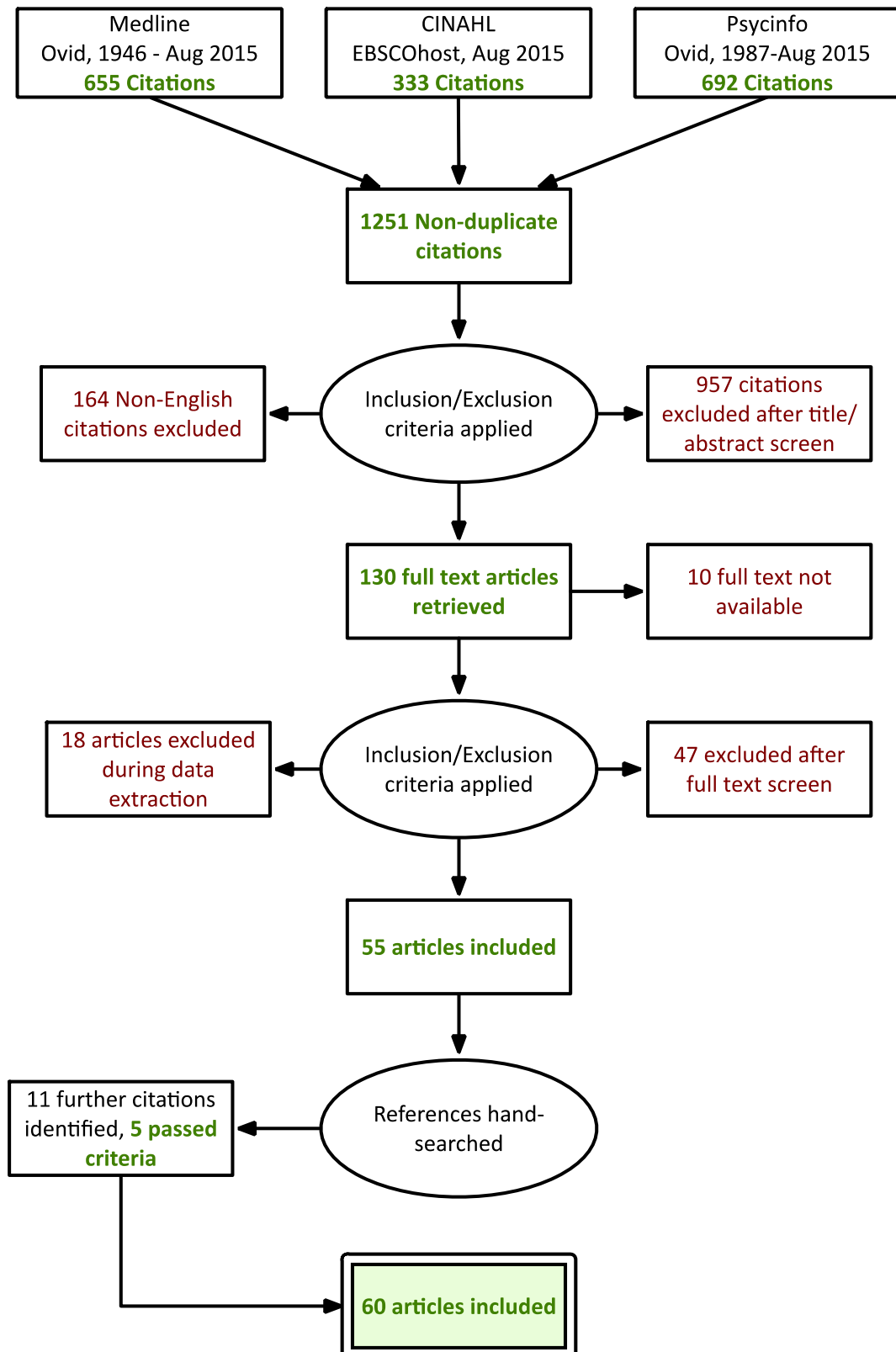
The final articles were categorised according to study type (Oxford Centre for Evidence Based Medicine - Levels of Evidence; 2011), treatment type (constraint or modality type) and outcome (Impairment, Activity/Participation, Quality of life, Carer burden). Author ROH resolved discrepancies during this process. The OCEBM levels do not have a manual that clearly describes categories and the Level "Case series" presumably includes pre/post group studies, case reports and Single Case Experimental Designs (SCEDs) that are not the traditional, medical N-of-1 design. However, there is a vast difference in rigour between SCEDs, which can provide evidence of cause and effect of a treatment, and other single participant studies such as case reports (Tate et al., 2013) or pre/post group designs. Therefore, we have included an additional level of evidence for SCEDs, below randomised trials and above non-randomised trials and have used this method in our categorisation. This more accurately represents the levels of rigour of study designs. Using the guidelines in Tate et al. (2013), this category included any multiple baseline, alternating treatments, withdrawal/reversal or changing criterion designs. It also included quasi-experimental bi-phasic (AB) designs.

Outcomes were classified as Activity/Participation where they rated or measured performance in either real or simulated everyday life activities. Quality of life outcomes were those using subjective ratings of life satisfaction, while carer burden outcomes were any that proposed to measure the carer's wellbeing or distress of any sort. An extraction template was created and included data fields for sample size, treatment, details of cueing and constraint used, and outcome measures. The completed data extraction for all 60 articles is included in Appendix 1.

Randomised controlled trials (RCTs) and nonrandomised controlled trials (nonRCTs) were assessed for methodological quality using the PEDro-P, an adapted version of the

PEDro scale (Sherrington, Herbert, Maher, & Moseley, 2000) used by the PsycBITE and speechBITE teams (see Fitzpatrick, 2008). Ratings by the first author were compared against those found on SpeechBITE and PsycBITE and any discrepancies were resolved with an independent rater from the SpeechBITE team.

The RoBiNT scale (Tate et al., 2015) was used to assess methodological quality for SCEDs. Non-experimental single case designs, including one phase designs, pre/post designs and case descriptions, were not rated. Pre/post group designs were also not rated as these have no experimental control and form a low level of evidence. Each study was rated by two authors and discrepancies were discussed until agreement was reached.

For RCTs and nonRCTs, scores five and above on PEDro-P were considered moderate to high quality based on commonly accepted consensus (Centre for Evidence-Based Physiotherapy, 2016). For the RoBiNT, benchmarks have not yet been formally established, but in a paper examining the reliability of the scale on a small number of papers (Tate et al., 2013), the mean score was 12 with the highest score being 18 (Tate et al., 2015). In lieu of existing benchmarks, SCEDs that scored 12 or higher were therefore considered moderate to high quality. Studies that did not reach these quality cutoff scores (≥5 and ≥12) are included in Appendix 1 but they were not considered further in the results.

Effect size calculation

Group Designs

We planned to calculate effect sizes for relevant group studies of Level 2 or 3 evidence (RCTs and nonRCTs); however, studies were too heterogeneous in terms of outcome measures and treatments (see results).

Single Case Experimental Designs

Calculation of effect sizes for SCEDs is a domain that continues to develop but it has a number of promising methods (Parker, Vannest, Davis, & Sauber, 2011). Tau-U is a recent effect size measure which is resistant to effects of autocorrelation, considers baseline trend, deals well with only few data points and has the highest power of the non-overlap indices (Brossart, Vannest, Davis, & Patience, 2014). Tau-U can also produce confidence intervals.

In this review, to be eligible for SCED effect size calculation, studies needed to be an experimental design (viz., we excluded one phase designs, pre/post designs and case descriptions), have raw data presented on case charts and have at least 2 data points in each of the baseline and intervention phases. Only the outcome measure *confrontation naming of treated items* was used for effect sizes, which included 94% of SCEDs with a RoBiNT score of 12 or greater. Data was extracted by the first author and calculated using the online Tau-U calculator at www.singlecaseresearch.org. A weighted average Tau-U was calculated across all relevant participants and word sets for each paper. Positive baseline trend was corrected where baselines had a Tau score of greater than .4 (Parker et al., 2011) and an increasing trend apparent on visual inspection. Negative baseline was not corrected, nor was trend in treatment phases, as this would have boosted effect sizes. These steps are consistent with recommendations to correct for trend conservatively and to check statistical results against visual analysis (Brossart et al., 2014; Parker et al., 2011). Maintenance and follow up phases were not included in calculations.

Results

As outlined in Figure 2-1, of 1680 original results (1251 non-duplicates), 60 papers met the criteria for this systematic review. Figure 2-2 summarises the number of studies for

each treatment approach.
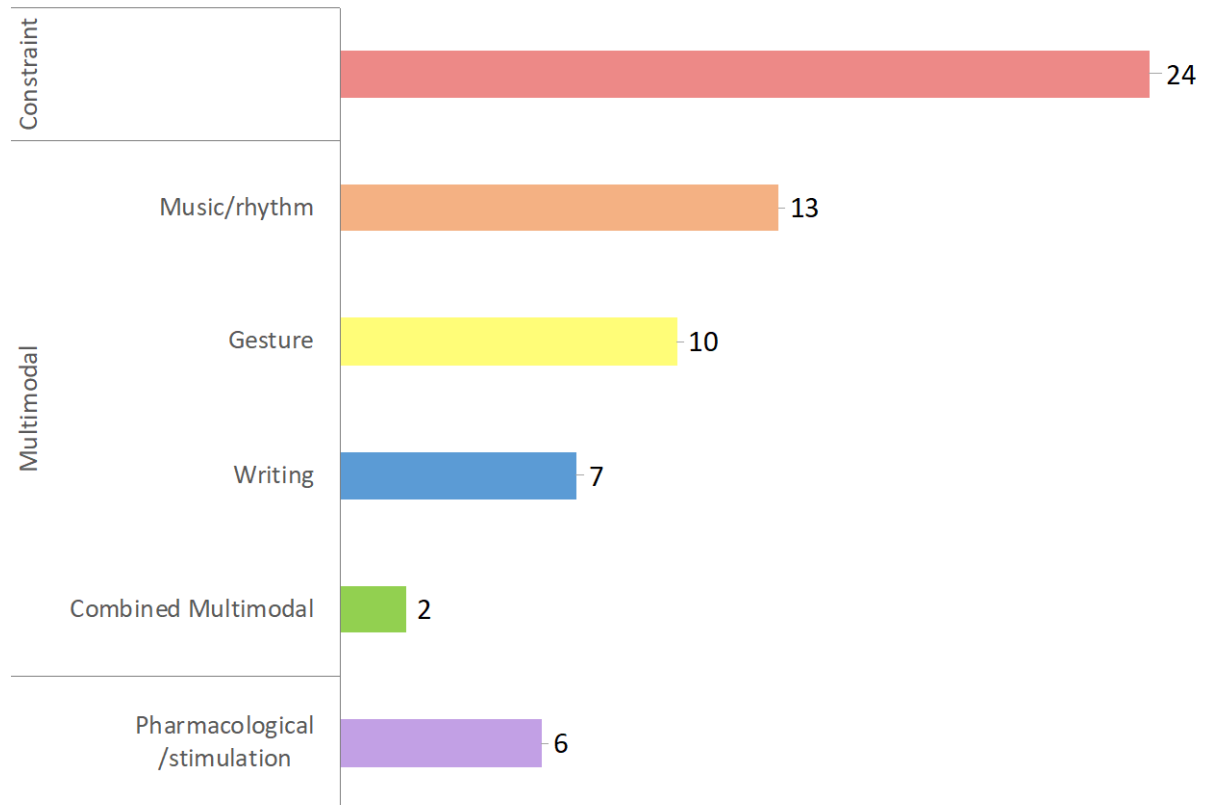
Figure 2-2 – Number of results by treatment approach

Descriptive characteristics of data

Overall, there were more multimodal treatment studies than constraint (36 vs. 24) but more constraint studies than studies on any single modality. The first constraint paper was published in 2001 (Pulvermüller et al.), while the earliest multimodal paper was published in 1974 (Sparks et al.), demonstrating the rapid growth in constraint therapy research.

No studies were identified that used drawing alone as a means of cueing verbal output. Six studies utilised neurological stimulation (e.g. rTMS) or pharmacological treatments (e.g., memantine) in combination with constraint or multimodal therapies. Two of these were RCTs (Barbancho et al., 2015; Berthier et al., 2009), one a SCED (Al-Janabi et

al., 2014) and three were pre/post designs (Abo et al., 2012; Martin et al., 2014; Vines, Norton, & Schlaug, 2011). While results of these mixed treatments were broadly positive, the contribution of the constraint and multimodal therapies could not be differentiated from the pharmacological and stimulation treatment aspects. These studies were therefore not considered further in this review.

The majority of the twenty-four constraint studies reported use of the original CIAT approach. Other variations included CIAT Plus (Attard et al., 2013; Meinzer et al., 2005), CIAT II (Johnson et al., 2014), lower intensity CIAT (Goral & Kempler, 2009; Kempler & Goral, 2011; Maul, Conner, Kempler, Radvanski, & Goral, 2014), or modifications such as inclusion of grammatical shaping (Faroqi-Shah & Virion, 2009) or drill tasks (Kempler & Goral, 2011).

The way "constraint" was applied varied considerably and was not well described in many studies. While all studies constrained communication between participants to the verbal modality only, three studies explicitly prevented participants from using gestures to self-cue (Breier, Maher, Novak, & Papanicolaou, 2006; L. M. Maher et al., 2006; Martin et al., 2014) whereas five studies allowed such gesture (Attard et al., 2013; MacGregor, Difrancesco, Pulvermüller, Shtyrov, & Mohr, 2015; Meinzer, Streiftau, & Rockstroh, 2007; Mohr, Difrancesco, Harrington, Evans, & Pulvermüller, 2014). The remaining 19 did not specify whether they allowed self-cueing or not, including the original CIAT paper (Pulvermüller et al., 2001). Similarly, while the majority of studies (14) allowed therapists to provide cueing when necessary, one reported not providing cues (Maul et al., 2014) and the remaining 11 provided no description of whether therapists provided cues.
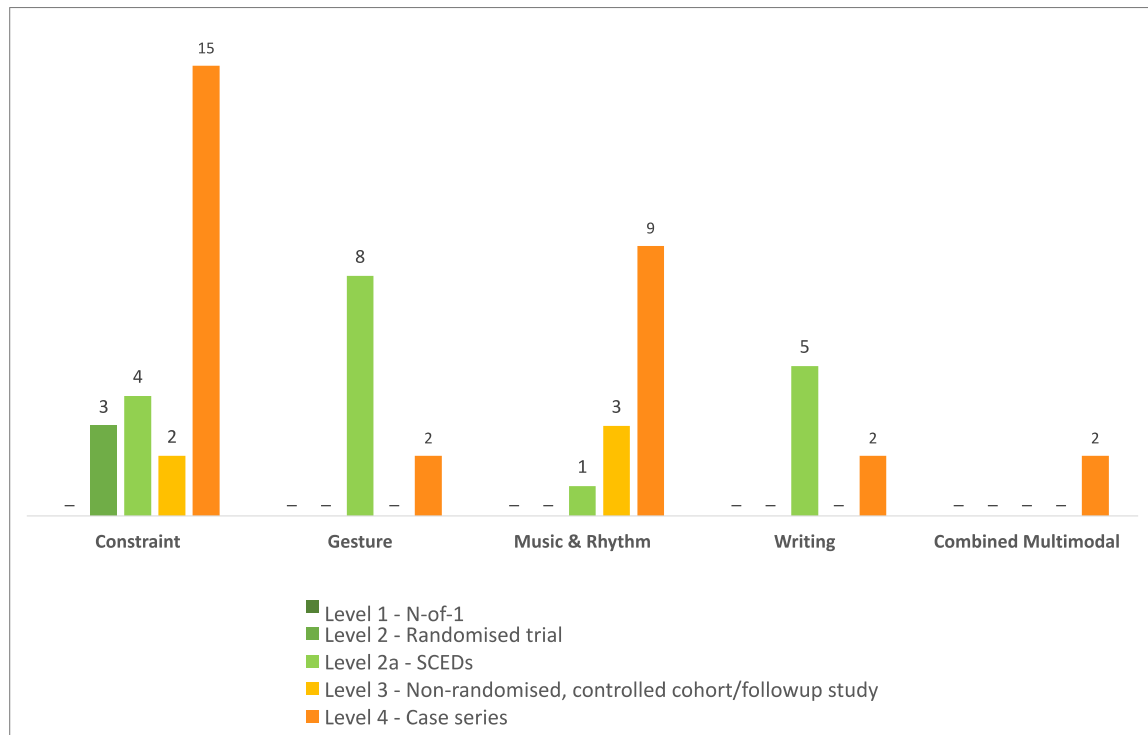
Figure 2-3 – Number of papers per level (OCEBM levels of evidence)

Note. Level 4 studies (Case Series) were not eligible for quality evaluation and meta-analysis. In addition to N-of-1 designs, Level 1 ordinarily includes Systematic Reviews but these were not eligible as they are not original data.

Figure 2-3 shows the levels of evidence found for each therapy type according to OCEBM levels. There were a very limited number of RCTs and nonRCTS – five for constraint and three for music – however, there were a number of SCEDs (19), particularly in gesture studies. The remaining papers were low quality designs.

Figure 2-4 presents the methodological scores for papers for controlled trials (PEDro-P, 2-4b) and SCEDs (RoBiNT, 2-4b) and the red line shows the cutoff scores for inclusion in further analysis. As visualised in Figure 2-4a, the methodological quality scores ranged from 3-7 on PEDro-P (/10), with half the eight controlled trials (Level 2-3) below the cutoff score of 5. Quality scores on the RoBiNT ranged from 9-21 of a possible 30, with three of the 18 SCED studies (Level 2a) below the cutoff score of 12 (Figure 2-4b).

Randomised/Non-randomised Controlled Trials - Pedro-P scores

Figure 4a.

Single Case Experimental Designs - RoBiNT scores
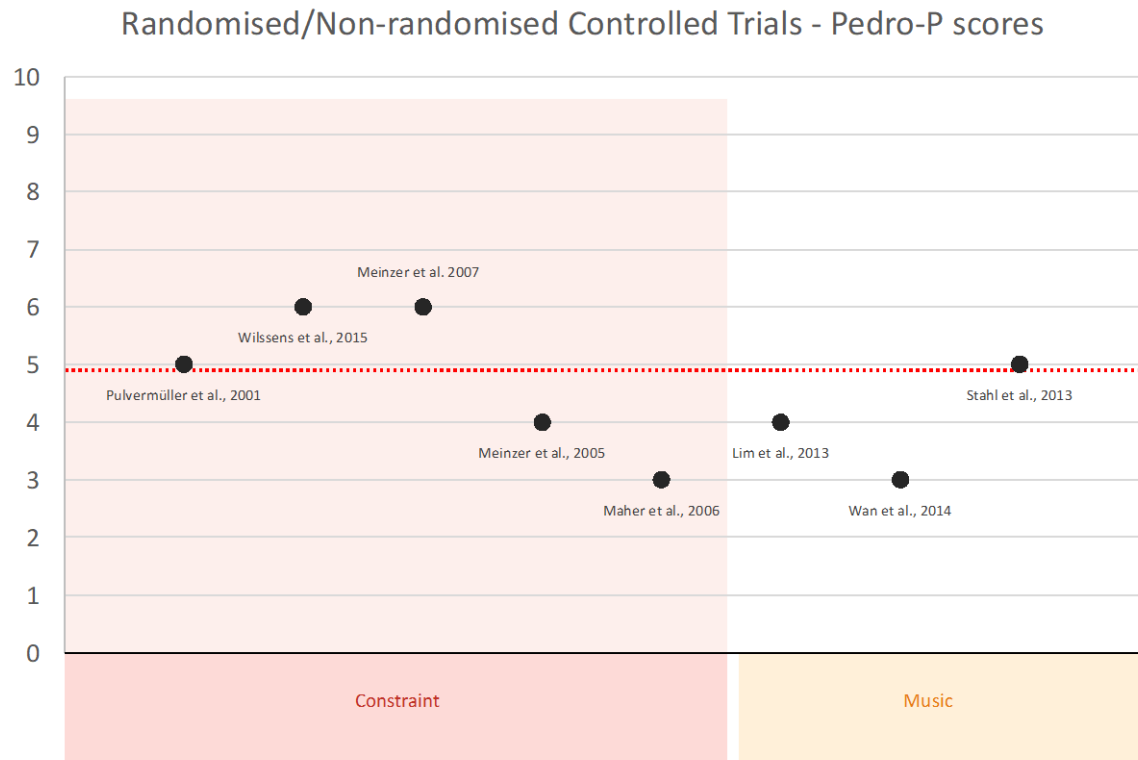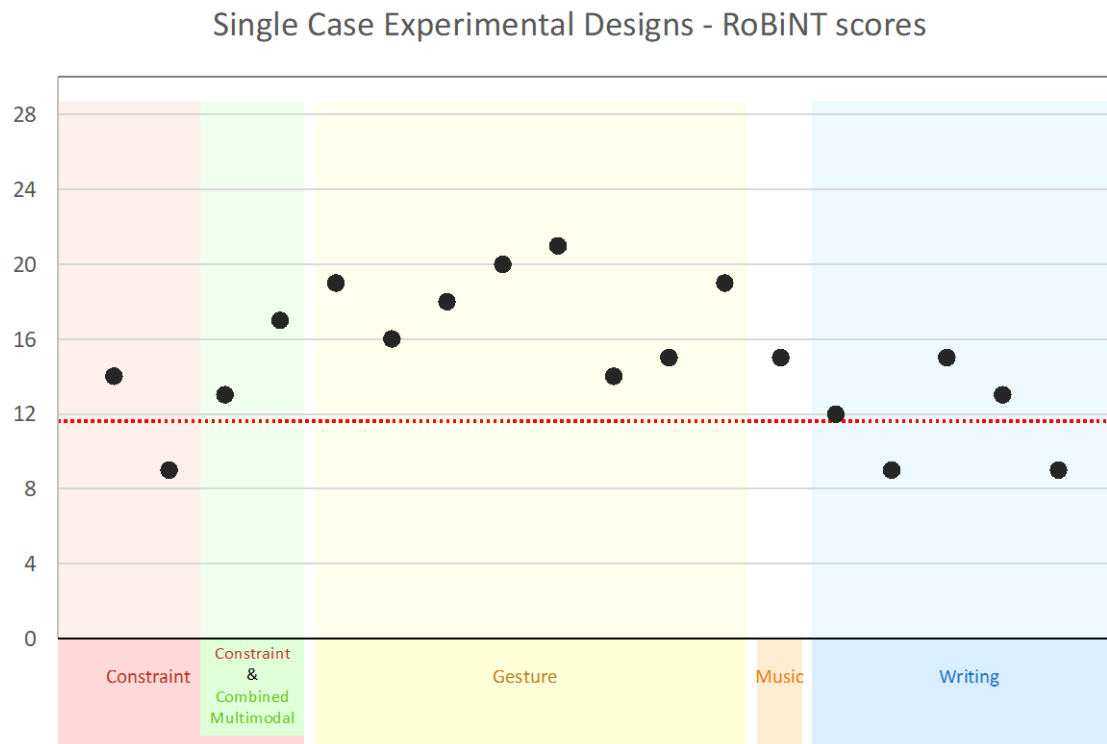
Figure 4b.

Figure 2-4 – Pedro-P and RoBiNT scores

Note. Red line indicates cutoff score for methodological quality.

The Tau-U effect sizes for confrontation naming of treated items across 14 eligible studies are shown in Figure 2-5. Many lacked regular intervals between probes and/or had significant gaps in probing, and these are noted. Interpretation of Tau-U effect sizes is difficult without benchmarks, which are not available for Tau-U at present. However, Tau-U scores are an indication of the percentage of data in the treatment phase that has improved over time compared to baseline (Parker et al., 2011). The limits are -1.0 and 1.0, which would indicate 100% reduction and improvement in scores respectively, while 0 would indicate no change in scores. A total effect size weighted by study was calculated for each modality. Studies with more data points, whether due to more participants or more probes during phases, have a greater weighting (Parker et al., 2011).

Results are described below in relation to the key questions of this review: by outcome level (Impairment, Activity/Participation, Quality of life, Carer burden) and within outcome level by treatment type (constraint, multimodal). Only studies Level 3 or higher which met minimum quality cutoff scores are discussed.

	Tau-U	95% C.I.	Weight	Sets calculated	Notes
Constraint					
Rose et al. 2013	**0.240**	0.020 - 0.460	3.3	CIAT-plus	Irregular probing frequency
Kurland et al. 2012	**0.692**	0.266 - 1.000	1.2	Constraint items only	Seven day intermission between baseline and treatment phase with reduced probe frequency
Attard et al. 2013	**1.000**	0.200 - 1.000	0.2	CIAT-plus	Irregular probing frequency
Combined Multimodal					
Rose et al. 2013	**0.684**	0.446 - 0.921	2.5	M-MAT	Irregular probing frequency
Attard et al. 2013	**0.813**	0.149 - 1.000	0.4	M-MAT	Irregular probing frequency
Gesture					
Raymer et al. 2006	**0.670**	0.528 - 0.812	21.0	Trained nouns & verbs	One month break beween baseline and treatment phases for some participants
Rose & Sussmilch 2008	**0.731**	0.543 - 0.918	14.1	Gest, Sem + Gesture	Delayed treatment of control set included
Raymer et al. 2012	**0.252**	0.045 - 0.460	13.6	Gest. facilitation	Duration of intermission period unclear
Boo & Rose 2011	**0.857**	0.616 - 1.000	12.8	Gest, sem + gesture	
Rose & Douglas 2002	**0.738**	0.482 - 0.995	7.6	Gest, Gest + Verbal	Delayed treatment of control set included
Ferguson et al. 2012	**-0.043**	-0.341 - 0.256	4.3	Pantomime set only	Seven day intermission between baseline and treatment phase for some participants
Rose & Douglas 2008	**0.913**	0.633 - 1.000	5.3	Gest, Gest + Verbal	
Rodriguez et al. 2006	**0.464**	0.175 - 0.753	4.2	Verbal naming scores, Verbal + Gesture	One month break between baseline and treatment phase for Participant 2
Writing					
Hillis 1989	**0.997**	0.669 - 1.000	4.7	P2, verb & noun verbal naming (treatment 2)	
Wright et al. 2008	**0.896**	0.612 - 1.000	3.1	All treated sets	Significant, irregular gaps in baseline probing frequency
Beeson & Egnor 2006	**0.864**	0.539 - 1.000	1.8	Verbal naming (CART + Repetition)	

Graph Generated by DistillerSR

Figure 2-5 – Tau-U meta-analysis

Note. Solid lines -1.0 and 1.0 indicate the possible limits of Tau-U scores – negative values demonstrate reduced naming performance and positive clues demonstrate improvement. Dotted lines with diamonds indicate the weighted average for each area or modality. Larger squares indicate larger relative weighting.

CIAT = Constraint Induced Aphasia Therapy, M-MMAT = Multi-Modal Aphasia Therapy, Gest = Gesture, Sem = Semantic

Impairment Outcomes

Constraint Treatment

Impairment-based results are detailed in Table 2-3. Constraint had the highest number of top tier studies according to the OCEBM levels; however, only three met quality criteria on the PEDro-P and all had small sample sizes (range 9 - 27). In addition, two of the three did not compare constraint to an equivalent non-constraint control. Meinzer, Streiftau and Rockstroh (2007) compared CIAT to CIAT run by family, and Pulvermüller et al. (2001) compared CIAT to conventional therapy of equal total hours but lower intensity. Wilssens et al. (2015) did compare CIAT to equal intensity conventional therapy and included impairment based assessments as secondary outcome measures.

Of the three constraint SCEDs that met quality criteria, two compared CIAT Plus to M-MAT (Attard et al., 2013; Rose et al., 2013) while the other (Kurland, Pulvermüller, Silva, Burke, & Andrianopoulos, 2012) compared CIAT to PACE (Promoting Aphasic Communicative Effectiveness, Davis, 2005). Figure 2-5 shows Tau-U effect sizes with constraint studies at the top followed by multimodal studies. In Attard et al. (2013), the Tau-U effect size for CIAT Plus (1.0) was greater than M-MAT (0.81), though M-MAT was within the 95% confidence interval of CIAT Plus. In Rose et al. (2013) the CIAT Plus effect size was markedly lower than M-MAT (.24 and.68) and the upper and lower confidence interval bounds just overlapped.

The weighted Tau-U effect size for all three high-quality constraint SCEDs was .374, which means that approximately 37% of treatment phase naming scores were higher than baseline scores. This was the lowest effect size relative to other modalities and was reduced by Rose et al. (2013) which, due to its larger number of participants and probes, held a

greater weighting. However, Rose et al. (2013) was an outlier and the other three constraint SCEDs were more in line with the effect sizes of other modalities (see Figure 2-5), though they had lower weighting and RoBiNT scores.

As displayed in Table 2-3, group studies showed improvements on secondary, pre/post impairment measures such as the Aachen Aphasia Test (AAT; Huber, Poeck, & Willmes, 1984). AAT results in Wilssens et al. (2015) appeared to favour CIAT over traditional therapy but subtest results were reported without any overall profile scores. There were indications of positive effects on blinded clinician ratings on the Communicative Activity Log (CAL; Pulvermüller et al., 2001), while Boston Naming Test changes were variable between studies (BNT; Kaplan, Goodglass, & Weintraub, 1983). Changes in pre/post measures within SCEDs were obscured due to the crossover designs.

Multimodal Treatment

Combined Multimodal. There were two studies on combined multimodal treatment that investigated M-MAT. Both used multiple impairment-based outcomes. As reported above, the Tau-U effect size for M-MAT in Attard et al. (2013) was .81, lower than CIAT Plus but suggesting a strong effect of the treatment compared to baseline nonetheless. The Tau-U effect size was .68 for Rose et al. (2013), notably higher than CIAT Plus. The combined, weighted Tau-U for combined multimodal was .70, the second highest of all total effect sizes calculated; however, an effect size based on only two studies is far from conclusive.

Aside from confrontation naming of treated items, other impairment-based measures were obscured by the difficulty of calculating pre/post changes in crossover designs.

Gesture. All studies using gesture to facilitate verbal output that met the quality and inclusion criteria for gesture were SCEDs. Most compared the efficacy of gesture to another treatment (see Appendix 1), including repetition, semantic treatment and intention gestures. All used impairment outcomes, with most investigating confrontation naming and/or the Western Aphasia Battery (WAB, Kertesz, 2007).

As all eight studies used confrontation naming of treated items for probes, Tau-U was calculated for each one. The narrower confidence intervals in some of the Tau-U effect sizes and high relative weightings reflect the high number of data points, due to more participants or more probes in each phase (Figure 2-5). The overall weighted effect size for the eight SCED studies on gesture was .62, indicating that approximately 62% of treatment phase naming scores were higher than baseline scores. This is promising for the effects of gesture on confrontation naming of verbs and nouns. However, this included one negative (Ferguson, Evans, & Raymer, 2012), low (Raymer et al., 2012) and moderate effect size (Rodriguez, Raymer, & Gonzalez Rothi, 2006). Gesture therefore had the widest range of Tau-U effect sizes. These lower effect sizes are not attributable to poor quality methodology, as Raymer et al. (2012) scored 19 on the RoBiNT.

The small number of secondary, impairment-based outcome measures taken before and after gesture treatment alone showed variable changes. Mean changes were small for naming batteries. The mean change in those using the WAB Aphasia Quotient (WAB-R AQ) reached the critical difference of five for one study (Raymer et al., 2012) but not the other (Raymer et al., 2006).

Music. Despite 13 articles for music meeting the initial inclusion criteria — the second highest yielding category — nine were pre/post designs, along with two low quality nonRCTs. Only two studies with impairment-based outcomes reached adequate methodological quality (Hough, 2010; Stahl, Henseler, Turner, Geyer, & Kotz, 2013).

While Stahl et al. (2013) aimed to disentangle the contributions of rhythm and singing in melodic therapy, the results of the rhythm arm are not reported for this review. The singing group rehearsed common phrases whilst the conventional therapy group received treatment on other stimuli. After treatment, patients were assessed in their ability to sing and speak in unison with recordings of both trained phrases and untrained phrases. Written prompts were also provided. Unsurprisingly, the music group who had practised the trained phrases showed greater improvement than the conventional therapy group. The conventional therapy group demonstrated gains on the untrained phrases, however, while the singing group did not improve.

Hough (2010) used successful *repetition* of phrases as the primary outcome for their single participant. Thus, while the data presented in their case chart appears to show improvement, it is not comparable to confrontation naming tasks, and a Tau-U effect size was not calculated. Regardless, improvement of repetition performance was highly significant for both common and personalised phrases (p < .0001), and the WAB-R AQ and CQ (cortical quotient) showed large improvements (13 and 13.6 respectively).

Neither repetition (Hough, 2010) nor speaking in unison and with prompts (Stahl et al., 2013) is a test of generative verbal output. These two papers show only that singing specific phrases improves participants' ability to produce those phrases with maximum modelling and prompting.

Writing. There were no group studies utilising writing as a prompt for verbal output. All three high quality SCEDs used the same primary outcome measure of confrontation naming of treated items. These SCEDs had a high overall Tau-U effect size of .94, greater than all other modalities and constraint. This is close to the ceiling of 1.0 and suggests that the majority of treatment probes were improved over baseline.

In regards to pre/post impairment-based outcomes, the two participants in the multiple baseline design of Wright et al. (2008) demonstrated improvement in pre/post WAB-R AQ (3.1 and 9.7) and contradictory improvement in pre/post BNT (-3 and 5).

Table 2-2 – Impairment-based outcomes of high-quality studies

Treatment	Authors	Outcomes (Italics = secondary outcome)	Results
Constraint	(Pulvermüller et al., 2001)	AAT (comprehension, repetition, naming, token test)	Group x Time effect in favour of CIAT (F[1,15]=5.0; P<0.04)
		CAL (blinded clinician ratings)	*7/10 CIAT participants improved. Significant but statistical testing only conducted on seven (F[1,6] 10.5, P<0.01). Control group CAL scores not reported.*

Constraint	(Meinzer et al., 2007)	AAT	Time effect significant for both groups
			    Clinicians: t(9) = 7.05, p<.0001
			    Laypersons: t(9) = 5.65, p<.002
			Group x Time effect non-significant (F [1,18] = 1.26; p>.2)
			19/20 improved as per critical difference in manual

Constraint	(Wilssens et al., 2015)	*AAT*	*All participants improved on at least one subtest, no overall profile scores reported.*
			    CIAT group: Statistically significant improvement on 4/5 subtests
			    BOX group: Statistically significant improvement on 1/5 subtests

		BNT (/60) *SAT* *PALPA 51, 49, 8, 5, 6)*	*Pre/post CIAT group (p = .004), Pre/post BOX group (p = .094), no between group testing.* *No significant change in 8/9, no between group testing.* *Pre/post semantic scores favoured BOX, Pre/post phonological scores favoured CIAT*
Constraint	(Attard et al., 2013)	Confrontation naming – treated items	Tau-U 1.00
		WAB-R AQ (0-100) *BNT (/60)* *Cinderella Narrative Retell* *Semi-structured conversation*	*No change pre/post CIAT Plus* *Non-significant changes pre/post CIAT Plus* *No statistically sig. improvement* *No statistically sig. improvement*
Constraint	(Rose et al., 2013)	Confrontation naming – treated items	Tau-U 0.24
		WAB-R AQ (0-100)[a] *BNT (/60)[a]*	*Mean change 2.47 (range -3.00 to 7.70)* *Mean change 1.9 (range -6.0 to 15.0)*

Treatment	Authors	Outcomes (Italics = secondary outcome)	Results
		Semi-structured conversation[a]	*Substantive nouns mean change 4.67 (range -64.0 to 76.0), substantive verbs unchanged overall*
Constraint	(Kurland et al., 2012)	Confrontation naming – treated items	Tau-U 0.69
		BNT (/60)	*P1: 25 to 35 (though pre-treatment was 32); P2: 23 to 33*
		BDAE-3	*P1: Unchanged; P2: Responsive naming 6 to 11, others largely unchanged*
Combined Multimodal	(Attard et al., 2013)	Confrontation naming – treated items	Tau-U 0.81
		WAB-R AQ (0-100)	*Non-significant changes pre/post M-MAT*
		BNT (/60)	*Non-significant changes pre/post M-MAT*
		Cinderella Narrative Retell	*No statistically sig. improvement*
		Semi-structured conversation	*No statistically sig. improvement*
Combined Multimodal	(Rose et al., 2013)	Confrontation naming – treated items	Tau-U 0.68
		WAB-R AQ (0-100)[a]	*Mean change 1.40 (range -4.2 to 9)*
		BNT (/60)[a]	*Mean change 2.2 (range -9.0 to 15.0).*
		Semi-structured conversation[a]	*Substantive nouns mean change 20.4 (range -24.0 to 87.0), substantive verbs largely unchanged*
Gesture	(Boo & Rose, 2011)[b]	Confrontation naming – treated items	Tau-U 0.85 (all gesture sets)
Gesture	(Ferguson et al., 2012)	Confrontation naming – treated items	Tau-U -0.04 (pantomime)
Gesture	(Rodriguez et al., 2006)[b]	Confrontation naming – treated items	Tau-U 0.46 (all gesture sets)

Treatment	Authors	Outcomes (Italics = secondary outcome)	Results
Gesture	(Raymer et al., 2006)	Confrontation naming – treated items *WAB-R AQ (0-100)* *BNT (/60)* *ANT (/60)*	Tau-U 0.67, one participant's results not published due to poor response *Mean change 4.8 (p = 0.15)* *Mean change 1.2 (range -3 to 9)* *Mean change -1.4 (range -11.6 to 7.4)*
Gesture	(Rose et al., 2002)	Confrontation naming – treated items	Tau-U 0.74
Gesture	(Raymer et al., 2012)	Confrontation naming – treated items *WAB-R AQ (0-100) (pre/post gesture)* *BNT (/60) (pre/post gesture)*	Tau-U 0.25 *Mean change 5.5 (range 0.2 to 16.1)* *Mean change 3 (range -15 to 13)*
Gesture	(Rose & Sussmilch, 2008) [b]	Confrontation naming – treated items	Tau-U 0.73
Gesture	(Rose & Douglas, 2008) [b]	Confrontation naming – treated items	Tau-U 0.91
Music	(Stahl et al., 2013)	All in unison with recordings & written prompts: Singing & speaking trained phrases (% correct) Singing & speaking untrained phrases (% correct)	 Singing therapy: Mean change = 36.47, 95% CI [28.24, 44.70]; Standard therapy: Mean change = 4.98, 95% CI [−3.25, 13.21] Singing therapy: Mean change = −0.36, 95% CI [−2.62, 1.90]; Standard therapy: Mean change = 6.21, 95% CI [3.96, 8.47]
Music	(Hough, 2010)	Repetition of treated phrases *WAB-R AQ (0-100)* *WAB CQ (0-100)*	Pre/post change (p < 0.0001) *Improved 13.0* *Improved 13.6*

Treatment	Authors	Outcomes (Italics = secondary outcome)	Results
Writing	(Wright et al., 2008)	Confrontation naming – treated items *WAB-R AQ (0-100)* *BNT (/60)*	Tau-U 0.90 *Pre/post improvement – P1: 3.1, P2: 9.7* *Pre/post improvement – P1: -3; P2: 5*
Writing	(Hillis, 1989)	Confrontation naming – treated items	Tau-U 0.99
Writing	(Beeson & Egnor, 2006)	Confrontation naming – treated items *PALPA 53 (/40)*	Tau-U 0.86 *Unchanged*

Note. AAT = Aachen Aphasia Test (Profile score), CAL = Communicative Activity Log, BNT = Boston Naming Test, SAT = Verbal Semantic Association Test, PALPA = Psycholinguistic Assessment of Language Processing in Aphasia, WAB-R AQ/CQ = Western Aphasia Battery [Revised] (Aphasia Quotient/Cortical Quotient), BDAE = Boston Diagnostic Aphasia Examination, ANT = Action Naming Test, VAST = Verb And Sentence Test, OANB = Object Action Naming Battery.

[a]The crossover design makes it difficult to isolate the contributions of pre/post outcome measures, but taking a conservative approach, where the treatment of interest was administered second, the post treatment measure of the first treatment was taken as the baseline.

[b]Pre/post measures represent changes from multiple treatments and so are not listed.

Activity/Participation Outcomes

Constraint Treatment

Four high quality constraint studies included activity/participation outcomes: two RCTs (Pulvermüller et al., 2001; Wilssens et al., 2015) and two SCEDs (Attard et al., 2013; Rose et al., 2013). As displayed in Table 2-4, outcomes included the Communication Effectiveness Index (CETI; Lomas, Pickard, Bester, & Elbard, 1989), the Amsterdam-Nijmegen Everyday Language Test (ANELT, Blomert, Kean, Koster, & Schokker, 1994), the Scenario Test, and one used the CAL, a customised measure.

The improvements on the CAL self-ratings in Pulvermüller et al. (2001) were significant for constraint and not the control group, but no between-group comparisons were made and the control group received lower intensity therapy.

Wilssens et al. (2015) was the only controlled trial to compare constraint to a non-constraint treatment of the same intensity, a Dutch drill-based lexical-semantic therapy program, BOX . They did not find a statistically significant *between* group change on the CETI ($p = .332$). However, within group changes were significant for the BOX group and not the CIAT group. The mean CETI change for the BOX group was also above the clinically significant improvement level of 12 while it was less for the CIAT group (Lomas et al., 1989). There were no significant between group differences on the ANELT.

There was no clinically significant improvement in CETI scores for CIAT in Attard, Rose & Lanyon (2013) or in mean changes in Rose et al. (2013). Scenario Test scores improved in both participants for CIAT in Attard, Rose & Lanyon but the mean change was negligible for Rose et al.

Multimodal Treatment

Combined Multimodal. The CETI score changes post M-MAT in Attard, Rose and Lanyon (2013) did not reach the minimum clinically significant change of 12. In Rose et al. (2013) the mean change in CETI scores was 8.5 (range -2 to 33), again lower than the clinically significant change. Scenario Test changes in both studies were minimal.

Gesture. There was limited use of activity/participation measures in gesture studies, but the three studies that did use these were moderate to high quality SCEDs (RoBiNT 14-21). However, two of these were comparisons of multiple treatments (Boo & Rose, 2011; Rose & Sussmilch, 2008), and therefore pre/post measures of activity/participation cannot be considered as they represent changes from all treatments. The third, Raymer et al. (2012), used family member ratings on two measures — the CETI and the Functional Outcomes Questionnaire for Aphasia (FOQ-A; Glueckauf et al., 2003). Changes on the CETI were inconsistent. Two participants had a negative change — one clinically significant — and two had a positive change — one clinically significant. The mean change was 5.20 (range -15 to 33). On the FOQ-A changes were positive but small, with only one score changing greater than one standard deviation from the original FOQ-A paper (Glueckauf et al., 2003). The mean change was 0.44 (-0.19 to 1.21).

Music. A single high quality study investigated activity/participation outcomes in a multiple baseline design (Hough, 2010). The CETI was completed by the participant's caregiver and improved by 28.2, which is well above the clinically significant change of 12. The ASHA FACS (American Speech-Language Hearing Association Functional Assessment of Communication

Skills; Frattali, Thompson, Holland, & Wohl, 1995) improvement was 2.05/7 (to our

knowledge there are no established benchmarks for clinically significant change for the

ASHA FACS).

Writing. No high quality studies on writing used activity/participation outcomes.

Table 2-3 – Activity/Participation outcomes of high-quality studies

Treatment	Authors	Outcomes (Italics = secondary outcome)	Results
Constraint	(Pulvermüller et al., 2001)	*CAL (self ratings)*	*Pre/post CIAT group (F[1,7]=25.0, P<0.001); Pre/post control group (F<1); no between group comparison.*
Constraint	(Wilssens et al., 2015)	CETI (/100) (self and family ratings)	No between-group difference in improvements (t(6) = 1.01, *p* = .332); however: Pre/post CIAT group (t(4) = 1.47, *p* = .216), < clinically significant difference Pre/post BOX group (t(2) = 7.40, *p* = .019), > clinically significant difference
		ANELT	Statistically significant improvement for both groups; no significant difference between groups t(7) = –0.85, *p* = .426
Constraint	(Attard et al., 2013)	*CETI (/100)*	*Both less than clinically significant difference (12). P1: 3 point increase; P2: 3 point increase*
		Scenario test (/54)	*Improved by 3 and 8 points*
Constraint	(Rose et al., 2013)	*CETI (/100)[a]* *Scenario test (/54)[a]*	*Mean change 4 points (range -3 to 13). 2/11 participants > clinically significant difference* *Mean change -0.1/54 (range -10.0 to 7.70)*
Combined Multimodal	(Attard et al., 2013)	*CETI (/100)[a]*	*Both less than clinically significant difference (12). P1: 8 point increase; P2: 9 point increase*
		Scenario test (/54)	*P1: +1 point; P2: -3 points*
Combined Multimodal	(Rose et al., 2013)	*CETI (/100)[a]*	*Mean change 8.5 points (range -2 to 33). 3/11 participants > clinically significant difference*
		Scenario test (/54)[a]	*Mean change 0.80/54 (range -3.0 to 9.0)*
Gesture	(Raymer et al., 2012)	*CETI (pre/post gesture) (/100)[a]*	*Mean change 5.20 points (range -15 to 33). One participant > clinically significant difference, one participant negative change > clinically significant difference.*
		FOQ-A (pre/post gesture) (/5)[a]	*Mean change 0.44 (-0.19 to 1.21)*
Gesture	(Rose & Sussmilch, 2008)	*LCQ (/90)[b]*	*n/a*

Treatment	Authors	Outcomes *(Italics = secondary outcome)*	Results
Gesture	(Boo & Rose, 2011)	*LCQ (/90)[b]*	*n/a*
Music	(Hough, 2010)	*CETI (caregiver rating) (/100)* *ASHA FACS (/7)*	*28.2 increase (> clinically significant difference)* *2.05 increase*

Note. ANELT = Amsterdam-Nijmegen Everyday Language Test, CAL = Communicative Activity Log, CETI = Communication Effectiveness Index, LCQ = La Trobe Communication Questionnaire, FOQ-A = Functional Outcomes Questionnaire for Aphasia, ASHA FACS = American Speech-Language Hearing Association Functional Assessment of Communication Skills.

[a]The crossover design makes it difficult to isolate the contributions of pre/post outcome measures, but taking a conservative approach, where the treatment of interest was administered second, the post treatment measure of the first treatment was taken as the baseline.

[b]Pre/post measures represent changes from multiple treatments and so are not listed.

Quality of Life/Carer Burden Outcomes

Carer Burden

No studies were retrieved that used assessments of carer burden.

Constraint Treatment

In two constraint studies (Attard et al., 2013; Rose et al., 2013), quality of life was measured using the Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39; Hilari, Byng, Lamping, & Smith, 2003), but no others investigated this domain. Both administered SAQOL-39 before and after their crossover design comparing CIAT Plus and M-MAT. As they were not measured between crossover of treatments, it is not possible to attribute changes to one of M-MAT or CIAT Plus.

Multimodal Treatment

Combined Multimodal. As reported above, Rose et al. (2013) and Attard, Rose and Lanyon (2013) did not measure the SAQOL-39 between CIAT and M-MAT, therefore individual contributions of the treatments cannot be determined.

Gesture. No gesture studies were found that investigated quality of life.

Music. One study meeting quality criteria investigated quality of life (Hough, 2010). The single participant demonstrated a 25 point increase on the ASHA QCL (Paul et al., 2003)

— a 64% improvement from baseline score.

Writing. No writing studies investigated quality of life.

Table 2-4 – Quality of life outcomes of high-quality studies

Treatment	Authors	Outcomes	Results
Constraint	(Attard et al., 2013)	*SAQOL-39[a]*	*n/a*
Constraint	(Rose et al., 2013)	*SAQOL-39 [a]*	*n/a*
Combined Multimodal	(Attard et al., 2013)	*SAQOL-39 [a]*	*n/a*
Combined Multimodal	(Rose et al., 2013)	*SAQOL-39 [a]*	*n/a*
Music	(Hough, 2010)	*ASHA-QCL (/90)*	*25 point increase*

Note. SAQOL-39 = Stroke and Aphasia Quality of Life Scale-39, ASHA-QCL = American Speech-Language Hearing Association Quality of Communication Life Scale.
[a]Pre/post measures represent changes from multiple treatments and so are not listed

Discussion

The aim of this systematic review was to examine and compare evidence for constraint and multimodal therapies in chronic aphasia. Such a review is important for the treatment of people living with aphasia long term and for the evidence-based application of learned nonuse and multimodal cueing. The effectiveness of these approaches also inform our theoretical understanding of language processing. This review shows that there is limited high quality evidence to support the use of either constraint or multimodal approaches for impairment, activity/participation and quality of life outcomes. The amount and strength of evidence varied between communication modalities.

Impairment outcomes

There were few high-quality studies comparing constraint therapies to equivalent intensity controls, and for those that did, results did not favour constraint. Indeed, Tau-U scores for one study favoured multimodal over constraint. The existing research therefore fails to demonstrate superiority of CIAT over any non-constraint treatment in chronic aphasia. Comparisons aside, overall evidence does suggest positive effects for impairment outcomes in constraint, but data is far from conclusive and further research is required.

There is scant evidence for combined multimodal treatment with only two SCED studies. While this preliminary evidence has positive results in impairment measures with an effect size of 0.702 for naming, the effect size could change with further research. Changes in noun production during semi-structured conversation also hinted

at positive outcomes, but more research is indicated before any conclusions can be drawn.

Gesture evidence came from high quality SCED papers, yet effect sizes varied widely for impairment outcomes. Further investigation is needed to explain this variability. In addition to improvements in confrontation naming, there were signs of positive effects on other impairment-based assessments. However, these data are only preliminary as results were inconsistent and obscured at times by crossover designs. Group designs with control groups would address the limitations of crossover designs and allow direct comparison of changes on multiple outcome measures.

It is also worth noting that all eight SCEDs for gesture were conducted by Rose or Raymer. Though a variety of participants and study designs were used, the evidence would be enhanced with research from other authors. Group designs such as RCTs are also needed to confirm current findings.

Despite MIT being a well-known treatment for aphasia, in chronic aphasia the evidence is of low quality in terms of both study design and methodology. Future studies into music-based treatment of verbal output in aphasia need to employ more rigorous research designs. Rather than investigating speech production in unison or repetition, research should investigate the presumed end goal of such treatments — the independent production of trained words or phrases — and probe for generalisation to discourse and conversation.

Results for impairment outcomes from writing are very positive thus far but inconclusive due to low replication. It is possible that research specifically examining the effects of writing on verbal output are limited because pairing writing with speech is already widely accepted and frequently embedded in cueing hierarchies. Nonetheless,

further research is necessary to confirm the impact of writing on verbal output, especially group studies.

No studies were found investigating drawing. Research on this promising modality is needed.

In summary, there is encouraging but insufficient evidence for constraint or multimodal treatments on impairment-based measures, in terms of both quality and quantity.

Activity & Participation

In group studies on constraint therapy, there was insufficient evidence for effects on Activity/Participation following treatment, both in terms of pre/post improvement and between-group comparisons. CIAT Plus and M-MAT changes in the CETI, though not clearly attributable to either treatment due to the crossover designs, suggested positive outcomes, again with high variability amongst participants. In gesture studies, the limited results for Activity/Participation measures were contradictory and inconclusive. Improvement and deterioration of scores effectively cancelled each other out, with disparity even between self and carer improvement ratings at times. In music studies, there was only quality data for a single participant, though this was clinically and statistically significant.

Quality of Life and Carer Burden

Quality of life outcomes were rarely investigated. A subjective improvement in a person's life experience should be a goal of aphasia therapy, yet in this review, only five studies employed quality of life outcomes, and only three were high quality.

Similarly, the effect of treatment on carer burden remains unexplored. Carers of people with aphasia experience changes such as loneliness, anxiety, increased responsibilities and need for support (Patrício et al., 2013), and supporting them should also be a crucial goal of rehabilitation (Rombough, Howse, & Bartfay, 2006). Yet the impact of treatment on carers and family was not investigated in any study.

Future research should include both quality of life and carer burden as outcome measures.

Systematic review design

In this systematic review, we chose to include high quality SCEDs. The inclusion of SCEDs allowed closer inspection of research areas that would have returned no results if a traditional evidence hierarchy was used. We made this decision based on increasing recognition that high quality SCEDs can have equal or even superior rigour to RCTs (e.g. Medical N-of-1 designs) (Tate et al., 2016).

Use of the RoBiNT scale allowed us to rank SCEDs according to specific features that contribute to internal and external validity. This approach is more fine-grained than simply classifying SCEDs by design subtype or as experimental/non-experimental. As RoBiNT is a relatively new scale there is not yet an established cutoff for what constitutes a quality study. Our use of 12 as the cutoff was based on early uses of this

new scale. While this formed a limited empirical basis, the widely accepted cutoff of 5

for the PEDro/PEDro-P is also based on common scores (Teasell et al., 2007). Neither

approach considers relative weighting of individual items. Nevertheless, higher scoring

studies will have stronger methodological quality and internal validity and in our results

there were no borderline papers on the RoBiNT that were excluded.

As far as we are aware, this is the first time that the Tau-U effect size

calculations have been applied within a systematic review of aphasia treatments. The

Tau-U effect size provided ranking of SCEDs based on improved data points in the

treatment phases, after correction for baseline trend. These effect sizes allowed

comparisons *between* papers in this study but we know of no external benchmarks for

Tau-U effect sizes. In addition, many SCEDs analysed with Tau-U had gaps in treatment

probe intervals (e.g. washout periods between crossover trials) or irregular probing,

which may have influenced the validity of resulting effect sizes. Another challenge in

utilising Tau-U is that there is no agreed protocol yet; for example, the threshold for

applying baseline trend correction varies between studies. The first author therefore

combined visual inspection with the baseline Tau to eliminate overcorrection (see

methods). While this did introduce a subjective element to calculation, the overall

process remains a defensible calculation of effect sizes in a family of designs previously

resistive to meta-analysis.

Finally, the exclusion of non-English articles is an unfortunate, but in our case,

unavoidable, criterion. It is possible that relevant research published in languages other

than English was missed by this systematic review.

Treatment reporting

Interventions should be reported thoroughly to allow future replication and synthesis of results (Hoffmann, Glasziou, Boutron, & Milne, 2014). A problem noted across retrieved constraint studies, regardless of quality or design, was the disparity in what was considered to constitute constraint. While the original authors have recently clarified that self-cueing with actions (e.g., gesture) is permitted as long as it is not communicative (Difrancesco et al., 2012), prior to this publication some protocols banned gesture and the majority made no comment on this aspect, including the original CIAT article (Pulvermüller et al., 2001). Likewise, whether or not therapists provided cues to participants was often not described in methods and is not addressed in Difrancesco, Pulvermüller and Mohr (2012). The term "constraint" therefore currently represents therapies with significant procedural differences. Without clearer reporting of methods, there is a risk of continued bleeding of the term "constraint" to an increasingly diverse range of game-based language treatments. All future constraint-based research should provide comprehensive description of methods and state explicitly in which ways they depart from the outline in Difrancesco, Pulvermüller and Mohr (2012). A template such as TIDieR (Hoffmann et al., 2014) also provides a framework for thorough reporting. Without this detail, it will be difficult to determine the aspects of CIAT that contribute to effectiveness.

Outcome measures

The majority of studies in this review focused on improvements at the Impairment level. However, outcomes in aphasia research should be those that are

important to people with aphasia and their families (Wallace, 2016). Recent work has shown that these include a range of outcomes across the ICF as well as quality of life and patient satisfaction with treatment (Wallace et al., 2016). While these constructs can be difficult to measure directly, especially in people with aphasia (e.g.; Szaflarski et al., 2015), without the inclusion of such outcomes in future, the benefit of these treatments to people living with aphasia will remain unknown.

There was also a high number of different outcomes within this review. Heterogeneity of outcome measures reduces research efficiency by limiting synthesis and meta-analysis of results, which is an important way to overcome the small sample sizes that are common in aphasia research (Brady et al., 2016). The development of a core outcome for aphasia research set is currently underway which will recommend preferred outcomes measures for the constructs identified as important to those with aphasia and other stakeholders (Wallace, Worrall, Rose, & Le Dorze, n.d.). The core outcome set should be adhered to in future research wherever possible.

Conclusion

Overall, this review has found a limited evidence base for constraint therapy in chronic aphasia, especially in proportion to its prominence in research and clinical practice. While studies indicated positive outcomes, there is a need for rigorous high level studies comparing CIAT and its derivatives to non-constraint therapies or controls. We also found a very limited evidence base for multimodal therapies. Studies on some modalities had limited research of any quality (drawing, writing, combined multimodal) while others had more research but little of adequate quality (music, gesture).

Accordingly, there is insufficient data to suggest superiority of either constraint or multimodal approaches in chronic aphasia. There were not enough comparable high-quality group studies to perform meta-analysis. Meta-analysis of SCEDs favoured multimodal treatments but this is not yet conclusive.

In addition, there was insufficient examination of "real world" endpoints. Aphasia research needs to expand beyond the use of basic impairment outcomes such as confrontation naming and toward consistent outcome measures based on the wishes of people with aphasia and their families.

Examination of constraint and multimodalities against control treatments of equal intensity and duration are needed, as well as direct, rigorous comparison between constraint and multimodal treatments. It is further recommended that future research provides a comprehensive description of treatment methods and readily accessible treatment materials arising from clinical trials to enable translation into practice.

Clinicians should not adopt either treatment approach exclusively until further research is published demonstrating the superiority of one treatment, or, more likely, the suitability of each to particular patient characteristics.

References

Abo, M., Kakuda, W., Watanabe, M., Morooka, A., Kawakami, K., & Senoo, A. (2012). Effectiveness of low-frequency rTMS and intensive speech therapy in poststroke patients with aphasia: A pilot study based on evaluation by fMRI in relation to type of aphasia. *European Neurology, 68*(4), 199–208. https://doi.org/10.1159/000338773

Al-Janabi, S., Nickels, L. A., Sowman, P. F., Burianová, H., Merrett, D. L., & Thompson, W. F. (2014). Augmenting melodic intonation therapy with non-invasive brain stimulation to treat impaired left-hemisphere function: Two case studies. *Frontiers in Psychology, 5*(37), 1–12. https://doi.org/10.3389/fpsyg.2014.00037

Attard, M. C., Rose, M. L., & Lanyon, L. E. (2013). The comparative effects of Multi-Modality Aphasia Therapy and Constraint-Induced Aphasia Therapy-Plus for severe chronic Broca's aphasia: An in-depth pilot study. *Aphasiology, 27*(1), 80–111. https://doi.org/10.1080/02687038.2012.725242

Ball, A. L., de Riesthal, M., Breeding, V. E., & Mendoza, D. E. (2011). Modified ACT and CART in severe aphasia. *Aphasiology, 25*(6-7), 836–848. https://doi.org/10.1080/02687038.2010.544320

Barbancho, M. A., Berthier, M. L., Navas-Sanchez, P., Dàvila, G., Green-Heredia, C., Garcia-Alberca, J. M., et al. (2015). Bilateral brain reorganization with memantine and constraint-induced aphasia therapy in chronic post-stroke aphasia: An ERP study. *Brain and Language*, 145-146, 1–10. https://doi.org/10.1016/j.bandl.2015.04.003

Beattie, G., & Shovelton, H. (2006). A critical appraisal of the relationship between speech and gesture and its implications for the treatment of aphasia. *International Journal of Speech-Language Pathology, 8*(2), 134–139. https://doi.org/10.1080/14417040600667392

Beeson, P. M., & Egnor, H. (2006). Combining treatment for written and spoken naming. *Journal of the International Neuropsychological Society, 12*, 816–827. https://doi.org/10.10170S1355617706061005

Berthier, M. L., Green, C., Lara, J. P., Higueras, C., Barbancho, M. A., Dàvila, G., & Pulvermüller, F. (2009). Memantine and constraint-induced aphasia therapy in chronic poststroke aphasia. *Annals of Neurology, 65*(5), 577–585. https://doi.org/10.1002/ana.21597

Blomert, L., Kean, M. L., Koster, C., & Schokker, J. (1994). Amsterdam—Nijmegen everyday language test: construction, reliability and validity. *Aphasiology, 8*(4), 381–407. https://doi.org/10.1080/02687039408248666

Bonakdarpour, B., Eftekharzadeh, A., & Ashayeri, H. (2003). Melodic Intonation Therapy in Persian aphasic patients. *Aphasiology, 17*(1), 75–95. https://doi.org/10.1080/02687030244000464

Boo, M., & Rose, M. L. (2011). The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in Broca's aphasia. *Aphasiology, 25*(2), 154–175. https://doi.org/10.1080/02687031003743789

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *The Cochrane Database of Systematic Reviews, 6*(6), CD000425–CD000425. https://doi.org/10.1002/14651858.CD000425.pub4

Breier, J. I., Juranek, J., & Papanicolaou, A. C. (2011). Changes in maps of language function and the integrity of the arcuate fasciculus after therapy for chronic aphasia. *Neurocase, 17*(6), 506–517. https://doi.org/10.1080/13554794.2010.547505

Breier, J. I., Juranek, J., Maher, L. M., Schmadeke, S., Men, D., & Papanicolaou, A. C. (2009). Behavioral and neurophysiologic response to therapy for chronic aphasia. *Archives of Physical Medicine and Rehabilitation, 90*(12), 2026–2033. https://doi.org/10.1016/j.apmr.2009.08.144

Breier, J. I., Maher, L. M., Novak, B., & Papanicolaou, A. C. (2006). Functional imaging before and after constraint-induced language therapy for aphasia using magnetoencephalography. *Neurocase, 12*(6), 322–331. https://doi.org/10.1080/13554790601126054

Breier, J. I., Maher, L. M., Schmadeke, S., Hasan, K. M., & Papanicolaou, A. C. (2007). Changes in language-specific brain activation after therapy for aphasia using magnetoencephalography: A case study. *Neurocase, 13*(3), 169–177. https://doi.org/10.1080/13554790701448200

Breier, J. I., Randle, S., Maher, L. M., & Papanicolaou, A. C. (2010). Changes in maps of language activity activation following melodic intonation therapy using magnetoencephalography: two case studies. *Journal of Clinical and Experimental Neuropsychology, 32*(3), 309–314. https://doi.org/10.1080/13803390903029293

Brossart, D. F., Vannest, K. J., Davis, J. L., & Patience, M. A. (2014). Incorporating nonoverlap indices with visual analysis for quantifying intervention effectiveness in single-case experimental designs. *Neuropsychological Rehabilitation, 24*(3-4), 464–491. https://doi.org/10.1080/09602011.2013.868361

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-Case Research Design and Analysis: New Directions for Psychology and Education* (pp. 187–212). Hillsdale, NJ: Lawrence Erlbaum Associates.

Carragher, M., Sage, K., & Conroy, P. (2013). The effects of verb retrieval therapy for people with non-fluent aphasia: Evidence from assessment tasks and conversation. *Neuropsychological Rehabilitation, 23*(6), 846–887. https://doi.org/10.1080/09602011.2013.832335

Centre for Evidence-Based Physiotherapy. (2016). PEDro Statistics. Retrieved June 14, 2016, from http://www.pedro.org.au/english/downloads/pedro-statistics/

Cubelli, R. (1995). More on drawing in aphasia therapy. *Aphasiology, 9*(1), 78–83. https://doi.org/10.1080/02687039508248693

Davis, G. A. (2005). PACE revisited. *Aphasiology, 19*(1), 21–38. https://doi.org/10.1080/02687030444000598

DeDe, G., Parris, D., & Waters, G. (2003). Teaching self-cues: A treatment approach for verbal naming. *Aphasiology, 17*(5), 465–480. https://doi.org/10.1080/02687030344000094

Difrancesco, S., Pulvermüller, F., & Mohr, B. (2012). Intensive language-action therapy (ILAT): The methods. *Aphasiology, 26*(11), 1317–1351. https://doi.org/10.1080/02687038.2012.705815

Farias, D., Davis, C., & Harrington, G. (2006). Drawing: Its contribution to naming in aphasia. *Brain and Language, 97*(1), 53–63. https://doi.org/10.1016/j.bandl.2005.07.074

Faroqi-Shah, Y., & Virion, C. R. (2009). Constraint-induced language therapy for agrammatism: Role of grammaticality constraints. *Aphasiology, 23*(7-8), 977–988. https://doi.org/10.1080/02687030802642036

Ferguson, N. F., Evans, K., & Raymer, A. M. (2012). A comparison of intention and pantomime gesture treatment for noun retrieval in people with aphasia. *American Journal of Speech-Language Pathology, 21*(2), s126–s139.

Feyereisen, P. (2006). How could gesture facilitate lexical access? *International Journal of Speech-Language Pathology, 8*(2), 128–133. https://doi.org/10.1080/14417040600667293

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61*(6), 825–850. https://doi.org/10.1080/17470210701623605

Fitzpatrick, R. B. (2008). PsycBITE™: Psychological Database for Brain Impairment Treatment Efficacy. *Journal of Electronic Resources in Medical Libraries, 5*(2), 171–178. https://doi.org/10.1080/15424060802064428

Flowers, H. L., Silver, F. L., Fang, J., Rochon, E., & Martino, R. (2013). The incidence, co-occurrence, and predictors of dysphagia, dysarthria, and aphasia after first-ever acute ischemic stroke. *Journal of Communication Disorders, 46*(3), 238–248. https://doi.org/10.1016/j.jcomdis.2013.04.001

Frattali, C. M., Thompson, C. M., Holland, A. L., & Wohl, C. B. (1995). *ASHA functional assessment of communication skills* (FACS). Rockville, MD: American Speech-Language Hearing Association.

Glueckauf, R. L., Blonder, L. X., Ecklund-Johnson, E., Maher, L. M., Crosson, B., & Gonzalez-Rothi, L. (2003). Functional Outcome Questionnaire for Aphasia: overview and preliminary psychometric evaluation. *NeuroRehabilitation, 18*(4), 281–290.

Goldfarb, R., & Bader, E. (1979). Espousing melodic intonation therapy in aphasia rehabilitation: A case study. *International Journal of Rehabilitation Research, 2*(3), 333–342.

Goral, M., & Kempler, D. (2009). Training verb production in communicative context: Evidence from a person with chronic non-fluent aphasia. *Aphasiology, 23*(12), 1383–1397. https://doi.org/10.1080/02687030802235203

Hadar, U., & Rumiati, R. I. (2006). I gesture therefore I speak: Cartesian cures. *International Journal of Speech-Language Pathology, 8*(2), 140–142. https://doi.org/10.1080/14417040600667301

Hilari, K. (2011). The impact of stroke: are people with aphasia different to those without? *Disability and Rehabilitation, 33*(3), 211–218. https://doi.org/10.3109/09638288.2010.508829

Hilari, K., Byng, S., Lamping, D. L., & Smith, S. C. (2003). Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of Acceptability, Reliability, and Validity. *Stroke, 34*(8), 1944–1950. https://doi.org/10.1161/01.STR.0000081987.46660.ED

Hillis, A. E. (1989). Efficacy and generalization of treatment for aphasic naming errors. *Archives of Physical Medicine & Rehabilitation, 70*(8), 632–636.

Hoffmann, T. C., Glasziou, P. P., Boutron, I., & Milne, R. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ, 348*, g1687–g1687.

Hough, M. S. (2010). Melodic Intonation Therapy and aphasia: Another variation on a theme. *Aphasiology, 24*(6-8), 775–786. https://doi.org/10.1080/02687030903501941

Huber, W., Poeck, K., & Willmes, K. (1984). The Aachen Aphasia Test. Advances in *Neurology, 42*, 291–303.

Hurkmans, J., de Bruijn, M., Boonstra, A. M., Jonkers, R., Bastiaanse, R., Arendzen, H., & Reinders-Messelink, H. A. (2012). Music in the treatment of neurological language and speech disorders: A systematic review. *Aphasiology, 26*(1), 1–19. https://doi.org/10.1080/02687038.2011.602514

Johnson, M. L., Taub, E., Harper, L. H., Wade, J. T., Bowman, M. H., Bishop-McKay, S., et al. (2014). An enhanced protocol for constraint-induced aphasia therapy II: A case series. *American Journal of Speech-Language Pathology, 23*(1), 60–72. https://doi.org/10.1044/1058-0360(2013/12-0168)

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston naming test.* Philadelphia, PA: Lea & Febiger.

Kavian, S., Khatoonabadi, A. R., Ansari, N. N., Saadati, M., & Shaygannejad, V. (2014). A single-subject study to examine the effects of constrained-induced aphasia therapy on naming deficit. *International Journal of Preventive Medicine, 5*(6), 782–786.

Kempler, D., & Goral, M. (2011). A comparison of drill- and communication-based treatment for aphasia. *Aphasiology, 25*(11), 1327–1346. https://doi.org/10.1080/02687038.2011.599364

Kertesz, A. (2007). *Western Aphasia Battery* (Revised). San Antonio, TX: PsychCorp.

Kiran, S. (2005). Training phoneme to grapheme conversion for patients with written and oral production deficits: A model-based approach. *Aphasiology, 19*(1), 53–76. https://doi.org/10.1080/02687030444000633

Kurland, J., Baldwin, K., & Tauer, C. (2010). Treatment-induced neuroplasticity following intensive naming therapy in a case of chronic Wernicke's aphasia. *Aphasiology, 24*(6-8), 737–751. https://doi.org/10.1080/02687030903524711

Kurland, J., Pulvermüller, F., Silva, N., Burke, K., & Andrianopoulos, M. (2012). Constrained versus unconstrained intensive language therapy in two individuals with chronic, moderate-to-severe aphasia and apraxia of speech: Behavioral and fMRI outcomes. *American Journal of Speech-Language Pathology, 21*(2), S65–S87. https://doi.org/10.1044/1058-0360(2012/11-0113)

Lam, J. M. C., & Wodchis, W. P. (2010). The Relationship of 60 Disease Diagnoses and 15 Conditions to Preference-Based Health-Related Quality of Life in Ontario Hospital-Based Long-Term Care Residents. *Medical Care, 48*(4), 380–387. https://doi.org/10.1097/MLR.0b013e3181ca2647

Lim, K.-B., Kim, Y.-K., Lee, H.-J., Yoo, J., Hwang, J. Y., Kim, J.-A., & Kim, S.-K. (2013). The therapeutic effect of neurologic music therapy and speech language therapy in post-stroke aphasic patients. *Annals of Rehabilitation Medicine, 37*(4), 556–562. https://doi.org/10.5535/arm.2013.37.4.556

Lomas, J., Pickard, L., Bester, S., & Elbard, H. (1989). The Communicative Effectiveness Index: Development and Psychometric Evaluation of a Functional Communication Measure for Adult Aphasia. *Journal of Speech and Hearing Disorders, 54*, 113–124.

Lorenz, A., & Nickels, L. A. (2007). Orthographic cueing in anomic aphasia: How does it work? *Aphasiology, 21*(6-8), 670–686. https://doi.org/10.1080/02687030701192182

Luria, A. R. (1970). *Traumatic aphasia: Its syndromes, psychology and treatment*. (D. Bowden, Trans.). The Hague, Netherlands: Mouton & Co.

MacGregor, L. J., Difrancesco, S., Pulvermüller, F., Shtyrov, Y., & Mohr, B. (2015). Ultra-rapid access to words in chronic aphasia: The effects of intensive language action therapy (ILAT). *Brain Topography, 28*(2), 279–291. https://doi.org/10.1007/s10548-014-0398-y

Maher, L. M., Kendall, D. L., Swearengin, J. A., Rodriguez, A., Leon, S. A., Pingel, K., et al. (2006). A pilot study of use-dependent learning in the context of Constraint Induced Language Therapy. *Journal of the International Neuropsychological Society, 12*(6), 843–852. https://doi.org/10.1017/S1355617706061029

Makuuchi, M., Kaminaga, T., & Sugishita, M. (2003). Both parietal lobes are involved in drawing: a functional MRI study and implications for constructional apraxia. *Cognitive Brain Research, 16*(3), 338–347. https://doi.org/10.1016/S0926-6410(02)00302-6

Marangolo, P., Bonifazi, S., Tomaiuolo, F., Craighero, L., Coccia, M., Altoè, G., et al. (2010). Improving language without words: First evidence from aphasia. *Neuropsychologia, 48*(13), 3824–3833. https://doi.org/10.1016/j.neuropsychologia.2010.09.025

Martin, P. I., Treglia, E., Naeser, M. A., Ho, M. D., Baker, E. H., Martin, E. G., et al. (2014). Language improvements after TMS plus modified CILT: Pilot, open-protocol

study with two, chronic nonfluent aphasia cases. *Restorative Neurology and Neuroscience, 32*(4), 483–505. https://doi.org/10.3233/RNN-130365

Maul, K. K., Conner, P. S., Kempler, D., Radvanski, C., & Goral, M. (2014). Using informative verbal exchanges to promote verb retrieval in nonfluent aphasia. *American Journal of Speech-Language Pathology, 23*(3), 407–420. https://doi.org/10.1044/2014_AJSLP-13-0004

McNeill, D. (1992). *Hand and Mind*. Chicago, IL: University of Chicago Press.

Meinzer, M., Djundja, D., Barthel, G., Elbert, T., & Rockstroh, B. (2005). Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy. *Stroke, 36*(7), 1462–1466. https://doi.org/10.1161/01.STR.0000169941.29831.2a

Meinzer, M., Streiftau, S., & Rockstroh, B. (2007). Intensive language training in the rehabilitation of chronic aphasia: Efficient training by laypersons. *Journal of the International Neuropsychological Society, 13*(5), 846–853. https://doi.org/10.1017/S1355617707071111

Mohr, B., Difrancesco, S., Harrington, K., Evans, S., & Pulvermüller, F. (2014). Changes of right-hemispheric activation after constraint-induced, intensive language action therapy in chronic aphasia: fMRI evidence from auditory semantic processing. *Frontiers in Human Neuroscience, 8*(309), 919. https://doi.org/10.3389/fnhum.2014.00919

Morrow-Odom, K. L., & Swann, A. B. (2013). Effectiveness of melodic intonation therapy in a case of aphasia following right hemisphere stroke. *Aphasiology, 27*(11), 1322–1338. https://doi.org/10.1080/02687038.2013.817522

Nickels, L. A. (1992). The autocue? self-generated phonemic cues in the treatment of a disorder of reading and naming. *Cognitive Neuropsychology, 9*(2), 155–182. https://doi.org/10.1080/02643299208252057

Nickels, L. A. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology, 16*(10-11), 935–979. https://doi.org/10.1080/02687030244000563

OCEBM Levels of Evidence Working Group. (2011). *The Oxford Levels of Evidence 2*. Oxford Centre for Evidence-Based Medicine. Retrieved from http://www.cebm.net/index.aspx?o=5653

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining Nonoverlap and Trend for Single-Case Research: Tau-U. *Behavior Therapy, 42*(2), 284–299. https://doi.org/10.1016/j.beth.2010.08.006

Patrício, B., Jesus, L., & Cruice, M. (2013). *Quality of life of the caregivers of people with aphasia. A systematic review*. [Master's Thesis]. Citeseer. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.926.7095&rep=rep1&type=pdf

Paul, D. R., Frattali, C. M., Holland, A. L., Thompson, C. K., Caperton, C. J., & Slater, S. C. (2003). *ASHA Quality of Communication Life Scale* (QCL). Rockville, MD: American Speech-Language-Hearing Association.

Pulvermüller, F., & Berthier, M. L. (2008). Aphasia therapy on a neuroscience basis. *Aphasiology, 22*(6), 563–599. https://doi.org/10.1080/02687030701612213

Pulvermüller, F., Hauk, O., Zohsel, K., Neininger, B., & Mohr, B. (2005). Therapy-related reorganization of language in both hemispheres of patients with chronic aphasia. *NeuroImage, 28*(2), 481–489. https://doi.org/10.1016/j.neuroimage.2005.06.038

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke, 32*(7), 1621–1626. https://doi.org/10.1161/01.STR.32.7.1621

Ravenek, M. J., Skarakis-Doyle, E., Spaulding, S. J., Jenkins, M. E., & Doyle, P. C. (2015). Enhancing the conceptual clarity and utility of the international classification of functioning, disability & health: the potential of a new graphic representation. *Disability and Rehabilitation, 35*(12), 1015–1025. https://doi.org/10.3109/09638288.2012.717582

Raymer, A. M., McHose, B., Smith, K. G., Iman, L., Ambrose, A., & Casselton, C. (2012). Contrasting effects of errorless naming treatment and gestural facilitation for word retrieval in aphasia. *Neuropsychological Rehabilitation, 22*(2), 235–266. https://doi.org/10.1080/09602011.2011.618306

Raymer, A. M., Singletary, F., Rodriguez, A., Ciampitti, M., Heilman, K. M., & Gonzalez Rothi, L. J. (2006). Effects of gesture+verbal treatment for noun and verb retrieval in aphasia. *Journal of the International Neuropsychological Society, 12*(6), 867–882. https://doi.org/10.10170S1355617706061042

Richter, M., Miltner, W. H. R., & Straube, T. (2008). Association between therapy outcome and right-hemispheric activation in chronic aphasia. *Brain, 131*(Pt 5), 1391–1401. https://doi.org/10.1093/brain/awn043

Rodriguez, A. D., Raymer, A. M., & Gonzalez Rothi, L. J. (2006). Effects of gesture+verbal and semantic-phonologic treatments for verb retrieval in aphasia. *Aphasiology, 20*(2-4), 286–297. https://doi.org/10.1080/02687030500474898

Rombough, R. E., Howse, E. L., & Bartfay, W. J. (2006). Caregiver strain and caregiver burden of primary caregivers of stroke survivors with and without aphasia. *Rehabilitation Nursing, 31*(5), 199–209.

Rose, M. L. (2006). The utility of arm and hand gestures in the treatment of aphasia. *Advances in Speech Language Pathology, 8*(2), 92–109. https://doi.org/10.1080/14417040600657948

Rose, M. L. (2013). Releasing the constraints on aphasia therapy: the positive impact of gesture and multimodality treatments. *American Journal of Speech-Language Pathology, 22*(2), S227–39. https://doi.org/10.1044/1058-0360(2012/12-0091)

Rose, M. L., & Douglas, J. (2008). Treating a semantic word production deficit in aphasia with verbal and gesture methods. *Aphasiology, 22*(1), 20–41. https://doi.org/10.1080/02687030600742020

Rose, M. L., & Sussmilch, G. (2008). The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology, 22*(7-8), 691–706. https://doi.org/10.1080/02687030701800800

Rose, M. L., Attard, M. C., Mok, Z., Lanyon, L. E., & Foster, A. M. (2013). Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: A phase 1 study. *Aphasiology, 27*(8), 938–971. https://doi.org/10.1080/02687038.2013.810329

Rose, M. L., Douglas, J., & Matyas, T. (2002). The comparative effectiveness of gesture and verbal treatments for a specific phonologic naming impairment. *Aphasiology, 16*(10-11), 1001–1030. https://doi.org/10.1080/02687030143000825

Sacchett, C. (2002). Drawing in aphasia: moving towards the interactive. International *Journal of Human-Computer Studies, 57*(4), 263–277. https://doi.org/10.1006/ijhc.2002.1018

Schlaug, G., Marchina, S., & Norton, A. (2008). From singing to speaking: why singing may lead to recovery of expressive language function in patients with broca's aphasia. *Music Perception, 25*(4), 315–323. https://doi.org/10.1525/MP.2008.25.4.315

Sherrington, C., Herbert, R. D., Maher, C. G., & Moseley, A. M. (2000). PEDro. A database of randomized trials and systematic reviews in physiotherapy. *Manual Therapy, 5*(4), 223–226. https://doi.org/10.1054/math.2000.0372

Smania, N. (2006). Constraint-induced movement therapy: an original concept in rehabilitation. *Europa Medicophysica, 42*(3), 239–240.

Sparks, R., Helm, N., & Albert, M. (1974). Aphasia rehabilitation resulting from melodic intonation therapy. Cortex, 10(4), 303–316.

Stahl, B., Henseler, I., Turner, R., Geyer, S., & Kotz, S. A. (2013). How to engage the right brain hemisphere in aphasics without even singing: Evidence for two paths of speech recovery. *Frontiers in Human Neuroscience, 7*(Article 35), 1–12. https://doi.org/10.3389/fnhum.2013.00035

Stahl, B., Kotz, S. A., Henseler, I., Turner, R., & Geyer, S. (2011). Rhythm in disguise: why singing may not hold the key to recovery from aphasia. *Brain, 134*(Pt 10), 3083–3093. https://doi.org/10.1093/brain/awr240

Sugishita, M., Seki, K., Kabe, S., & Yunoki, K. (1993). A material-control single-case study of the efficacy of treatment for written and oral naming difficulties. *Neuropsychologia, 31*(6), 559–569.

Szaflarski, J. P., Ball, A. L., Vannest, J., Dietz, A. R., Allendorfer, J. B., Martin, A. N., et al. (2015). Constraint-Induced Aphasia Therapy for Treatment of Chronic Post-Stroke Aphasia: A Randomized, Blinded, Controlled Pilot Trial. *Medical Science Monitor, 21*, 2861–2869. https://doi.org/10.12659/MSM.894291

Szaflarski, J. P., Ball, A., Grether, S., Al-fwaress, F., Griffith, N. M., Neils-Strunjas, J., et al. (2008). Constraint-induced aphasia therapy stimulates language recovery in

patients with chronic aphasia after ischemic stroke. *Medical Science Monitor, 14*(5), CR243–CR250.

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., et al. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. *Aphasiology, 30*(7), 862–876. https://doi.org/10.1080/02687038.2016.1178022

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: The 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation, 23*(5), 619–638. https://doi.org/10.1080/09602011.2013.824383

Tate, R. L., Rosenkoetter, U., Wakim, D., Sigmundsdottir, L., Doubleday, J., Togher, L., et al. (2015). *The Risk of Bias in N-of-1 Trials (RoBiNT) Scale: An expanded manual for the critical appraisal of single-case reports*. Sydney, Australia: Author.

Taub, E. (1976). Movement in nonhuman primates deprived of somatosensory feedback. *Exercise and Sport Sciences Reviews, 4*, 335–374.

Taub, E., Uswatte, G., Mark, V. W., & Morris, D. M. (2006). The learned nonuse phenomenon: implications for rehabilitation. *Europa Medicophysica, 42*(3), 241–255.

Teasell, R., Bayona, N., Marshall, S., Cullen, N., Bayley, M., Chundamala, J., et al. (2007). A systematic review of the rehabilitation of moderate to severe acquired brain injuries. *Brain Injury, 21*(2), 107–112. https://doi.org/10.1080/02699050701201524

van der Meulen, I., van de Sandt-Koenderman, M. E., & Ribbers, G. M. (2012). Melodic Intonation Therapy: Present controversies and future opportunities. *Archives of Physical Medicine and Rehabilitation, 93*(S1), S46–S52. https://doi.org/10.1016/j.apmr.2011.05.029

Vines, B. W., Norton, A. C., & Schlaug, G. (2011). Non-invasive brain stimulation enhances the effects of melodic intonation therapy. *Frontiers in Psychology, 2*(Article 230), 230. https://doi.org/10.3389/fpsyg.2011.00230

Wallace, S. (2016, December 18). *Improving Research Outcome Measurement in Aphasia (ROMA): Development of a Core Outcome Set*. Retrieved from http://espace.library.uq.edu.au/view/UQ:415571

Wallace, S. J., Worrall, L. E., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., et al. (2016). Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disability and Rehabilitation, 91*, 1–16. https://doi.org/10.1080/09638288.2016.1194899

Wallace, S., Worrall, L. E., Rose, T., & Le Dorze, G. (n.d.). *Improving Research Outcome Measurement in Aphasia (ROMA): Development of a Core Outcome Set*. Retrieved March 13, 2017, from http://www.cometinitiative.org/studies/details/287

Wan, C. Y., Zheng, X., Marchina, S., Norton, A., & Schlaug, G. (2014). Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. *Brain and Language, 136*, 1–7. https://doi.org/10.1016/j.bandl.2014.03.011

Weill-Chounlamountry, A., Capelle, N., Tessier, C., & Pradat-Diehl, P. (2013). Multimodal therapy of word retrieval disorder due to phonological encoding dysfunction. *Brain Injury, 27*(5), 620–631. https://doi.org/10.3109/02699052.2013.767936

Wilson, S. J., Parsons, K., & Reutens, D. C. (2006). Preserved singing in aphasia: A case study of the efficacy of Melodic Intonation Therapy. *Music Perception, 24*(1), 23–36.

Wilssens, I., Vandenborre, D., van Dun, K., Verhoeven, J., Visch-Brink, E., & Marien, P. (2015). Constraint-induced aphasia therapy versus intensive semantic treatment in fluent aphasia. *American Journal of Speech-Language Pathology, 24*(2), 281–294. https://doi.org/10.1044/2015_AJSLP-14-0018

Wright, H. H., Marshall, R. C., Wilson, K. B., & Page, J. L. (2008). Using a written cueing hierarchy to improve verbal naming in aphasia. *Aphasiology, 22*(5), 522–536. https://doi.org/10.1080/02687030701487905

Zumbansen, A., Peretz, I., & Hébert, S. (2014a). Melodic intonation therapy: back to basics for future research. *Frontiers in Neurology, 5*(7), 1–11. https://doi.org/10.3389/fneur.2014.00007/abstract

Zumbansen, A., Peretz, I., & Hébert, S. (2014b). The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca's aphasia. *Frontiers in Human Neuroscience, 8*(Article 592). https://doi.org/10.3389/fnhum.2014.00592/abstract

Chapter 3 – What is Meant by 'Multimodal Therapy' for Aphasia?

Preface

      For the systematic review presented in the previous chapter, the initial yield for the search term "multimodal" revealed that our understanding of multimodal treatment was one of many. This necessitated a definition, which we created for the review, and demonstrated that the term "multimodal" was insufficient and was not commonly understood. As a result, we decided to embark on a scoping review to elucidate any patterns in use of the term within aphasia research.

Publication
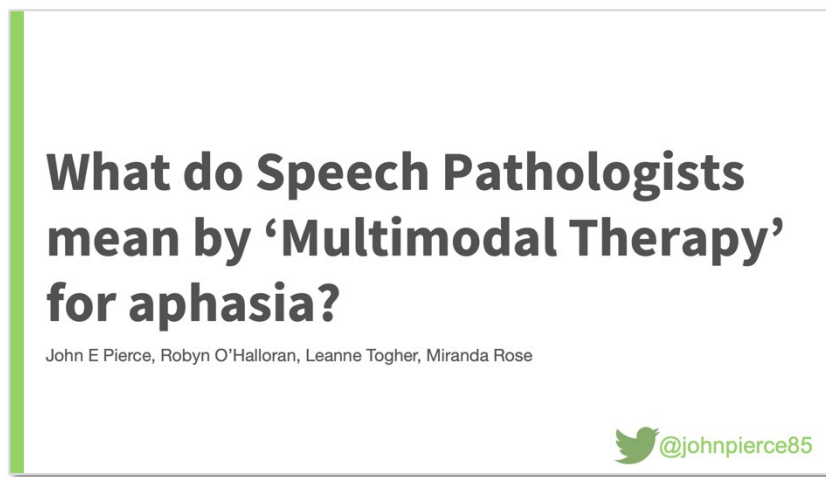
Video abstract – click to follow link

What is meant by 'multimodal therapy' for aphasia?

John E. Pierce[a,b], Robyn O'Halloran[a], Leanne Togher[c] and Miranda L. Rose[a]*

[a]*School of Allied Health, La Trobe University, Melbourne, Australia;* [b]*Speech Pathology, Cabrini Health, Melbourne, Australia;* [c]*Speech Pathology, Faculty of Health Sciences, The University of Sydney, Lidcombe, Australia*

John E Pierce. 494 Glen Huntly Road, Elsternwick 3185, Australia. ORCID identifier: 0000-0001-5164-5106 Twitter:@johnpierce85 pierce.john.e@gmail.com

Robyn O'Halloran. Health Sciences 1, La Trobe University, Plenty Road, Bundoora 3083, Australia. ORCID identifier: 0000-0002-2772-2164 R.OHalloran@latrobe.edu.au

Leanne Togher. Faculty of Health Sciences, The University of Sydney, 75 East St, Lidcombe 2141, Australia. ORCID identifier: 0000-0002-4518-6748

Leanne.togher@sydney.edu.au

*Miranda L. Rose. School of Allied Health, La Trobe University, Plenty Road, Bundoora 3083, Australia. ORCID identifier: 0000-0002-8892-0965 M.rose@latrobe.edu.au

Abstract

Purpose: "Multimodal therapy" is a frequent term in aphasia literature but it has no agreed

upon definition. Phrases such as *multimodal therapy* and *multimodal treatment* are applied to a

range of aphasia interventions as if mutually understood and yet the interventions reported in

the literature differ significantly in methodology, approach and aims. This inconsistency can be

problematic for researchers, policy makers and clinicians accessing the literature and

potentially compromises data synthesis and meta-analysis. A literature review was conducted

to examine what types of aphasia treatment are labelled multimodal and determine whether

any patterns are present.

Methods: A systematic search was conducted to identify literature pertaining to aphasia that

included the term *multimodal therapy* (and variants). Sources included literature databases,

dissertation databases, textbooks, professional association websites and Google Scholar.

Results: Thirty-three original research papers were identified, as well as another 31 sources

referring to multimodal research, all of which used a variant of the term 'multimodal therapy'.

Treatments had heterogeneous aims, underlying theories and methods. The rationale for using

more than one modality was not always clear, nor was the reason each therapy was considered

to be multimodal when similar treatments had not used the title. Treatments were noted to

differ across two key features. The first was whether the ultimate aim of intervention was to

improve total communication, as in Augmentative and Alternative Communication approaches,

or to improve one specific modality, as when gesture is used to improve word retrieval. The

second was the point in the treatment that the non-speech modalities were employed.

Discussion: Our review demonstrated that references to 'multimodal' treatments represent very different therapies with little consistency. We propose a framework to define and categorise 'multimodal' treatments which is based both on our results and on current terminology in speech language pathology.

Introduction

Broadly, a modality is a channel of communication, also described as a mode or medium (Crystal, 2011; Ferguson & Thomson, 2008), but understanding of the term varies subtly across fields. In education, the modality is the medium of teaching students; for example, lecture, seminar, self-directed learning (Armour, Schneid, & Brandl, 2016; Ilic & Maloney, 2014). In computing, modalities are different channels of communication between a device and its user, such as text, video, audio (Schroeder, 2010), or gestures/movement (Mäntyjärvi, Kela, Korpipää, & Kallio, 2004). Biology views modalities in terms of the primary senses used by the receiver, that is, auditory, visual, tactile, olfactory, or taste (Lawrence, 2011; Partan & Marler, 1999).

In speech language pathology, there is no formal definition of modality, but the term is commonly used to describe any method of communication between people. Modalities described in speech language pathology include:

- Speech/oral (Speech Pathology Australia, 2011)

- Graphic (Iacono, Mirenda, & Beukelman, 2009)

- Augmentative and alternative (Speech Pathology Australia, 2011)

- Gesture/manual (Rose, Raymer, Lanyon, & Attard, 2013b)

- Writing (Beeson & Egnor, 2006)

- Reading (Howard, Patterson, Franklin, Orchard-lisle, & Morton, 1985)

- Drawing (Purdy & Van Dyke, 2011)

- Music/melody (Pierce, Menahemi-Falkov, O'Halloran, Togher, & Rose, 2017)

- Facial expression (Iacono et al., 2009)

- Repetition[2] (Tanemura, 1999)

The above list is unlikely to be exhaustive or universally accepted by speech language pathologists. An explicit, finite set of modalities is rarely produced within a research field because it is assumed that modalities are "unproblematic and self-evident" (Bateman, 2011, p. 17). However, some observations can be made. The modalities listed above include linguistic communication modalities as well as non-linguistic modalities such as gesture, drawing, pictures and facial expression. By and large the modalities are employed in intentional communication, although in instances, some may be used without the intention to communicate (e.g., unconscious facial expression, drawing for pleasure, gesturing while talking on the telephone). Incidental and unconsciously produced messages such as body language are certainly recognised in speech language pathology but are rarely described as modalities, likely because intentional communication is most commonly the focus of speech language pathology practice. Interestingly, reading and writing are often described as independent modalities but are in fact the receptive and expressive components of the same modality, orthography.

In speech language pathology literature, *multi*modal typically refers to communication in any modality outside of speech, regardless of how many modalities are used (e.g., Speech Pathology Australia, 2011). Strictly speaking, the term multimodal refers to communication of the same message via more than one channel either

---

[2] Repetition is not a communication channel in itself but rather a way of eliciting speech, frequently used in language interventions. Nevertheless, many papers describe repetition as a modality (e.g., Howard et al., 1985; Kiran, 2005; Tanemura, 1999)

simultaneously or serially (Partan & Marler, 2005). Thus, using writing alone to

communicate, for example, could be regarded as unimodal (Partan & Marler, 2005).

However, this distinction is rarely made in speech language pathology. Even in the field

of Augmentative and Alternative Communication, multimodal communication typically

refers to communication using multiple *non-speech* systems, whereas a single system

combined with speech is considered unimodal; for example, speech plus signs (Iacono

et al., 2009; Sigafoos et al., 2007).

With limited agreement on the terms modality and multimodal, what then, do

speech language pathologists mean when describing multimodal *treatment*? Aphasia is

a multimodal disorder in that it affects communication across multiple communicative

systems (Hallowell & Chapey, 2001), so use of the term to describe treatment is

common in aphasia research. However, despite the frequency, there is little consistency

in its use. The label of *multimodal* is applied to a diverse range of interventions and yet

used in the speech language pathology literature as if understood mutually by all. For

example, the following article titles refer to very different treatments:

1. *An investigation of the communicative use of trained symbols following*

   multimodality training (Purdy, Duffy, & Coelho, 1994) - therapy to improve

   conversational use of gesture, speech and a picture board through practise of

   each.

2. ***Multimodal therapy*** *of word retrieval disorder due to phonological encoding*

   dysfunction (Weill-Chounlamountry, Capelle, Tessier, & Pradat-Diehl, 2013) -

   treatment of the phonological output lexicon through computerised

   phonological/orthographic tasks.

3. *Comparing uni-modal and **multi-modal therapies** for improving writing in acquired dysgraphia after stroke* (Thiel, Sage, & Conroy, 2015) – intervention aimed at improving spelling accuracy through semantic, phonological and orthographic distractor decisions.

As can be seen with only three examples, there is little agreement. Some published multimodal treatments target improvement of spoken output, some spelling, and some both verbal and non-verbal communication. The most straightforward explanation is to consider multimodal as merely an adjective; describing any treatment that uses more than one modality. However, this application of the term would include the majority of treatments for aphasia. Repetition and responsive naming treatment tasks use only a single modality (speech) for both input and response but most other treatments use one or more modalities as input and expect one or more modalities as patient responses. Thomson (2012) reviewed 453 anomia treatment instances and found that only 21 (4.6%) used a single modality for input, all pictures. Of these 21 treatment instances, all used speech for participant output. This demonstrates that all anomia treatments might be called multimodal if the term is taken to mean simply the use of 'more than one modality' within a task. This definition would therefore be so broad as to be useless.

There are consequences of ambiguous terminology. Use of specific definitions in science progresses theory, research and practice, in contrast to vague or convention-based terminology (McNeil & Pratt, 2001; Schindler, 2009; Walsh, 2009), which creates a "breakdown in communication and the exchange of ideas" (Walsh, 2009, p. 67). As one example, the use of one term to denote multiple, dissimilar treatments presents difficulty in summarising and comparing treatments. In an early meta-analysis of

aphasia treatment effectiveness, Robey (1998) had to develop a classification system in order to group and distinguish treatment approaches based on their underlying methodology, though papers did not always fall easily into these categories. Specific, defined labels for approaches or underlying theories, if not for the treatments themselves, would facilitate grouping for synthesis.

Patients and policy makers would also benefit from clear terminology (Madden, Robinson, & Kendall, 2017). Tracking outcomes is easier when treatments are distinguishable without inside knowledge of theories and approaches. With increasing value placed on person-centred healthcare (Wallace et al., 2016), clients should have a greater voice in treatment decisions but cannot be expected to disentangle various interpretations of multimodal approaches.

Differentiation between multimodal treatments for aphasia is often possible by reading the introduction and methods sections of papers as they may outline the underlying theoretical rationale and therapy specifics. However, finding time to read research is the most consistently reported barrier to implementing evidence-based practice for clinicians (O'Connor & Pettigrew, 2009) and thus, they are less likely to read papers in detail. Clinicians and other consumers of research may therefore impute their own understanding of the treatment being used based on the term 'multimodal therapy' in the title or abstract, or on websites and in textbooks, and make erroneous conclusions about treatment effectiveness and applicability.

In order to gain greater clarity in the use of terminology, we aimed to investigate and map the various interpretations of multimodal treatment within the aphasia literature through a scoping review. The question examined was, "What types of

aphasia therapy are labelled as multimodal?" Original research was sought as well as secondary sources; that is, literature referencing or discussing multimodal treatment.

Method

A systematic search was conducted by the first author between September and October 2017 to identify English language literature pertaining to people with aphasia that included the term *multimodal therapy* (and variants). Grey literature was also included in order to build a comprehensive picture of current use of the term(s). Primary progressive aphasia was included. Databases searched were Medline (Ovid, 1946 - Sep 2017), CINAHL, PsycINFO (Ovid, 1806 - 2017) and Proquest Dissertations and Theses. Google Scholar was also searched. Other sources included aphasia textbooks and speech language pathology association websites. There was no limit on the publication date.

Search strategies

Databases and dissertations

For databases (Medline, CINAHL, PsycINFO and Proquest Dissertations and Theses), the two search concepts *aphasia* and *multimodal* were used to generate the search operators as listed below.

aphasia (MeSH term) OR aphasia OR dysphasia OR anomia

AND

multimodal* OR multi-modal*

No limiters were applied outside of the search terms. Results from databases were imported into citation management software before screening.

Google Scholar

A search was conducted in Google Scholar using the terms "multimodal" and "aphasia" with the operator "allintitle:" to ensure both terms appeared in the article's title. This search strategy yields more grey literature results (Haddaway, Collins, Coughlin, & Kirk, 2015). All results were screened. The two terms were also searched without the "allintitle:" operator and the first 300 results were screened (based on title and preview), as recommended in Haddaway, et al. (2015).

Textbooks and association websites

Full text searches for the words 'multi-modal' or 'multimodal' were conducted in 28 e-books relating to aphasia. Indexes of five aphasia textbooks were searched for the term 'multimodal' and corresponding pages were screened manually. Multiple speech language pathology association websites were searched using Google: American Speech-Language-Hearing Association, Speech Pathology Australia, Royal College of Speech and Language Therapists, South African Speech-Language-Hearing Association, Irish Association of Speech & Language Therapists, New Zealand Speech Therapists' Association, and Speech-Language & Audiology Canada. The term 'multimodal' was searched in combination with an operator that limited results to each website; e.g., *site:asha.org multimodal*. Google-generated previews and titles were used for screening.

Study selection

Initial results were screened by the first author according to the following inclusion criteria:

i.   English language,

ii.   uses the term multimodal in reference to treatment; for example,

multimodal + program / strategy / approach / treatment / therapy /

cueing,

iii.   the participants had aphasia or the topic was aphasia.

Data Extraction

Papers were categorised by the first author into publication type (article, dissertation, conference abstract, etc.) and then into original research or secondary sources. To examine term usage, the key phrase(s) containing *multimodal* were extracted from each paper, typically located within the title or abstract.

Within original research, the target of intervention, the underlying rationale and outcome measures of treatment were recorded. The modalities used in papers were also extracted. This data was not sufficient to describe and classify how the term multimodal was being used in aphasia intervention research. Therefore, treatment designs were further divided into three elements of a) input, b) therapist cueing and c) participant output/response and the modalities used for each element were recorded. We also examined the timing of modalities, for example, used for each target production or only on errors. This complement of data sufficiently discerned the most useful dimensions with which to categorise treatments.

Results

Figure 3-1 – Diagram of search procedures
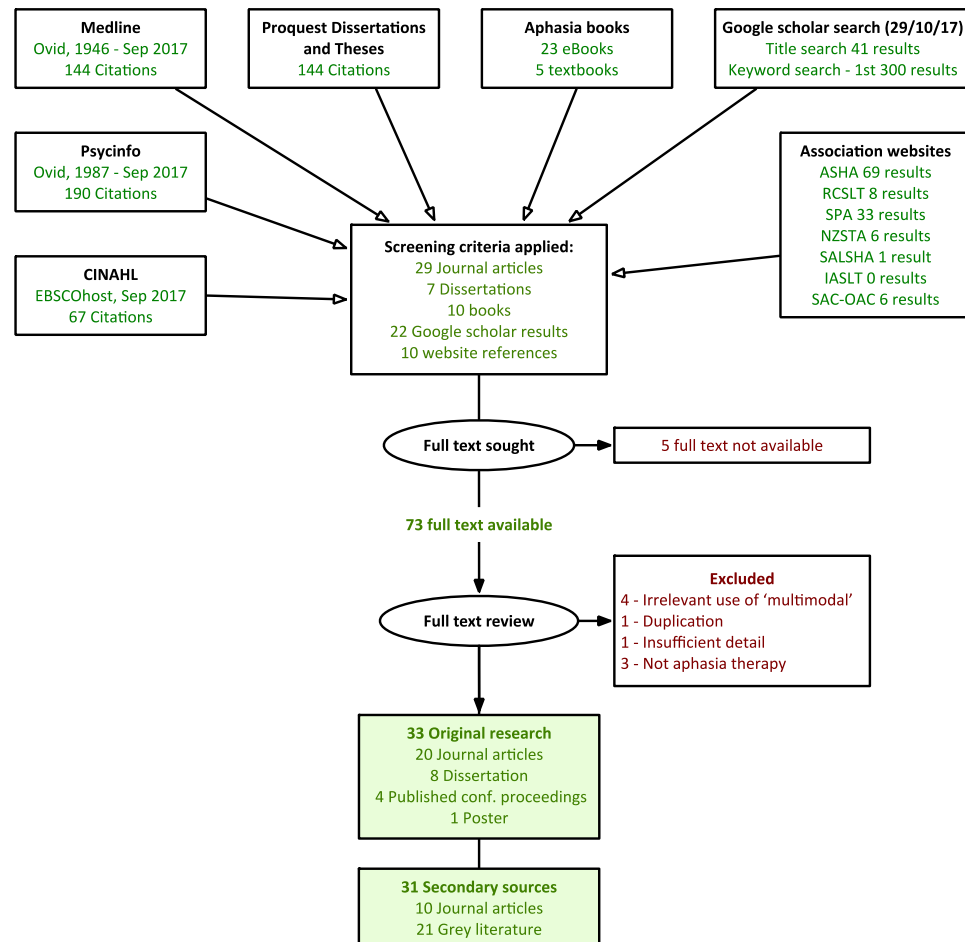
Figure 3-1displays the search results. Initial yields for databases were Medline (144), PsycINFO (190), CINAHL (67), Proquest (144). After screening titles and abstracts, a total of 29 journal articles and seven dissertations were included. The Google Scholar search yielded 41 results for the title search. After screening these and the first 300 results of the keyword Google Scholar search there were 22 results. There were ten

references to multimodal treatments in books. Association website searches found 123

results and ten met inclusion criteria.

The yield was thus 73 references which were further examined, with 9 excluded

due to: irrelevant use of 'multimodal' (4), duplication (1), insufficient information (1), or

unrelated to aphasia therapy (3). Finally, there were 33 original research papers and 31

secondary sources.

Secondary Sources

Secondary sources are found in Supplementary Appendix 2A. There were 10

narrative or systematic reviews and 21 grey literature items including book chapters,

posters, conference proceedings and web pages.

Original research

Of the 33 papers with original data and sufficient detail to examine methods,

there were 20 journal articles, eight dissertations, four published conference

proceedings and one poster. Only one paper had a participant with primary progressive

aphasia (Rebstock, 2014); participants in all other papers had acquired aphasia.

Supplementary Appendix 2B contains data extraction results.

A wide variety of phrases containing multimodal were used. Some of the phrases

were formal titles of established therapies or approaches such as Multimodal

Communication Program (4 papers), Multi-Modality Aphasia Therapy (3) or Music and

Multimodal Stimulation (1) whereas the majority appeared to be ad hoc descriptions of

the treatment without a working definition.

Extracting data for whether modalities were used as input, cueing or output was necessarily challenging, given the lack of agreement of what constitutes a modality. For example, when a participant is asked to name a picture using an AAC picture board, they point (gesture) at a symbol (visual) which may include text (written). An electronic device may also have text to speech capability which could be considered spoken output. We tried to remain descriptive for this process which resulted in a large number of modalities. However, some modalities were seen frequently across papers.

A) Input

Input was considered as any stimulus or material presented to a participant in order to elicit a response. Most commonly, input was pictures alone (19) or in combination with other modalities (10). Speech (11) and written (10) input were used in approximately one third of studies, though most often in conjunction with other modalities. One study used pictures, objects, videos and sensory information (hot/cold) to stimulate increased comprehension (Henning, 2016), and another used pictures, objects and spoken and written words as naming stimuli (Denman, 2017). Four did not report on inputs used, though two were case reports rather than prospective interventions (Beeson & Ramage, 2000; Lasker, LaPointe, & Kodras, 2005)

B) Therapist cueing

Therapist cueing covered anything the therapist (or software) provided in addition to input: modelling, shaping, correction or cueing. Clinician cueing modalities were more widely distributed than input: drawing (12), speech (26), written (12), gesture (17), symbol/picture boards including software (10) and one each of melodic speech and objects. However, five did not clearly report on cueing.

C) Participant output

All studies without exception required participants to produce some form of spoken output, whether naming, sentence production, repetition or phonemes/syllables. Spoken output was not always the *target* of intervention, but all treatments requested or allowed it. Gesture (20), writing (18), drawing (14), and symbol/picture boards (10) were also frequently reported as outputs.

During extraction of data (details in Supplementary Appendix 2B), the primary aim of interventions emerged as a key dimension which differentiated the treatments within two broad categories. Another dimension was the timing of multimodal involvement, which formed a number of subcategories within each treatment aim category. The section below describes these dimensions.

1. Ultimate aim of improving total communication

In multimodal papers including what we termed *total communication approaches*, the emphasis was on successful communication of the message in any modality rather than improvement of a particular modality. Fifteen such papers were found, designed to teach participants to use non-speech modalities such as gesture or picture boards to communicate. These modalities were employed either for augmentation of remaining speech or as an alternative channel to speech. As an example of augmentation, Carlomagno et al. (2013) trained two people with aphasia in "multimodal communication therapy" to implement gesture *alongside* speech to add semantic information for the listener. In contrast, Purdy, Duffy and Coelho (1994) used "multimodality training" for alternative communication wherein people with aphasia were trained in the use of three different modalities – speech, gesture production and pointing to pictures on a communication board. Gestures and use of the communication

board were trained independently of speech with the aim of providing an alternative communication channel for participants to switch to if speech failed. Many of these total communication approaches were based on Promoting Aphasic Communicative Effectiveness (PACE, Davis, 2005), a treatment for aphasia where the participant "is allowed free choice with respect to selection of communicative channels" (Pulvermüller & Roth, 1991, p. 40). Interestingly, PACE is not referred to as multimodal by its original authors (Davis, 2005).

Timing of modalities

Within total communication approaches, the timing of modality training was a key aspect that differed across papers, specifically, whether modalities were trained simultaneously (e.g., producing gesture and speech within a sentence), separately (e.g., treating drawing in one session and writing in another) or consecutively (e.g., spoken naming, then written naming, repetition and symbol pointing for the same target word in one session). We found that, of the fifteen papers aiming to improve total communication:

One trained modalities simultaneously

Two trained modalities separately

One trained modalities separately before combining them

Nine trained multiple modalities consecutively for each target

Two allowed the participant to choose the modality or modalities used for each target

2. Ultimate aim of improving speech (or, another specific modality)

In the second category of treatment aims, alternative modalities were used explicitly as a means to improve spoken output. There were 17 such studies. The theoretical premise was not always stated explicitly but papers predominantly invoked the principle that other, less impaired modalities have sufficient neural links with damaged linguistic representations to aid their retrieval and production. Luria was an early proponent of facilitation across modalities in what he described as *intersystemic reorganisation* (Luria, 1970; Pierce et al., 2017). For example, there is research demonstrating that gesture is used by both aphasic and non-aphasic speakers which may assist word retrieval (Rose, Attard, Mok, Lanyon, & Foster, 2013a).

Papers within this category did not all cite the same intersystemic links as rationales for their designs. Some papers proposed intermodal links at the semantic-conceptual level to promote activation of the spoken modality (e.g., Dunn, 2010; McCarthy, 2004), while some suggested multiple links, including semantic, orthographic and phonological (e.g., Brookshire, Conway, Pompon, Oelke, & Kendall, 2014; Thiel et al., 2015). Other authors cited theoretical and empirical support for links between more specific systems: gesture and verbal lexical retrieval (Rose & Sussmilch, 2008), phonemes and graphemes (Weill-Chounlamountry et al., 2013) and language-action links (Grechuta et al., 2016).

In addition to the 17 aimed at improving speech, three papers used the same principal of facilitation across modalities but targeted impairments of non-speech modalities. Thiel, Sage and Conroy (2015) targeted written output using written and spoken matching and copying/repetition. In Brookshire et al. (2014), the aim was improved reading comprehension via improved phonological processing, while Henning

(2016) targeted both spoken output and auditory comprehension within their program based on Melodic Intonation Therapy (MIT, Sparks, Helm & Albert, 1974).

It is important to note that the two categories of intervention aims described were not mutually exclusive. A small number of papers explicitly stated aims in both categories of total communication and improving speech. Attard, Rose and Lanyon (Attard, Rose, & Lanyon, 2013) had a primary aim of improved word retrieval but noted that the M-MAT protocol could also support enhanced alternative communication. Rebstock (2014) investigated outcomes of both word retrieval and modality switching.

Although all studies with the aim of facilitation across modalities were based on the same underlying theory of links between modalities, there were significant differences in the way this was interpreted in treatment. The timing and number of modalities used in each element of input, therapist cueing and participant output were examined. This revealed three groupings.

2.1 Multimodal cueing and output

Input	Therapist cueing	Participant output
Unimodal	Multimodal	Multimodal

One method of facilitating speech had the participant receiving one modality as input and producing several modalities as an *output* and this typically involved cueing or modelling from the therapist (or software, in some studies). For example, in M-MAT, if the participant was unable to name the picture stimulus, they produced written, gestural and drawn representations of the target while repeating a spoken model from the therapist (Attard et al., 2013). Modelling was provided by the therapist in all modalities as needed. Seven results fitted this category.

Four of these seven multimodal cueing + output studies provided cueing only on errors (Attard et al., 2013; Rose & Sussmilch, 2008; Rose et al., 2013a; Rose, Mok, Carragher, Katthagen, & Attard, 2015), whereas three routinely provided multimodal cueing for each item presented (Hoodin & Thompson, 1983; Kearns, Simmons, & Sisterhen, 1982; Rebstock, 2014).

2.2 Multimodal input

Input	Therapist cueing	Participant output
Multimodal	Unimodal	Unimodal

The complement to multimodal cueing and output was the use of non-speech modalities as *input* or stimulation for participants. In other words, the clinician presented the participant with multiple modalities to elicit a response in a single modality. There were three studies in this subcategory (Denman, 2017; Henning, 2016; Thompson & McReynolds, 1986) which used a selection of objects, pictures, written words, videos and sensory/tactile cues as input.

Thompson & McReynolds (1986) based their design on the "stimulation approach," attributed to Schuell and Wepman (Robey, 1998). The underlying assumption of the stimulation approach is that the person with aphasia has not lost language but only *access* to the language, and that sufficient activation from the environment in multiple modalities can enhance access to target words (Duffy & Coelho, 2001). Robey's literature review (1998) grouped treatments in this approach under the banner *Schuell-Wepman-Darley Multimodality treatment* or *multimodal stimulation* and this label is used in other textbooks and reviews. Denman (2017) was presented as a poster and thus did not provide a comprehensive rationale but appeared

to rely on the same multimodal stimulation approach. Lastly, the treatment in Henning (2016) was based on Melodic Intonation Therapy with the addition of multimodal stimuli, which were employed to improve comprehension of word meaning and thus improve output. There was no clear explanation of the assumed mechanism behind this simulation.

2.3 Multiple multimodal tasks

Task 1			Task 2			Task n
Input	Therapist cueing	Participant output	Input	Therapist cueing	Participant output	etc.
Unimodal or Multimodal	Unimodal or Multimodal	Unimodal or Multimodal	Unimodal or Multimodal	Unimodal or Multimodal	Unimodal or Multimodal	etc.

A third approach to implementing facilitation between modalities was use of multiple modalities *across therapy tasks* for each target. Our data extraction found six such studies. Studies using this approach did not necessarily use multiple modalities for each task, but still term themselves "multimodal" because the following task employed a different modality. As one example, Thomson (2012) treated word retrieval with a series of tasks completed for each noun. These tasks included naming, letter scrambles, reading aloud, repetition, semantic feature analysis, and written/spoken word matching. Weill-Chounlamountry et al. (2013) described their therapy software "Au Fil de Mots" as multimodal. Au Fil de Mots has several structured steps in which participants complete word scrambles, repeat phonemes and words, and type the target.

Miscellaneous

Two studies did not fit within the classification system described above but employed the use of "multimodal cueing". Dunn (2010) investigated the addition of therapist-produced gesture to semantic and phonological cueing in picture naming, but cues were to aid word retrieval and the participant was not required to imitate these. Thus, stimuli and output were unimodal while therapist cueing was multimodal.

Fink et al. (2005) reported on multimodal cueing and "multimodal exercise" on computer therapy software. The software allows clinicians to select the modalities for the stimuli, choices and cueing for matching tasks. Describing each task as multimodal is technically correct, as the input, cueing and responses allow different modalities. However, as described earlier, nearly all anomia treatments use more than one modality between input and output.

Discussion

This review has demonstrated that the term 'multimodal treatment' and similar iterations represent very different therapies with little consistency. First, it is not clear from the term what the purpose of intervention is. Some aimed to improve total communication through any modality, while others aimed to use intact modalities to facilitate a damaged modality (most often, speech). Second, the component of the treatment task using multiple modalities was not consistent. Some presented the participant with multiple simultaneous stimuli as input, while others had participants producing multiple modalities in response to a single stimulus. Still others included multiple, discrete tasks which each used variety of modalities. Further differences were seen in whether participants were given clinician cueing or shaping only when errors

were made, or routinely for each target. Finally, there was variation in whether those studies designed to improve total communication presented modalities simultaneously, separately or sequentially. There are many other papers with the same design and principles as those found in this review which could equally be termed multimodal treatments but are not labelled in this way by the authors. This is evident with the PACE approach. Multiple studies in our review were investigations of PACE, yet there are a multitude of PACE studies which do not use the label *multimodal*, including the original paper. Agreement between researchers on the term is clearly lacking.

Consequently, aphasia researchers need to be cautious about describing treatments as multimodal as if its meaning is evident, particularly in article titles and treatment descriptions. With a few exceptions (M-MAT, M-STIM, MCP, multimodal stimulation), *multimodal* does not have a set meaning as a treatment approach and thus describing therapy as multimodal brings no clarity to the reader.

There is also a need to clearly describe the various dimensions of treatment as outlined in this article – the aim of intervention, the rationale, the timing of modalities and who (therapist, client) produced them, and which modalities were used. In most cases we were able to discern the treatment by reading the papers in full, but for some, the aim of intervention or the theoretical principle was not clearly stated and thus needed to be inferred. Ambiguous terminology or labelling makes the use of clear therapy descriptions more important in order to differentiate between approaches. Use of reporting guidelines such as the TIDieR checklist (Hoffman et al., 2014) assists in clarifying most components of treatment, including the rationale (Item 2), materials (Item 3) and procedures, including prompts and cueing (Item 4).

It is surprising that this review found only 33 results (20 papers, 8 dissertations, 4 conference proceedings, 1 poster) with original intervention research explicitly described by the authors as multimodal. We expected many more results, considering the widespread use of the term in speech language pathology clinical practice – the search for secondary sources resulted in 31 references to multimodal treatment in aphasia, and there are likely to be many more in textbooks, non-association websites and journal articles without indexed full-text. The review was also limited to English literature and it is not known if similar problems with defining multimodal treatments exist in other languages.

Another possible weakness of this review is that screening of articles against inclusion criteria was conducted by a single author only. Best practice for scoping reviews does call for two authors to screen search results. However, as the criteria relied on identifying the language of the article, the presence of the keywords 'multimodal therapy' and confirming the topic of aphasia, there was minimal subjective decision making and it was felt that one author was sufficient.

The complexity of human communication is an acknowledged challenge in developing accurate terminology for speech language pathology (Walsh, 2009), so further classification is necessarily difficult. Nonetheless, it is imperative that we clarify as best as possible what speech-language pathologists and aphasia researchers mean when discussing multimodal treatment. Ideally, consensus is needed on a definition of multimodal therapy which is specific and represents the supporting theories while also providing inclusion and exclusion criteria (McNeil & Pratt, 2001). As we have demonstrated, current use of the term not only represents multiple theories but

provides no inclusion or exclusion criteria and the majority of aphasia treatments could arguably be defined as multimodal.

Proposed framework

We now propose a broad framework to categorise multimodal treatments. This framework is based on the themes identified in this review but also incorporates one very common interpretation of the term multimodal among speech language pathologists, which is that multimodal refers to any non-verbal production (SPA, 2011). 'Verbal' is used here in the sense of being word-based, consisting of speech and/or orthography (Crystal, 2011). The modalities of speech and orthography (whether reading or writing) are therefore not considered to be multimodal in this definition while modalities such as gesture, drawing, singing/rhythm, symbol boards, etc. are included. This conflicts with some previous perspectives but the distinction is necessary to provide a definition of multimodal that does not encompass the majority of aphasia treatment. For the same purpose, the use of images for stimuli or cueing does not necessarily qualify as multimodal treatment. A treatment requiring confrontation naming of images and including orthographic cueing, for example, is excluded despite the use of images and reading. While this review has demonstrated that some authors would consider such a treatment to be multimodal, in general, speech language pathology literature classifies this as traditional therapy.

Figure 3-2 illustrates the proposed framework, which includes two categories within multimodal treatment and one verbal treatment category. Multimodal treatments include two subcategories. The key differentiating features within the model

are whether non-verbal modalities are employed, and the primary aim of the

intervention – either to improve verbal communication (speech/orthographic) or to

improve speech and/or one or more nonverbal modalities.

Figure 3-2 – Proposed framework for classifying multimodal treatments.
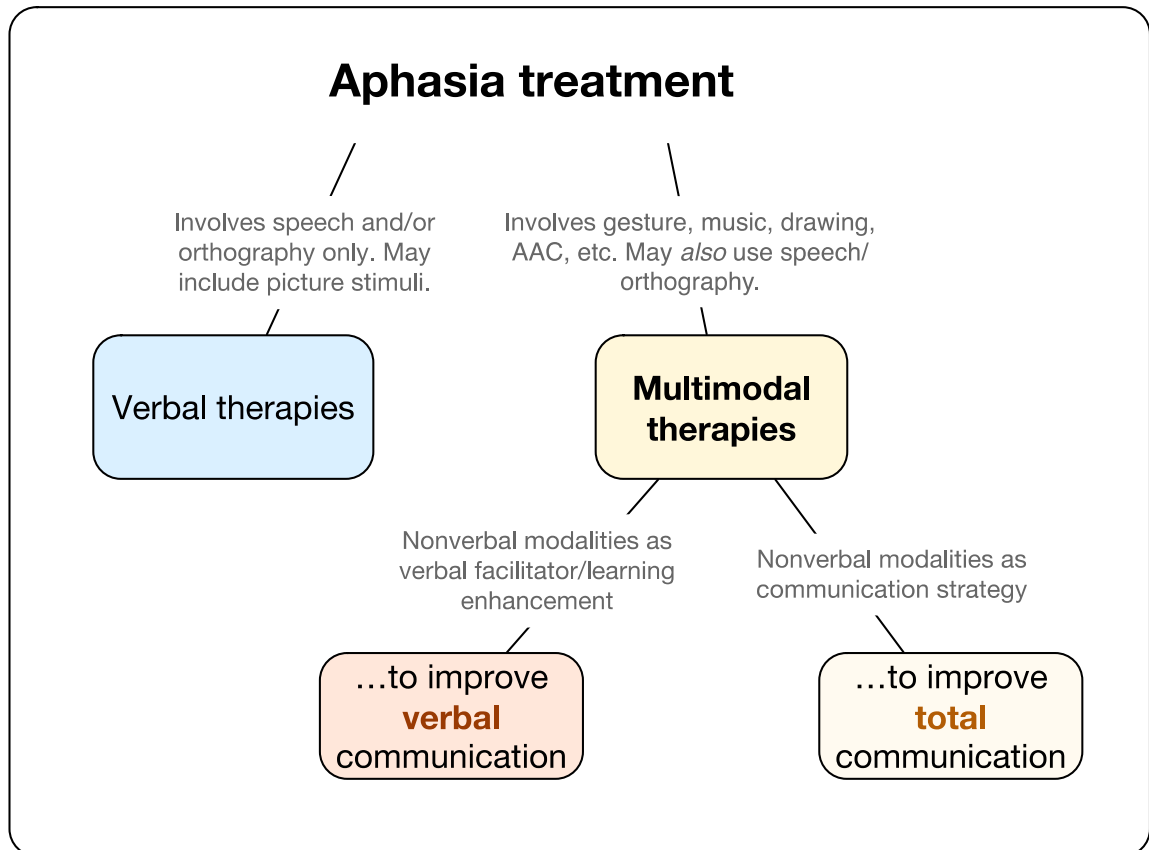
Two key features of treatments are examined and produce two categories: 1.
Multimodal therapy to improve *verbal* communication, 2. *Multimodal therapy* to
improve *total* communication.
Verbal (speech or reading/writing) therapy to improve *verbal* communication is not
considered multimodal.

1. Multimodal therapy …to improve <u>*verbal*</u> communication

As described earlier, treatments fitting this category use at least one non-verbal

modality (e.g., gesture) to facilitate improvement of verbal abilities. Commonly, the

target is speech, but writing, auditory comprehension and reading can also be targeted.

As the goal is verbal communication, treatments in this category typically combine nonverbal and verbal modalities. Importantly, the use of nonverbal modalities is a means to an end and not the primary goal. Examples include MIT (using melody to enhance speech production) (Sparks et al., 1974) and M-MAT (using gesture, drawing, writing, reading and verbal repetition to improve word retrieval and sentence production) (Rose & Attard, 2011).

2. Multimodal therapy …to improve _total_ communication

Treatments in this category target nonverbal modalities as communicative actions in their own right, rather than a means to an end. By definition, AAC approaches to aphasia fit into this category. Such treatments may or may not combine verbal modalities with the non-verbal modalities.

Modalities may not only be trained as stand-alone communication channels, but also to augment remaining speech or writing. This remains different to category 1 in that, in category 1 (facilitation/learning), nonverbal modalities are primarily intended to assist the person with aphasia to access/learn speech or writing and not to communicate with the conversation partner. In contrast, in category 2 (total communication) nonverbal modalities are a vital part of message transfer.

This proposed framework presents categories which incorporate the findings of our review with existing interpretations. We believe speech language pathologists and researchers will find the framework intuitive, as the aim of an intervention is typically readily identifiable. Two key questions can be asked about each therapy to categorise it – Does this treatment employ nonverbal modalities such as gesture, music or drawing,

and if so, does it aim to improve verbal communication or improve total communication of a message?

Our framework was recently utilised in the categorisation of speech and language therapy interventions for aphasia within REhabiliation and recovery of peopLE with Aphasia after StrokE (RELEASE). This big data project aims to synthesize individual participant data from multiple primary research studies (Collaboration of Aphasia Trialists, n.d.). The availability of the multimodal definition facilitated synthesis of highly complex therapy interventions and in turn meta-analysis. The framework was quick to apply and complemented other more commonly applied definitions of therapy approaches such as CIAT or MIT.

Naturally, discussion and consensus from the field is required for such a framework to be adopted and this proposal might be further developed to capture details identified in this review such as the timing of modalities or their use in input/cueing/output. Nonetheless, we suggest that it broadly captures the array of 'multimodal treatments' currently found within aphasia while giving clarity to their differing approaches and intended outcomes. A more coherent picture of such treatments benefits patients and stakeholders and may allow more precise reviews and meta-analyses in the future.

References

Armour, C., Schneid, S. D., & Brandl, K. (2016). Writing on the board as students' preferred teaching modality in a physiology course. *Advances in Physiology Education*, *40*(2), 229–233. https://doi.org/10.1152/advan.00130.2015

Attard, M. C., Rose, M. L., & Lanyon, L. E. (2013). The comparative effects of Multi-Modality Aphasia Therapy and Constraint-Induced Aphasia Therapy-Plus for severe chronic Broca's aphasia: An in-depth pilot study. *Aphasiology*, *27*(1), 80–111. https://doi.org/10.1080/02687038.2012.725242

Bateman, J. A. (2011). The Decomposability of Semiotic Modes. In K. O'Halloran & B. A. Smith (Eds.), *Multimodal Studies - exploring issues and domains* (pp. 17–38). New York & London: Routledge.

Beeson, P. M., & Egnor, H. (2006). Combining treatment for written and spoken naming. *Journal of the International Neuropsychological Society*, *12*, 816–827. https://doi.org/10.10170S1355617706061005

Beeson, P. M., & Ramage, A. E. (2000). Drawing from experience: the development of alternative communication strategies. *Topics in Stroke Rehabilitation*, *7*(2), 10–20.

Brookshire, C. E., Conway, T., Pompon, R. H., Oelke, M., & Kendall, D. L. (2014). Effects of intensive phonomotor treatment on reading in eight individuals with aphasia and phonological alexia. *American Journal of Speech-Language Pathology*, *23*(2), S300–11. https://doi.org/10.1044/2014_AJSLP-13-0083

Carlomagno, S., Zulian, N., Razzano, C., De Mercurio, I., & Marini, A. (2013). Coverbal gestures in the recovery from severe fluent aphasia: A pilot study. *Journal of Communication Disorders*, *46*(1), 84–99. https://doi.org/10.1016/j.jcomdis.2012.08.007

Collaboration of Aphasia Trialists. (n.d.). *REhabilitation and recovery of peopLE with Aphasia after StrokE (RELEASE).* http://www.aphasiatrials.org/release

Crystal, D. (2011). Dictionary of Linguistics and Phonetics. John Wiley & Sons.

Davis, G. A. (2005). PACE revisited. *Aphasiology*, *19*(1), 21–38. https://doi.org/10.1080/02687030444000598

Denman, A. (2017, September). Multi-modal errorless learning functional naming therapy - a single case study [poster]. *RCSLT Conference*. Glasgow, Scotland.

Duffy, J. R., & Coelho, C. A. (2001). Schuell's stimulation approach to rehabilitation. In R. Chapey (Ed.), *Language interventions strategies in aphasia and related neurogenic communication disorders* (4 ed., pp. 341–379). Baltimore: Lipincott Williams & Wilkins.

Dunn, I. (2010). *The effects of multimodality cueing on lexical retrieval in aphasic speakers (Doctoral Thesis)*. The William Paterson University of New Jersey.

Ferguson, A., & Thomson, J. (2008). Systemic Functional Linguistics and Communication Impairment. In *The handbook of clinical linguistics* (pp. 130–143). Cambridge: Blackwell.

Fink, R., Brecher, A., Sobel, P., & Schwartz, M. (2005). Computer-assisted treatment of word retrieval deficits in aphasia. *Aphasiology*, *19*(10-11), 943–954.

Grechuta, K., Rubio, B., Duff, A., Oller, E. D., Pulvermüller, F., & Verschure, P. (2016). Intensive language-action therapy in virtual reality for a rehabilitation gaming system (Vol. 9, pp. 243–254). Presented at the 10th International Conference of Disability, Virtual Reality & Associated Technologies.

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PloS One*, *10*(9), e0138237. https://doi.org/10.1371/journal.pone.0138237

Hallowell, B., & Chapey, R. (2001). Introduction to language intervention strategies in adult aphasia. In R. Chapey (Ed.), *Language interventions strategies in aphasia and related neurogenic communication disorders* (4 ed., pp. 3–19). Baltimore, MD: Lipincott Williams & Wilkins.

Henning, D. M. (2016). *Music and multimodal stimulation (M-STIM): A dynamic approach to increasing expressive and receptive language in severe global aphasia (Masters thesis)*. Northern Illinois University.

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., … Lamb, S. E. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, *348*(mar07 3), g1687–g1687. https://doi.org/10.1136/bmj.g1687

Hoodin, R. B., & Thompson, C. K. (1983). Facilitation of verbal labeling in adult aphasia by gestural, verbal, or verbal plus gestural training (pp. 62–64). Presented at the Clinical Aphasiology Conference, Phoenix, AZ.

Howard, D., Patterson, K., Franklin, S., Orchard-lisle, V., & Morton, J. (1985). The facilitation of picture naming in aphasia. *Cognitive Neuropsychology*, *2*(1), 49–80. https://doi.org/10.1080/02643298508252861

Iacono, T., Mirenda, P., & Beukelman, D. (2009). Comparison of unimodal and multimodal AAC techniques for children with intellectual disabilities. *Augmentative and Alternative Communication*, *9*(2), 83–94. https://doi.org/10.1080/07434619312331276471

Ilic, D., & Maloney, S. (2014). Methods of teaching medical trainees evidence-based medicine: a systematic review. *Medical Education*, *48*(2), 124–135. https://doi.org/10.1111/medu.12288

Kearns, K. P., Simmons, N., & Sisterhen, C. (1982). Gestural sign (Amer-Ind) as a facilitator of verbalization in patients with aphasia (pp. 183–191). Presented at the Clinical Aphasiology Conference, Oshkosh, WI. Retrieved from http://eprints-prod-05.library.pitt.edu/725/1/12-23.pdf

Kiran, S. (2005). Training phoneme to grapheme conversion for patients with written and oral production deficits: A model-based approach. *Aphasiology*, *19*(1), 53–76. https://doi.org/10.1080/02687030444000633

Lasker, J., LaPointe, L., & Kodras, J. (2005). Helping a professor with aphasia resume teaching through multimodal approaches. *Aphasiology*, *19*(3-5), 399–410.

Lawrence, E. (2011). Henderson's Dictionary of Biology. Pearson Higher Ed. Retrieved from http://www.myilibrary.com?ID=311493

Luria, A. R. (1970). Traumatic aphasia: Its syndromes, psychology and treatment. (D. Bowden, Trans.). The Hague, Netherlands: Mouton & Co.

Madden, E. B., Robinson, R. M., & Kendall, D. L. (2017). Phonological Treatment Approaches for Spoken Word Production in Aphasia. *Seminars in Speech and Language*, *38*(1), 62–74. https://doi.org/10.1055/s-0036-1597258

Mäntyjärvi, J., Kela, J., Korpipää, P., & Kallio, S. (2004). Enabling fast and effortless customisation in accelerometer based gesture interaction (pp. 25–31). Presented at the 3rd international conference on Mobile and ubiquitous multimedia, College Park, MD: ACM Press. https://doi.org/10.1145/1052380.1052385

McCarthy, S. E. (2004). *The effects of a multimodality approach on sentence production using response elaboration training with a reading component on aphasic patients (Doctoral dissertation)*. East Tennessee State University.

McNeil, M. R., & Pratt, S. R. (2001). Defining aphasia: Some theoretical and clinical implications of operating from a formal definition. *Aphasiology*, *15*(10-11), 901–911. https://doi.org/10.1080/02687040143000276

O'Connor, S., & Pettigrew, C. M. (2009). The barriers perceived to prevent the successful implementation of evidence-based practice by speech and language therapists. *International Journal of Language & Communication Disorders*, *44*(6), 1018–1035. https://doi.org/10.1080/13682820802585967

Partan, S. R., & Marler, P. (2005). Issues in the Classification of Multimodal Communication Signals. *The American Naturalist*, *166*(2), 231–245. https://doi.org/10.1086/431246

Partan, S., & Marler, P. (1999). Communication Goes Multimodal. *Science*, *283*(5406), 1272–1273. https://doi.org/10.1126/science.283.5406.1272

Pierce, J. E., Menahemi-Falkov, M., O'Halloran, R., Togher, L., & Rose, M. L. (2017). Constraint and multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation*, 1005–1041. https://doi.org/10.1080/09602011.2017.1365730

Pulvermüller, F., & Roth, V. M. (1991). Communicative aphasia treatment as a further development of PACE therapy. *Aphasiology*, *5*(1), 39–50. https://doi.org/10.1080/02687039108248518

Purdy, M., & Van Dyke, J. A. (2011). Multimodal Communication Training in Aphasia: A Pilot Study. *Journal of Medical Speech-Language Pathology*, *19*(3), 45–53.

Purdy, M., Duffy, R., & Coelho, C. A. (1994). An investigation of the communicative use of trained symbols following multimodality training, *22*, 345–256.

Rebstock, A. M. (2014). *Effects of semantic feature analysis + multimodal communication program for word retrieval and switching behavior in primary progressive aphasia (Doctoral dissertation)*. Duquesne University.

Robey, R. R. (1998). A Meta-Analysis of Clinical Outcomes in the Treatment of Aphasia. *Journal of Speech, Language, and Hearing Research*, *41*(1), 172–16. https://doi.org/10.1044/jslhr.4101.172

Rose, M. L., & Attard, M. C. (2011). *M-MAT Procedure Manual [Manual]*. La Trobe University.

Rose, M. L., & Sussmilch, G. (2008). The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology*, *22*(7-8), 691–706. https://doi.org/10.1080/02687030701800800

Rose, M. L., Attard, M. C., Mok, Z., Lanyon, L. E., & Foster, A. M. (2013a). Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: A phase 1 study. *Aphasiology*, *27*(8), 938–971. https://doi.org/10.1080/02687038.2013.810329

Rose, M. L., Mok, Z., Carragher, M., Katthagen, S., & Attard, M. C. (2015). Comparing multi-modality and constraint-induced treatment for aphasia: a preliminary investigation of generalisation to discourse. *Aphasiology*, 1–21. https://doi.org/10.1080/02687038.2015.1100706

Rose, M. L., Raymer, A. M., Lanyon, L. E., & Attard, M. C. (2013b). A systematic review of gesture treatments for post-stroke aphasia. *Aphasiology*, *27*(9), 1090–1127. https://doi.org/10.1080/02687038.2013.805726

Schindler, A. (2009). Terminology in speech pathology: Old problem, new solutions. *Advances in Speech Language Pathology*, *7*(2), 84–86. https://doi.org/10.1080/14417040500125327

Schroeder, R. (2011). Modalities of Communication. In *Being There Together: Social Interaction in Shared Virtual Environments* (pp. 177–202). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195371284.003.0006

Sigafoos, J., Didden, R., Schlosser, R., Green, V. A., O'Reilly, M. F., & Lancioni, G. E. (2007). A Review of Intervention Studies on Teaching AAC to Individuals who are Deaf and Blind. *Journal of Developmental and Physical Disabilities*, *20*(1), 71–99. https://doi.org/10.1007/s10882-007-9081-5

Sparks, R., Helm, N., & Albert, M. (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex*, *10*(4), 303–316.

Speech Pathology Australia. (2011). Competency-Based Occupational Standards for Speech Pathologists (CBOS). Melbourne: The Speech Pathology Association of Australia Ltd.

Tanemura, J. (1999). Aphasia Therapy Using the Deblocking Method and Kanji/Kana Issues. *Topics in Stroke Rehabilitation*, *6*(1), 23–32. https://doi.org/10.1310/1HNC-DUKX-XY9E-T4W2

Thiel, L., Sage, K., & Conroy, P. (2015). Comparing uni-modal and multi-modal therapies for improving writing in acquired dysgraphia after stroke. *Neuropsychological Rehabilitation*, 1–29. https://doi.org/10.1080/09602011.2015.1026357

Thompson, C. K., & McReynolds, L. V. (1986). Wh interrogative production in agrammatic aphasia: an experimental analysis of auditory-visual stimulation and

direct-production treatment. *Journal of Speech and Hearing Research*, *29*(2), 193–206.

Thomson, Jennifer. (2012). *Assessing the benefits of multimodal rehabilitation therapy for aphasia [Masters Dissertation]* (pp. 1–143). University of Manchester.

Wallace, S. J., Worrall, L. E., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., et al. (2016). Which outcomes are most important to people with aphasia and their families? An international nominal group technique study framed within the ICF. *Disability and Rehabilitation*, *91*, 1–16. https://doi.org/10.1080/09638288.2016.1194899

Walsh, R. (2009). Meaning and purpose: A conceptual model for speech pathology terminology. *Advances in Speech Language Pathology*, *7*(2), 65–76. https://doi.org/10.1080/14417040500125285

Weill-Chounlamountry, A., Capelle, N., Tessier, C., & Pradat-Diehl, P. (2013). Multimodal therapy of word retrieval disorder due to phonological encoding dysfunction. *Brain Injury*, *27*(5), 620–631. https://doi.org/10.3109/02699052.2013.767936

Chapter 4 – What is already known about higher and lower frequency interventions for chronic aphasia?

Publication

Pierce, J. E., O'Halloran, R., Menahemi-Falkov, M., Togher, L., & Rose, M. L.

(2020). Comparing higher and lower weekly treatment intensity for chronic aphasia: A

systematic review and meta-analysis. *Neuropsychological Rehabilitation*, *34*(2), 1–25.

https://doi.org/10.1080/09602011.2020.1768127

Video abstract – click to follow link
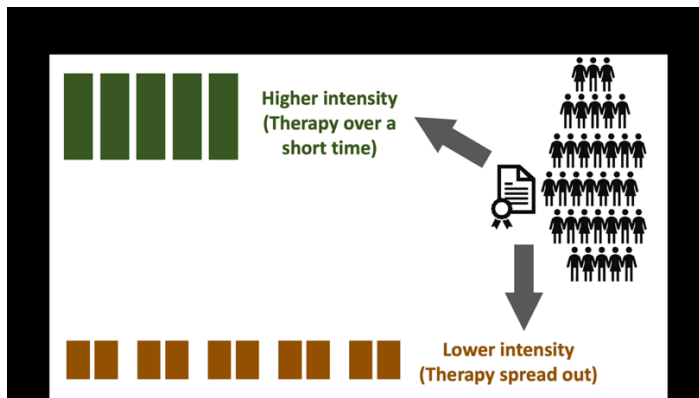
Video abstract (aphasia-friendly version)

Accepted Manuscript included here, in accordance with publisher's policy
https://authorservices.taylorandfrancis.com/publishing-open-access/open-access-cost-finder/

Comparing higher and lower weekly treatment intensity for chronic aphasia: A systematic review and meta-analysis.

John E. Pierce[a,b,c], Robyn O'Halloran[a,c], Maya Menahemi-Falkov[a,c], Leanne Togher[c,d] and Miranda L. Rose[a,c]*

[a]*School of Allied Health, La Trobe University, Melbourne, Australia;* [b]*Speech Pathology, Cabrini Health, Melbourne, Australia;* [c]*Centre of Research Excellence in Aphasia Recovery and Rehabilitation, Australia;* [d]*Speech Pathology, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia*

John E Pierce. 2-6 Hopetoun Street, Elsternwick 3185, Australia. +61430 130 893. ORCID identifier: 0000-0001-5164-5106 Twitter:@johnpierce85 pierce.john.e@gmail.com

Robyn O'Halloran. Health Sciences 1, La Trobe University, Plenty Road, Bundoora 3083, Australia. ORCID identifier: 0000-0002-2772-2164 R.OHalloran@latrobe.edu.au

Maya Menahemi-Falkov. Health Sciences 1, La Trobe University, Plenty Road, Bundoora 3083, Australia. mayam10@gmail.com

Leanne Togher. Faculty of Medicine and Health, The University of Sydney, 75 East St, Lidcombe 2141, Australia. ORCID identifier: 0000-0002-4518-6748 Leanne.togher@sydney.edu.au Twitter: @LeanneTogher

Miranda L. Rose. Health Sciences 1, La Trobe University, Plenty Road, Bundoora 3083, Australia. ORCID identifier: 0000-0002-8892-0965 M.rose@latrobe.edu.au

Abstract

Optimising intensity for aphasia treatment is a high priority research issue for people with aphasia, their families and clinicians, and could result in healthcare cost savings. An important aspect of intensity is the frequency of intervention, or how regularly treatment should be provided each week. While principles of neuroplasticity endorse massed practice, cognitive psychology has established superiority of distributed practice within normal learning. Neither concept has been conclusively tested in aphasia. There have been many literature reviews of intensity in aphasia intervention, but most have not investigated treatment intensity whilst also ensuring that therapy dose and treatment type are identical between study groups. Some have also combined studies across acute, subacute and chronic aphasia. We searched systematically for studies directly comparing higher and lower weekly treatment frequency in chronic aphasia. Eight studies were retrieved and rated for methodological quality. Meta-analysis was completed for group and single case experimental designs. Results showed that there are few studies investigating treatment frequency in chronic aphasia and their quality is low-moderate. Meta-analyses were inconclusive due to limited data, but there was no indication of either schedule being superior. Further research directly comparing treatment schedules is needed.

Word Count: 6348 (inc. tables, figures and citations)

Keywords: Aphasia, Systematic Review, Intensity, Therapy, chronic

Introduction

The most recent Cochrane systematic review demonstrated the effectiveness of aphasia therapy after stroke, but concluded that establishing optimum intensity of treatment should be a key aim for future research (Brady, Kelly, Godwin, Enderby, & Campbell, 2016). Intensity is a crucial component of aphasia treatment, with one author describing it as, "possibly the biggest challenge facing speech-language pathologists today — that is, how much treatment is enough to be effective?" (Togher, 2012, p. 438). The Canadian Stroke Network's consensus study placed the optimal intensity of aphasia therapy as the third priority for general stroke research (Bayley et al., 2007).

There are two underlying theories for rehabilitation which provide conflicting guidance for treatment intensity. Results from cognitive psychology research assert that lower frequency training schedules (distributed practice) are best for long-term retention of new skills or knowledge. Rehabilitation can be considered as a form of learning and thus, distributed treatments should result in superior retention in aphasia rehabilitation (For a comprehensive overview, see Dignam, Rodriguez, & Copland, 2016b). In contrast, principles of neuroplasticity promoted within neuroscience, which are largely derived from animal models and studies of human motor and sensory rehabilitation, hold that higher intensity schedules/massed practice facilitate better recovery or learning than lower intensity schedules/distributed practice (Dignam, Rodriguez, & Copland, 2016b). However, neither theory has been conclusively evaluated in aphasia rehabilitation.

Interpreting intensity

Intensity is a poorly defined term in speech pathology (Baker, 2012a), sometimes referring to the overall concept of treatment scheduling and duration, and at other times, to individual components. At a minimum, the variables that allow accurate reporting of intervention scheduling and intensity are:

1. The duration of each intervention session

2. The number of sessions in a given intervention period – the frequency of therapy sessions – most frequently described by the number of hours or sessions per week (e.g., Brady, Kelly, Godwin, & Enderby, 2012; Cherney, Patterson, & Raymer, 2011; Sage, Snell, & Lambon Ralph, 2011)

3. The duration of the whole intervention, typically number of weeks

In addition, some authors have called for treatment to be described more accurately by reporting treatment in terms of the number of "active ingredients" (e.g., one sentence production attempt) rather than by minutes or hours (Warren, Fey, & Yoder, 2007). In aphasia treatment, active ingredients have not yet been fully defined, but frameworks such as the Rehabilitation Treatment Taxonomy (Turkstra, Norman, Whyte, Dijkers & Hart, 2016) may provide a way to isolate and define such episodes. In this paper, we consider intensity as the amount of intervention provided in a given window, which can be considered at the level of *sessions* (active ingredients per session), *weeks* (active ingredients per week or hours per week), and the *total treatment* period (active ingredients or hours over the total duration of intervention). We will be investigating intensity at the level of each week rather than across the total treatment or within the session, though these are intrinsically linked, because weekly intensity is a commonly reported and practical measure (e.g. scheduling summaries in

Cherney, Patterson & Raymer, 2011; Brady et al., 2016; Dignam, Rodriguez & Copland, 2016b). We will use the term 'weekly intensity' as we feel is intuitive and transparent to clinicians and researchers alike.

'High' versus 'low' intensity

Weekly intensity has been of great interest to clinicians and researchers in recent years, possibly due to increasing research and implementation of high intensity treatments such as Constraint Induced Aphasia Therapy (CIAT, Pulvermüller et al., 2001). In simple terms, clinicians need to know whether a lot of therapy in a short period of time (high intensity) does more than the same amount in a longer period of time (low intensity). Even within this concept, high and low intensity is described within research using multiple terms, as shown in Table 4-1.

Table 4-1 – Terminology for high and low intensity

High intensity	Low intensity
Massed practice	Distributed practice
Intensive	Non-intensive
Intense	Non-intense

While most clinicians and researchers would agree that the 15 hours/week prescribed in CIAT is high intensity, there are no standard cut-offs for what constitutes 'low intensity'. Similarly, many papers discuss the efficacy of high intensity treatment for aphasia, yet 'high intensity' does not have a specific meaning or range. Previous research has created arbitrary boundaries for the comparison of high and low intensity (Cherney et al., 2011), with high and low intensity treatments grouped into:

- 3.5-10 versus 2 hours/week (Brady et al., 2016)

- 4-20 versus 1-4 hours/week (Brady et al., 2012)

- >8.8 versus <8.8 hours/week (Bhogal, Teasell, Foley, & Speechley, 2003a)

- 25 versus 4 hours/week (Hinckley & Carr, 2005)

- 5 versus 2 hours/week(Ramsberger & Marie, 2007)

Rather than considering weekly intensity in binary categories of high and low frequency, more accurate terminology might be 'high*er*' and 'low*er*' intensity as in reality, schedules exist on a continuum. At the upper end of the continuum lie intensive treatments such as CIAT which are typically 15 hours per week, with some Intensive Comprehensive Aphasia Programs providing up to 24 hours per week on average (Rose, Worrall & Cherney, 2013). Multiple studies have shown that outpatient aphasia treatment is rarely provided at more than 2-3 hours per week (Code & Heron, 2003; Mackenzie et al., 1993; Palmer, Witts, & Chater, 2018; Verna, Davidson, & Rose, 2009) which is the lower end of the continuum. With such a contrast between clinical practice and some high intensity aphasia treatments utilised within research, there will be significant implications for aphasia clinical practice worldwide if research does demonstrate superiority of higher intensity intervention.

Theoretical bases for massed and distributed schedules

Superiority of distributed practice has been demonstrated experimentally for learning and recall with a range of cognitive, verbal and motor activities (Dignam, Rodriguez, & Copland, 2016b). Distributed practice is thought to allow for more rehearsal between practice sessions and thus, deeper encoding (Moulton et al., 2006).

In addition, longer intervals between sessions reduces the ability to rely on priming from the previous session and encourages true recall, promoting deeper changes to the underlying representations (Sage et al., 2011).

In contrast, it has also been shown that neural connections are created or strengthened when events occur simultaneously and the strength of such connections increases proportionally with the frequency of occurrences (Pulvermüller & Berthier, 2008). For aphasia, the concept has been described as the Massed Practice Principle; that is, the hypothesis that more treatment and higher treatment intensity result in superior gains compared to less treatment and/or lower intensity (Pulvermüller & Berthier, 2008). The same reasoning leads to the proposal of a minimum threshold of treatment within a given timeframe that needs to be exceeded for the neural system to activate repair of connections or establish new pathways (Dignam et al., 2015; Harnish, Neils-Strunjas, Lamy, & Eliassen, 2008; Kleim & Jones, 2008).

The presence or absence of an activation threshold for aphasia treatment is a crucial piece of knowledge for clinicians. Therapy provided below this proposed minimum threshold would provide suboptimal results, or at worst, be completely ineffective (Baker, 2012a). Current low rates of treatment intensity (<3 hours/week) in clinical practice may sit below such a threshold. The risk of ineffective treatment is one reason treatment intensity is regarded as such a fundamental and pressing question for speech pathology (Baker, 2012a) and some have argued that if true, lower intensity would be unethical (Togher, 2012). However, neuroplasticity theory is largely based on research from motor actions and animal models to date (Kleim & Jones, 2008), and it remains unclear whether neuroplasticity or cognitive psychology models are more suitable for language recovery.

As a counterpart to a minimum threshold of activation, there might also be a ceiling level of treatment per week. In a synthesis of paediatric treatment for phonological awareness and print concepts, Schmitt and Justice (2012) noted that increasing the total dose does not indefinitely result in superior outcomes. They predicted diminishing returns after a point, where more treatment would not necessarily be better. Assuming a parallel within aphasia intervention, it is probable that there is an upper limit of effectiveness in terms of weekly dose, perhaps due to redundancy and patient fatigue. In this case speech pathology intervention provided *more* frequently than required would also be a waste of resources (Baker, 2012a), potentially stressful or even harmful.

Practical considerations

Conflicting cognitive psychology and neuroscience models notwithstanding, there are practical considerations for both higher and lower intensity of treatment. The higher treatment intensity of 15 hours per week found in some research studies is unlikely to be feasible within current healthcare models. In a survey of clinicians in the USA, 60% reported CIAT would be *very difficult* or *extremely difficult* to administer in their facility and 90% felt they would be unable to implement CIAT at all (Page & Wallace, 2014). Although this survey specifically enquired about CIAT, responses centred around the challenges of high intensity rather than specific treatment components of CIAT and the results are therefore likely to apply to other intensive treatments. Reimbursement from health insurers was also a common concern. Some authors predict that the financial constraints of healthcare are unlikely to change

significantly in the future (Code & Petheram, 2011) which would preclude uptake of more expensive treatments.

While feasibility does require consideration when developing treatments, the architecture of current service models should not solely dictate what treatments are developed and researched (the "tail wagging the dog"). The other counter-argument to higher intensity treatment being too challenging to implement is that, should higher intensity treatment prove to be superior, redesigning services to provide the *same* total dose of treatment in a shorter span of time could be an economically rational measure, as there would be greater recovery for the same hours. Given significant funding and resource limitations for aphasia rehabilitation worldwide (Rose, Ferguson, Power, Togher, & Worrall, 2013), funding could be best spent on shorter but more frequent intervention (Harnish et al., 2008). Such comparative effectiveness and economic data is not yet available.

Treatment adherence is another potential barrier to higher intensity treatments. Within physiotherapy, a survey found that the majority of patients were unwilling to participate in Constraint Induced Movement Therapy (CIMT) for post-stroke hemiparesis (Page, Levine, Sisto, Bond, & Johnston, 2002). The intensity of CIMT, which involves 6 hours/day over 2 weeks, was a major concern for these respondents. Surveyed clinicians also predicted their patients would be unlikely to adhere to CIMT. In a parallel survey for CIAT, more than 60% of clinicians believed their patients would be *unlikely* or *very unlikely* to adhere to the treatment protocol (Page & Wallace, 2014). Data on dropouts from a systematic review of aphasia treatment supports this view, finding significantly higher dropouts from more intensive treatments ($p = .03$) in acute and subacute aphasia (Brady et al., 2016). However, the analysis showed no significant

difference in dropouts between higher and lower intensity treatments in chronic

aphasia. This result might mean that people with chronic aphasia can tolerate intensive

treatments more readily than those in acute or sub-acute phases of recovery. Another

explanation is that people in acute and subacute phases are in the process of adjusting

to the stroke and learning about their capacity to participate in a range of life activities

including research activities. This may mean they are more likely to consent initially

without fully appreciating the demands of the research and later drop out. People with

chronic aphasia may be better able to predict the feasibility of attending intensive

treatments at the time of consenting. Additionally, people with chronic aphasia

generally cannot be routinely approached to participate in research and so researchers

would likely recruit enthusiastic participants who were motivated to approach the

researcher. Hence, the acceptability of higher intensity treatments across the

population with chronic aphasia is not yet clear.

Another concern reported by clinicians when asked about higher intensity

treatment was burden on caregivers in transporting patients to and from each session

(Page & Wallace, 2014). New models of care such as telehealth offer a possible solution

to travel and distance. However, in two studies of clinician perceptions after

administration of higher intensity treatments, other negatives reported were patient

and clinician fatigue (Gunning et al., 2016), frustration in patients who make limited

progress (Gunning et al., 2016) and unrealistic patient expectations of progress (Babbitt,

Worrall, & Cherney, 2013). The time required to plan and provide treatments was also a

concern (Babbitt et al., 2013). However, clinicians also identified that the progress seen

in patients was highly motivating and helped them appreciate the gains that are

possible (Babbitt et al., 2013). They described increased patient confidence (Gunning et

al., 2016) and improved relationships between patients, family and clinicians (Babbitt et

al., 2013; Gunning et al., 2016). For the clinicians themselves, reported rewards

included better teamwork and support and learning new techniques and clinical skills

(Babbitt et al., 2013; Gunning et al., 2016), though some of these benefits might be

attributable to the group aspect of therapy rather than the higher intensity. Finally, in

one of the studies, clinicians reported that returning to typical (non-intense) clinical

practice was difficult and even "depressing," as they felt that the designs of their

services presented fewer opportunities to offer meaningful treatment gains and high

quality therapy (Babbitt et al., 2013).

Previous research on intensity

So far, theoretical and practical arguments have been discussed for higher

versus lower intensity treatments, but in clinical research, is higher or lower weekly

intensity more efficacious for patient outcomes? There have been multiple previous

reviews on intensity, both systematic and narrative, with conflicting results. However,

the methods in these reviews do not allow conclusions to be made about treatment

frequency in chronic aphasia.

Most past reviews have examined studies that compared different treatment

frequencies, but also different total durations of treatment (Brady et al., 2016; Cherney,

Patterson, Raymer, Frymark, & Schooling, 2008). For example, the most recent

Cochrane review concluded that there is some tentative evidence for higher intensity

being more efficacious, yet the higher intensity arms of these studies provided a mean

of 84 total hours per participant, while lower intensity arms provided 44 hours (Brady et

al., 2016). Thus, with higher intensity treatments providing nearly double the treatment hours, the effects of *more* therapy overall and the *frequency* of therapy are conflated.

Previous reviews have also included papers that provide a *different* therapy in each arm (Brady et al., 2016; Cherney et al., 2008). For example, Pulvermüller et al. (2001) is often included as evidence in favour of high intensity treatment (as in Bhogal, Teasell, & Speechley, 2003b), yet that research compared two different treatment approaches in each arm and thus the contribution of the intensity and the type of treatment cannot be separated. To isolate the effect of treatment intensity, the same treatment should be offered at each intensity.

Finally, some reviews of intensity aggregated data from all phases of aphasia recovery, including acute, subacute and chronic (Bhogal et al., 2003b). Based on previous analyses of treatment effect sizes, early and chronic aphasia should not be expected to improve by the same magnitude (Brady et al., 2016; Robey, 1998). The specific response to higher or lower intensity of treatment could also differ between chronic and acute/subacute patients (Cherney et al., 2011). For example, the principles of neuroplasticity that promote higher intensity therapy might be more relevant to a brain which has recently been injured, and the cognitive psychology findings related to enhanced learning in a lower intensity might apply more to the more 'stable' neurophysiology of a brain with chronic aphasia.

In order to accurately determine whether 'a lot in a little time does more than a lot in a longer time,' treatment schedules need to be altered between treatment arms with all other variables controlled, including the type and total dose of therapy and the participant characteristics. One review to date has been careful to include only studies with these criteria (Dignam, Rodriguez, & Copland, 2016b). This paper was a narrative

review of research up to November 2014 and identified four papers. The authors concluded that there is some preliminary suggestion of superiority of lower intensity treatment when considering the longer-term maintenance timepoints. As a narrative review, and no doubt due to the low yield of papers, Dignam et al. did not attempt a meta-analysis. As a high priority topic, it is likely that further relevant evidence has been published in the past five years. The aim of this paper is to systematically review papers that directly compare higher and lower weekly intensity treatments while controlling other variables and, if feasible, meta-analyse results of high-quality papers.

Methods

Four major databases were searched in June 2018: Medline (OVID, 1946-present), CINAHL (Ebscohost), PubMed and Psycinfo (OVID, 1987-present). The search strategy combined two key concepts, **aphasia** and **treatment intensity**, and a variety of search terms to represent these were used. An iterative approach using known results was used to ensure all relevant terms were included. An example of the search strategy employed is found in Table 4-2. Medical Subject Headings (MeSH) and keyword searches were used. Searches were limited to English only and terms were exploded where databases allowed. Results were imported into citation management software to identify duplicates manually and automatically, and non-English papers were excluded. Results were then exported into Rayyan (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016), online systematic review software, for screening.

Table 4-2 – Search strategy example, Medline (OVID, 1946-present)

Keywords:	aphasi*	**AND**	duration
	dysphasi*		total hours
	anomic		total therapy
	anomia		total treatment

	NOT "progressive aphasia"		total intervention dose cumulative intervention intensity intensive amount of therapy amount of treatment amount of intervention intense
	OR		
Subject headings:	*exp: Aphasia/* *exp: anomia/* *NOT exp: Aphasia, primary progressive/*		

Note. All subject headings were exploded, and all subheadings were included.

A search was also conducted in Google Scholar in August 2018 using a variety of search phrases (see Table 4-3). As recommended in systematic review research (Haddaway, Collins, Coughlin, & Kirk, 2015), the first 300 results were collected, and then the search was repeated using the operator "AllInTitle:" which only returns results with search terms in the titles.

The first and third author independently screened remaining articles on Rayyan using the following inclusion criteria:

1. English language

2. Adults with chronic stroke-induced aphasia (≥ 6months)

3. Original research data

4. Two schedule (intensity) conditions with the same total dose of therapy (e.g. 30 hrs at each intensity)

After screening, the two authors discussed and resolved discrepancies in inclusion and exclusion decisions. Full texts of the remaining articles were then obtained

and further examined according to the inclusion criteria. Reference lists of resulting

articles were inspected by the first author to identify any additional studies.

Table 4-3 – Search strategy, Google Scholar

Search 1:	"aphasia duration" OR "total hours" OR "total therapy" OR "total treatment" OR "total intervention" OR dose OR "cumulative intervention" OR intensity OR intensive OR "amount of therapy" OR "amount of treatment" OR "amount of intervention" OR intense"
Search 2:	Allintitle: "aphasia duration" OR "total hours" OR "total therapy" OR "total treatment" OR "total intervention" OR dose OR "cumulative intervention" OR intensity OR intensive OR "amount of therapy" OR "amount of treatment" OR "amount of intervention" OR intense"

Included articles were categorised by study type using a modified version of the

Oxford Centre for Evidence Based Medicine levels of evidence (OCEBM Levels of

Evidence Working Group, 2011). An additional level was included for Single Case

Experimental Designs (SCEDs) outside of the n-of-1 design, including multiple baseline,

changing criterion, alternating treatment and withdrawal designs. Rather than

considering mean changes within or between groups, SCEDs repeatedly assess

individuals over time and determine whether manipulation of treatment (e.g.

commencing and withdrawing) results in clear changes to outcomes. When conducted

rigorously, SCEDs are recognised alongside n-of-1 designs as comprising a high level of

experimental control (Tate et al., 2013). The first and third authors categorised papers

and reached agreement.

Randomised and non-randomised controlled trials were rated for

methodological rigour with the PEDro-P scale (Murray et al., 2013). SCEDs were rated

for methodological rigour with the RoBiNT scale (Tate et al., 2015). PEDro-P was rated

by the first and third authors and SCEDs by the first and last authors. Discrepancies were resolved by discussion. Remaining study types, including pre/post case series and non-experimental single case designs such as single phase or biphase designs, were not rated as these typically contain no experimental control and form a low level of evidence (Tate et al., 2013).

Data from each paper was extracted to a spreadsheet including participant characteristics, treatment type, outcome measures, therapy schedule, and results.

Effect size calculation - group studies

The primary outcome measures at the post-intervention timepoint were meta-analysed as all were measures of expressive language. The mean and standard deviation of change scores per arm were calculated and analysed using RevMan (The Cochrane Collaboration, 2014). A random effects model was applied with a standardised mean difference, given the different outcome measures across studies. Only data from the first phase was considered in crossover studies.

For the maintenance timepoint, studies containing data at one month post intervention were meta-analysed. For Mozeiko et al. (2015), the one month follow-up scores and mean baseline scores were used to calculate a change score as a percentage of baseline within each participant. This approach was more conservative than using Tau-U which would have resulted in ceiling (1.0) or floor (-1.0) scores due to there being only one data point at maintenance. A forest plot was generated using Revman.

Effect size calculation - single case experimental designs

Data points were measured graphically using software image measurements when not obvious directly from the chart. Baselines were corrected for trend if the

baseline phase Tau was >0.40 *and* a trend was apparent on visual analysis. Tau was

consistent with visual analysis on each occasion. Tau-U was calculated using the

software at www.singlecaseresearch.org. To meta-analyse the maintenance timepoint,

only SCEDs that included "withdrawal" phases (Ramsberger & Marie, 2007; Raymer,

Kohen, & Saffell, 2006) had Tau-U applied, and this was conducted as baseline

compared to maintenance data. This comparison was made because language

behaviour is unlikely to rapidly react to withdrawal of treatments and the improvement

compared to pre-treatment is the most relevant outcome. A forest plot was generated

using the DistillerSR Forest Plot Generator from Evidence Partners (Evidence Partners,

2019) and overall effect sizes were calculated weighted by the number of pairs.

Results

Results of the search and screening process are displayed in Figure 4-1. The

initial yield was high (2462) but after screening, eight papers remained which

represented seven studies (two papers reported on different aspects of the same

study). Hand searching reference lists revealed two additional references. One was

added to the review while the other was a conference abstract and the authors stated

the data was not ready to be released (Rochon et al., 2016). The final yield therefore

included nine articles reporting on eight studies. It should be noted that Stahl et al.

(2018) was a comparison of different schedules *and* total dose, but published data

allowed comparison of the two groups at a timepoint where each had received 24 hours

of treatment.

Figure 4-1 - Search strategy results

Medline
Ovid, 1946 - June 2018
6/6/2018
429

CINAHL
EBSCOhost, Aug 2017
32 Citations

Psycinfo
Ovid, 1806 - June 2018
6/6/2018
698 Citations

Pubmed
June 2018
1076

1775 non-duplicate citations

687 duplicates

Google Scholar
August 2018
227

Title and abstract screening

1757 excluded as irrelevant
1 duplicate

17 full text articles for review

8 articles excluded:
Not chronic aphasia 1
Inconsistent dose 3
No dose comparison 1
Unpublished 1
No full text 1
Inconsistent treatment 1

Inclusion/Exclusion criteria applied

8 articles included (7 studies)

2 further citations identified**:**
1 passed criteria
1 unpublished

References hand-searched

9 articles included

Level of evidence

Results included one randomised controlled trial (Stahl et al., 2018) and one non-randomised controlled trial represented by two papers (Dignam et al., 2015; Dignam, Copland, Rawlings, OBrien, et al., 2016a). Sage, Snell and Lambon Ralph (2011) employed a crossover design with randomisation. When considering only the data from the first condition for each participant, this design can be considered a randomised controlled trial. There were three SCEDs: two crossover multiple baseline designs (Ramsberger & Marie, 2007; Raymer et al., 2006) and a biphasic design (Mozeiko, Coelho, & Myers, 2015). Two papers were pre/post case series (Harnish et al., 2008; Marcotte et al., 2018) and were not considered in meta-analysis.

Descriptive characteristics of data

The Appendix 3 contains extracted data for retrieved studies. Overall, there were 92 participants across the studies, with a mean reported age of 59 (SD 13). There were 60 males and 32 females, a higher proportion of males (1.875 : 1) than overall stroke populations (1.4 : 1) (Appelros, Stegmayr, & Terént, 2009). The mean number of months post stroke in papers reporting this data was 48 (range 4 months to 21 years), though these were not normally distributed, with 65/83 (78%) participants less than six years post stroke. Data from participants who were less than six months post stroke were excluded from meta-analysis.

Four studies investigated traditional, cueing hierarchy-based naming treatments (two using software), two Constraint Induced Aphasia Therapy, one an Intensive Comprehensive Aphasia Program and one Phonological Components Analysis. The total dose ranged from 10-48 hours, assuming sessions were one hour, as some reported in

sessions per week rather than hours. The weekly intensity for the higher intensity schedules (as labelled by the authors) was 3-16 hours per week and for the lower schedules, 1-6 hours. Within each study, the weekly dose for higher intensity schedules was at least double the lower intensity schedule, with one paper providing five times more per week (Mozeiko et al., 2015).

The six studies used a range of impairment-based measures as their primary outcome, including picture naming of treated items (3), a naming test (1), subtests from a language battery (1) and a discourse measure (1). Only two studies investigated activity/participation outcomes (Dignam et al., 2015; Mozeiko et al., 2015), both as secondary outcomes, while only Dignam et al. investigated quality of life.

Methodological quality

Methodological ratings for papers are displayed in Table 4-4. For group trials, Stahl et al. (2018) was rated as high quality methodology (7/10) while Dignam et al. (2015) and Sage et al. (2011) were rated as fair (4/10), as per commonly accepted benchmarks for the Pedro scale (Stroke Engine, n.d.). For the RobinT, a recently published algorithm indicates the risk of bias for internal validity (Perdices, Tate & Rosenkoetter, 2019). The algorithm gives higher weighting to items considered more important for internal validity. Using this algorithm, the SCEDs in this review had low (Ramsberger & Marie, 2007; Raymer et al., 2006) and very low (Mozeiko et al., 2015) methodological rigour for internal validity. The RobinT does not yet have consensus on a cut-off or benchmark of scores for overall methodology that includes external validity.

Table 4-4 - Methodological ratings

PEDro-P	Total	Eligibility criteria + source	Random Allocation	Allocation concealed	Baseline Similarity	Participant Blinding	Therapist blinding	Assessor blinding	85% retention	ITT Analysis	b/w group Comparisons	Point estimates and variability
Sage et al., 2011	4	✓	✓	×	×	×	×	×	✓	×	✓	✓
Stahl et al., 2018	7	✓	✓	✓	✓	×	×	✓	✓	×	✓	✓
Dignam et al., 2015	4	×	×	×	✓	×	×	×	✓	×	✓	✓

RobinT	Total	Design with control	Randomisation	Sampling behaviour	Participant Blinding	Assessor blinding	Interrater agreement	Treatment Adherence	Baseline characteristics	Setting	Dependent variable defined	Independent variable defined	Raw data record	Data analysis	Replication	Generalisation
Mozeiko et al., 2015	15	×	×	1	×	×	1	×	2	1	2	1	×	2	2	1
Ramsberger & Marie, 2007	17	1	1	×	×	×	1	×	2	1	2	2	2	2	2	1
Raymer et al., 2006	17	1	1	1	×	×	1	×	2	×	1	2	×	2	2	2

Meta-analysis - immediately post intervention

The forest plot for outcomes immediately post intervention for the three group trials is shown in Figure 4-2. The total overall estimate indicates that a distributed, lower weekly intensity schedule is superior (p = .02), though this result is based on a small data set (n = 70). The Tau-U meta-analysis chart for three SCEDs is shown in Figure 4-3 (see note on interpretation). Visual comparison of higher and lower intensity effect sizes reveals that higher intensity results had a marginally higher weighted effect size, but the confidence intervals of both schedules overlap, indicating no significant differences.

Figure 4-2 – Forest plot for group studies, post treatment

Study or Subgroup	Higher Mean	SD	Total	Lower Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Sage 2011	22	6.6	3	20.8	3.8	5	11.4%	0.21 [−1.23, 1.65]
Stahl 2018	1.6	0.6	15	2.2	0.8	15	41.9%	−0.83 [−1.58, −0.08]
Dignam 2015	3.2	3.4	16	5.4	3.8	16	46.7%	−0.59 [−1.30, 0.12]
Total (95% CI)			34			36	100.0%	−0.60 [−1.09, −0.11]

Heterogeneity: Tau² = 0.00; Chi² = 1.57, df = 2 (P = 0.46); I² = 0%
Test for overall effect: Z = 2.42 (P = 0.02)
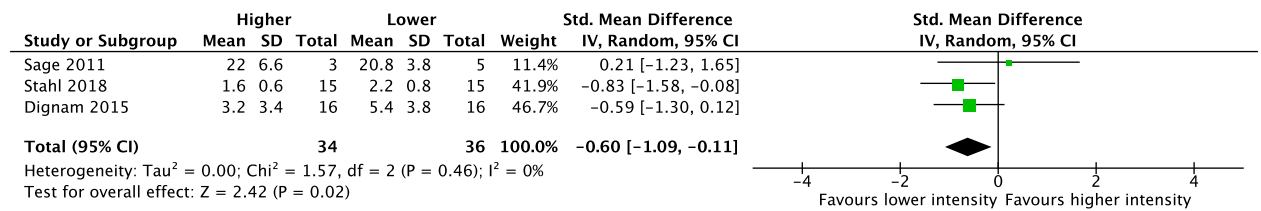
Figure 4-3 – Tau-U plot for single case experimental designs, post treatment
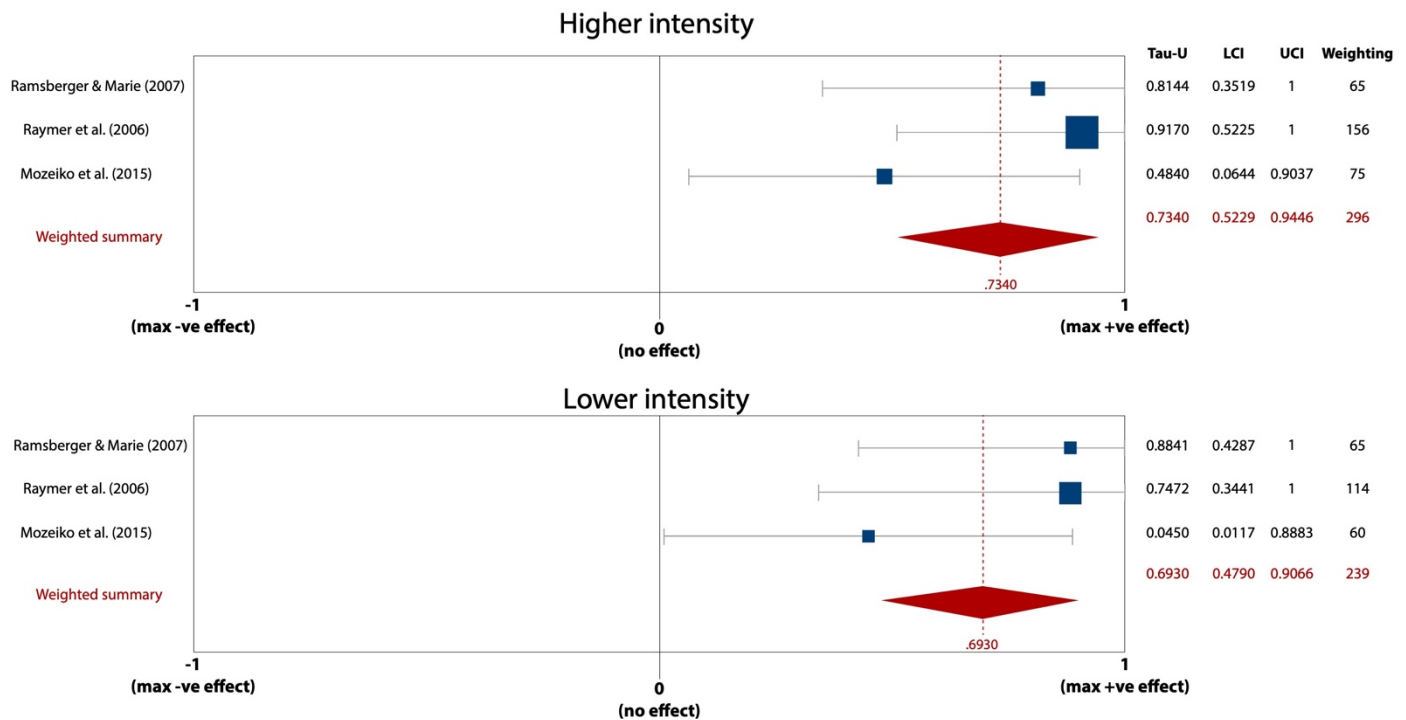
Note: Upper plot is higher intensity phases, lower plot is lower intensity. Lines −1.0 and 1.0 indicate the possible limits of Tau-U scores; positive values demonstrate improvement relative to baseline. Dotted lines indicate the weighted average and diamond width indicates the 95% confidence interval. Square sizes indicate relative weighting for each study.

Meta-analysis - one month post intervention

The total overall estimate in the forest plot (Figure 4-4) shows no significant

difference between intensities at this timepoint, again based on limited data (*n* = 47). In

the Tau-U chart for two SCEDs (Figure 4-5), both charts show a large effect size with no

appreciable difference between intensities.

Figure 4-4 – Forest plot one month follow-up

Study or Subgroup	Higher intensity Mean	SD	Total	Lower intensity Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Dignam 2015	2.7	4.5	16	7.7	4.8	15	47.3%	−1.05 [−1.81, −0.29]
Sage 2011	15.7	7.2	3	18.8	4	5	26.0%	−0.51 [−1.98, 0.96]
Mozeiko 2015	1	1.2	4	0.4	0.4	4	26.7%	0.58 [−0.86, 2.03]
Total (95% CI)			23			24	100.0%	−0.47 [−1.43, 0.49]

Heterogeneity: Tau² = 0.36; Chi² = 3.90, df = 2 (P = 0.14); I² = 49%
Test for overall effect: Z = 0.97 (P = 0.33)
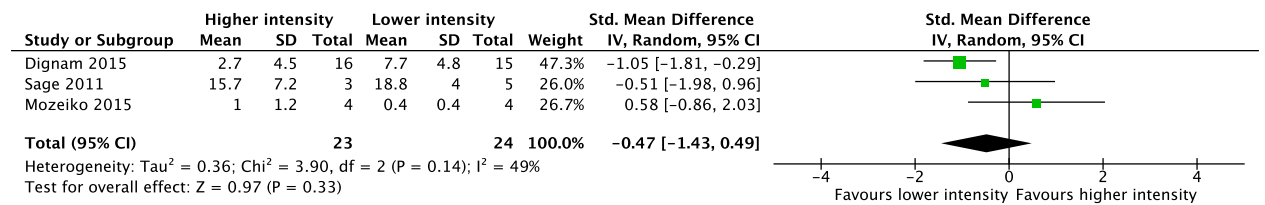
Figure 4-5 – Tau-U plot for Single Case Experimental Designs, one month follow-up
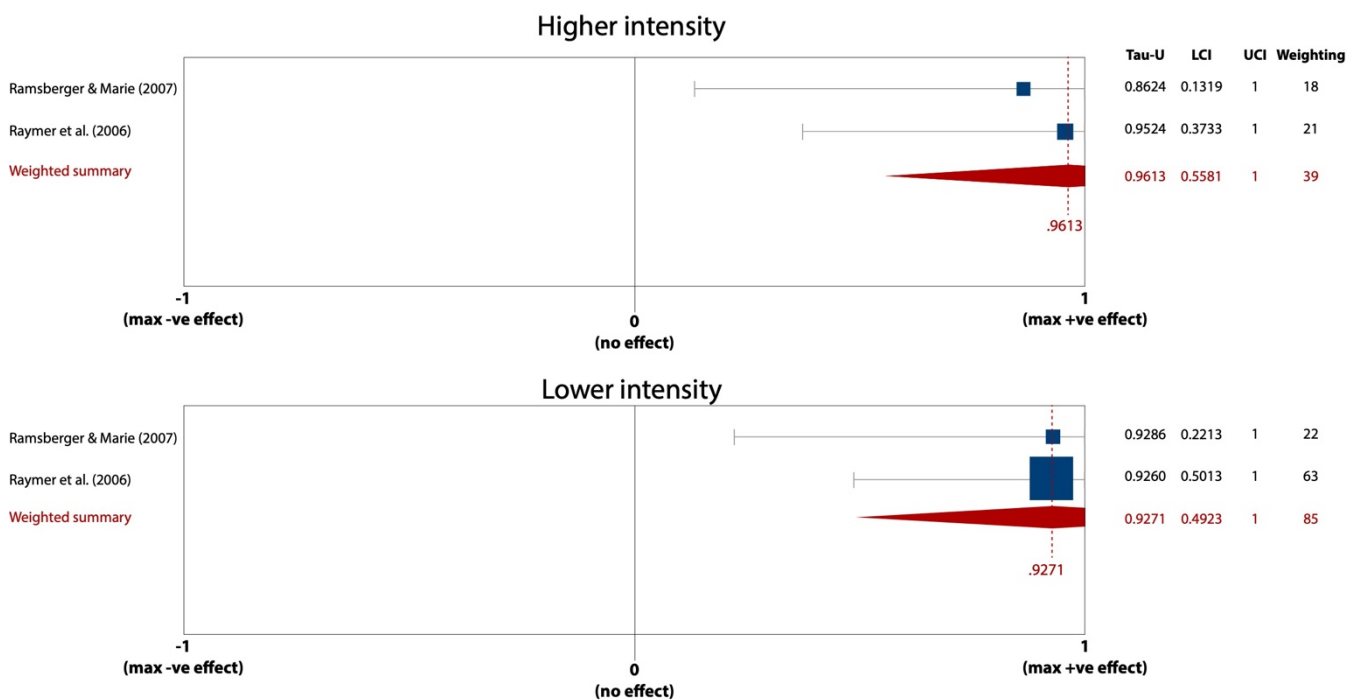
Secondary Outcomes

In the two studies employing activity/participation outcomes (Dignam et al., 2015; Mozeiko et al., 2015), results do not clearly favour either intensity. Mozeiko et al. (2015) used the Communication Activities of Daily Living-2 (CADL-2) which assesses functional communication in any modality across a range of simulated scenarios. Participants in the higher intensity arm were unchanged or improved (0, +6, +11, +25 point changes) and mixed in the lower intensity arm (-14, -12, +3, +38 point changes). Dignam et al. (2015) found no significant between group differences on the Communication Effectiveness Index ($p = .05$ post, $p = .21$ Follow-Up) or the Communication Confidence Rating Scale for Aphasia ($p = .79$ post, $p = .48$ Follow-Up), although both improved compared to baseline. In the one study that employed a quality of life outcome measure (Dignam et al. 2015), no difference was found between higher and lower intensity groups ($p = .37$ post, $p = .75$ Follow-Up).

Discussion

Considering the importance of the question of weekly intensity for clinicians, patients and health service funding bodies, this systematic review found a surprisingly low number of studies examining the question directly in chronic aphasia. However, the eight studies retrieved in this review includes four new studies (Dignam et al., 2015; Marcotte et al., 2018; Mozeiko et al., 2015; Stahl et al., 2018) since the recent narrative review by Dignam, Rodrizuez and Copland (2016b), particularly considering we limited our search to chronic aphasia whereas they included all phases. More may have been published in acute or subacute aphasia since this time. It is unclear why there have not been larger scale studies regarding this question. Perhaps the relative efficacy of different approaches – another important topic for aphasia – is considered by

researchers to be a higher priority than scheduling of any single treatment. Whatever the reason for the low number of studies, the demand for guidance on intensity demonstrates a disconnect between research priorities and practice.

In our review results, both the SCEDs and the group trials had high risk of bias. To more stringently examine the influence of weekly intensity on outcomes, group studies should randomise participants (2/3 group studies were randomised) and use blinded assessors (1/3). An intention to treat approach (0/3) is also recommended, as the acceptability and feasibility of different treatment schedules has not been clearly established for chronic aphasia. SCEDs addressing this question should employ more stringent designs such as Multiple Baseline Designs, collect a minimum of five data points per phase, and randomise the point at which treatment commences.

Impairment based outcome measures were employed with all retrieved studies, however, only two used activity/participation level outcomes and one used a quality of life measure; all as secondary outcomes. People with aphasia and their families have identified that impairment, activity/participation and wellbeing levels are important priorities for treatment outcomes (Wallace et al., 2016). There were no clear differences in the data retrieved in this review, but it is possible that higher and lower intensity schedules have different effects on these outcome domains compared to impairment, and thus more investigation of such outcomes with contrasting weekly intensities is required. Specifically, a Core Outcome Set was recently developed based on patient, family and clinician input and use of these outcome measures will ensure that future research is consistent, comparable and covers key domains (Wallace et al., 2016).

Although studies employed a disparate range of weekly schedules, within each study the schedules were sufficiently varied to allow comparison, with higher intensity

arms at least double the weekly dose as the lower intensity arms. In addition, 6/8 studies used <3 hours per week for the lower intensity arm which is a schedule similar to most clinical practice worldwide (Mackenzie et al., 1993; Palmer et al., 2018; Verna et al., 2009). Comparison of typical clinical schedules to higher intensity means that results can be used to aid decision making with regards to increasing intensity from current clinical levels.

Importantly, for most papers *both* treatment intensities showed a positive mean change for the primary outcome measure. Overall, results tentatively suggested that no schedule was ineffective and thus current clinical practice, although considered low intensity, is likely superior to no treatment – a finding consistent with the most recent Cochrane review of aphasia treatment (Brady et al., 2016).

The meta-analysis conducted in this review is preliminary and limited by a low yield of papers, low participant numbers and insufficiently rigorous designs. The immediate post-treatment meta-analysis for group studies was drawn from a very small pool of data, but it is interesting that the lower weekly intensity had a stronger effect at this timepoint. If this finding was replicated in future research, this would contradict neuroplasticity principles which predict better immediate performance with intensive practice. However, the forest plot does show wide confidence intervals which indicates uncertainty of the result, and the results shown on the SCED figure did not favour either intensity at the same timepoint.

For the meta-analysis of the maintenance timepoint (1 month), there was no apparent difference between schedules on either meta-analysis, yet cognitive psychology predicts that lower intensity would be more favourable in the longer term. If intensity had a simple linear correlation with outcomes, some sort of trend in results

might be expected given the significant difference in schedules between participants, even considering the limitations of this meta-analysis. Thus, the inconclusive result of this review is an interesting outcome in itself.

Higher versus lower intensity is a crucial issue, but it should not be taken as the only important treatment variable. Many authors have suggested that there are multiple interacting factors that moderate response to intensity (Baker, 2012b; Cherney et al., 2011; Dignam, Rodriguez, & Copland, 2016b). As one example, Baker (2012b) described a finding in paediatric treatment where the mean response to 1/week versus 5/week schedules was non-significant, but those without Down Syndrome and those who played with many objects did respond more to higher intensity (Yoder et al. 2012, as cited in Baker, 2012b). Within this review, Ramsberger and Marie (2007) had one participant respond more favourably to higher intensity while the remaining three showed no difference between schedules. Sage, Snell and Lambon-Ralph (2011) also reported one participant who demonstrated better naming at both timepoints for the wordlist treated at higher intensity, while the other participants showed no significant overall difference. There were no obvious factors in this participant to explain their stronger response.

Kiran and Thompson (2019) outlined a list of variables that influence recovery from stroke related aphasia. These included aphasia severity, cognitive reserve, psychosocial environment, age at time of stroke, lesion volume, lesion location and impact on connectivity, perfusion characteristics and treatment variables. The authors point out that these factors are not linearly correlated with recovery and form a complex picture. Any one of these factors, and probably many of them, may interact with intensity so that particular patient profiles respond better to higher intensity and

others to lower intensity (Brady et al., 2016). These questions require large datasets to solve.

Further to moderating factors, intensity is not simply a matter of "counting the hours" (Togher, 2012). For example, the active ingredient is important to report as it provides a more exact estimate and description of the treatment provided (Baker, 2012b). In addition, the number of active ingredients per session provided to participants is likely to vary across therapist and patient, depending on factors such as talkativeness and severity. As an extreme, one hour of therapy might contain, for example, several hundred naming attempts or only a handful. In comparative studies, active ingredients may therefore be an uncontrolled variable where not reported. While all studies retrieved in this review reported on the number of hours or sessions provided to participants, the active ingredient was not described with the exception of Dignam et al. (2015) who showed that the session dose and total dose were not significantly different between groups. Randomised trials of sufficiently large sample size are likely to control for this variance.

Limitations

This review considered only studies which compared different weekly intensities in a binary manner. Other reviews have also included papers in which treatment hours varied between each participant, allowing a correlational analysis of intensity (Cherney et al., 2011). In our case we were looking for maximally contrasting schedules in order to find any clear indications of higher or lower intensity being superior to one another. For some retrieved papers, the duration of each session was not reported and we had to assume these were one hour. Although this assumption might be imprecise, our

analysis was based on the comparison *within each study*, thus the units involved are not crucial. Assuming a simple linear relationship (e.g. higher intensity = better outcomes), each paper should have found better outcomes for the higher intensity treatment regardless of units. As higher and lower intensities overlapped between retrieved studies, this review is unable to provide guidance on *optimal* weekly intensity but instead allowed us to investigate the simple linear predictions of neuroplasticity and cognitive learning.

Conclusions

Results of this review are preliminary, given the limited number of studies retrieved, yet they did not support the predictions of neuroplasticity or cognitive psychology. If anything, lower intensity was marginally more favourable immediately after treatment while no difference was demonstrated at one month follow up. This is in contrast to many current perspectives – higher intensity is often described as superior as a foregone conclusion, yet this review demonstrates that there is no strong evidence for this in chronic aphasia.

As described above, many authors have concluded that treatment response is likely to emerge as a highly complicated system of interacting factors. Speech pathologists seeking an answer to the question of how often to schedule therapy may find this frustrating. Nonetheless, even lower intensity treatments had favourable outcomes and therefore clinicians who are limited to these treatment schedules due to funding and policy constraints can be confident in providing lower intensity services. It will be many years before large scale models are produced which provide accurate predictions of treatment response based on patient and treatment factors – an example of such work is in progress from the RELEASE collaboration (Brady et al., 2019).

However, despite the complexity of treatment scheduling, weekly intensity remains an important piece of the treatment puzzle. The call for larger RCTs directly comparing aspects of intensity, put forward by many before (Cherney et al., 2011; Dignam, Rodriguez, & Copland, 2016b), remains. Larger sample sizes with low risk of bias will control for the variance between clinicians, patient demographics and aphasia and stroke profiles and ultimately add data to this question.

Declaration of Interest

The authors report no conflicts of interest

References

Appelros, P., Stegmayr, B., & Terént, A. (2009). Sex differences in stroke epidemiology. *Stroke*, *40*(4), 1082–1090.

Babbitt, E. M., Worrall, L. E., & Cherney, L. R. (2013). Clinician perspectives of an intensive comprehensive aphasia program. *Topics in Stroke Rehabilitation*, *20*(5), 398–408. https://doi.org/10.1310/tsr2005-398

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

Baker, E. (2012b). Optimal intervention intensity in speech-language pathology: Discoveries, challenges, and unchartered territories. *International Journal of Speech-Language Pathology*, *14*(5), 478–485. https://doi.org/10.3109/17549507.2012.717967

Bayley, M. T., Hurdowar, A., Teasell, R., Wood-Dauphinee, S., Korner-Bitensky, N., Richards, C. L., et al. (2007). Priorities for stroke rehabilitation and research: results of a 2003 Canadian Stroke Network consensus conference. *Archives of Physical Medicine and Rehabilitation*, *88*(4), 526–528. https://doi.org/10.1016/j.apmr.2007.01.005

Bhogal, S. K., Teasell, R. W., Foley, N. C., & Speechley, M. R. (2003a). Rehabilitation of aphasia: more is better. *Topics in Stroke Rehabilitation*, *10*(2), 66–76. https://doi.org/10.1310/RCM8-5TUL-NC5D-BX58

Bhogal, S. K., Teasell, R., & Speechley, M. (2003b). Intensity of aphasia therapy, impact on recovery. *Stroke; a Journal of Cerebral Circulation*, *34*(4), 987–993. https://doi.org/10.1161/01.STR.0000062343.64383.D0

Brady, M. C., Kelly, H., Godwin, J., & Enderby, P. (2012). Speech and language therapy for aphasia following stroke. *The Cochrane Database of Systematic Reviews*, *5*(5), CD000425. https://doi.org/10.1002/14651858.CD000425.pub3

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *The Cochrane Database of Systematic Reviews*, *6*(6), CD000425–CD000425. https://doi.org/10.1002/14651858.CD000425.pub4

Cherney, L. R., Patterson, J. P., & Raymer, A. M. (2011). Intensity of aphasia therapy: evidence and efficacy. *Current Neurology and Neuroscience Reports*, *11*(6), 560–569. https://doi.org/10.1007/s11910-011-0227-6

Cherney, L. R., Patterson, J. P., Raymer, A. M., Frymark, T., & Schooling, T. (2008). Evidence-Based Systematic Review: Effects of Intensity of Treatment and Constraint-Induced Language Therapy for Individuals With Stroke-Induced Aphasia. *Journal of Speech, Language, and Hearing Research*, *51*(5), 1282–1299. https://doi.org/10.1044/1092-4388(2008/07-0206)

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), 1231–1237.
https://doi.org/10.1080/09638280310001599961

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

Dignam, J. K., Copland, D., McKinnon, E., Burfein, P., OBrien, K., Farrell, A., & Rodriguez, A. (2015). Intensive Versus Distributed Aphasia Therapy A Nonrandomized, Parallel-Group, Dosage-Controlled Study. *Stroke*, *46*(8), 2206–2211.
https://doi.org/10.1161/STROKEAHA.115.009522/-/DC1

Dignam, J. K., Copland, D., Rawlings, A., OBrien, K., Burfein, P., & Rodriguez, A. D. (2016a). The relationship between novel word learning and anomia treatment success in adults with chronic aphasia. *Neuropsychologia*, *81*, 186–197.
https://doi.org/10.1016/j.neuropsychologia.2015.12.026

Dignam, J. K., Rodriguez, A. D., & Copland, D. (2016b). Evidence for Intensive Aphasia Therapy: Consideration of Theories from Neuroscience and Cognitive Psychology. *American Academy of Physical Medicine and Rehabilitation*, *8*(3), 254–267. https://doi.org/10.1016/j.pmrj.2015.06.010

Evidence Partners. (2019). DistillerSR Forest Plot Generator. Retrieved from https://www.evidencepartners.com/resources/forest-plot-generator/

Gunning, D., Wenke, R., Ward, E. C., Chalk, S., Lawrie, M., Romano, M., et al. (2016). Clinicians' perceptions of delivering new models of high intensity aphasia treatment. *Aphasiology*, *31*(4), 406–426.
https://doi.org/10.1080/02687038.2016.1236359

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PloS One*, *10*(9), e0138237. https://doi.org/10.1371/journal.pone.0138237

Harnish, S. M., Neils-Strunjas, J., Lamy, M., & Eliassen, J. C. (2008). Use of fMRI in the study of chronic aphasia recovery after therapy: a case study. *Topics in Stroke Rehabilitation*, *15*(5), 468–483. https://doi.org/10.1310/tsr1505-468

Hinckley, J., & Carr, T. (2005). Comparing the outcomes of intensive and non-intensive context-based aphasia treatment. *Aphasiology*, *19*(10-11), 965–974.
https://doi.org/10.1080/02687030544000173

Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of Language Networks in Aphasia: Advances, Updates, and Future Challenges. *Frontiers in Neurology*, *10*, S225.
https://doi.org/10.3389/fneur.2019.00295

Kleim, J. A., & Jones, T. A. (2008). Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. (Vol. 51, pp. S225–39). Presented at the Journal of speech, language, and hearing research : JSLHR.
https://doi.org/10.1044/1092-4388(2008/018)

Mackenzie, C., Le May, M., Lendrem, W., McGuirk, E., Marshall, J., & Rossiter, D. (1993). A survey of aphasia services in the United Kingdom. *European Journal of*

Disorders of Communication : the Journal of the College of Speech and Language Therapists, London, *28*(1), 43–61.

Marcotte, K., Laird, L., Bitan, T., Meltzer, J. A., Graham, S. J., Leonard, C., & Rochon, E. (2018). Therapy-Induced Neuroplasticity in Chronic Aphasia After Phonological Component Analysis: A Matter of Intensity. *Frontiers in Neurology*, *9*, 225–7. https://doi.org/10.3389/fneur.2018.00225

Moulton, C.-A. E., Dubrowski, A., Macrae, H., Graham, B., Grober, E., & Reznick, R. (2006). Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial. *Annals of Surgery*, *244*(3), 400–409. https://doi.org/10.1097/01.sla.0000234808.85789.6a

Mozeiko, J., Coelho, C. A., & Myers, E. B. (2015). The role of intensity in constraint-induced language therapy for people with chronic aphasia. *Aphasiology*, *30*(4), 339–363. https://doi.org/10.1080/02687038.2015.1070949

Murray, E., Power, E., Togher, L., McCabe, P., Munro, N., & Smith, K. (2013). The reliability of methodological ratings for speechBITE using the PEDro-P scale. *International Journal of Language \& Communication Disorders / Royal College of Speech \& Language Therapists*, *48*(3), 297–306. https://doi.org/10.1111/1460-6984.12007

OCEBM Levels of Evidence Working Group. (2011). *The Oxford Levels of Evidence 2*. Oxford Centre for Evidence-Based Medicine. Retrieved from http://www.cebm.net/index.aspx?o=5653

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Page, S. J., & Wallace, S. E. (2014). Speech language pathologists' opinions of constraint-induced language therapy. *Topics in Stroke Rehabilitation*, *21*(4), 332–338. https://doi.org/10.1310/tsr2104-332

Page, S. J., Levine, P., Sisto, S., Bond, Q., & Johnston, M. V. (2002). Stroke patients" and therapists" opinions of constraint-induced movement therapy. *Clinical Rehabilitation*, *16*(1), 55–60. https://doi.org/10.1191/0269215502cr473oa

Palmer, R., Witts, H., & Chater, T. (2018). What speech and language therapy do community dwelling stroke survivors with aphasia receive in the UK? *PloS One*, *13*(7), e0200096. https://doi.org/10.1371/journal.pone.0200096

Pulvermüller, F., & Berthier, M. L. (2008). Aphasia therapy on a neuroscience basis. *Aphasiology*, *22*(6), 563–599. https://doi.org/10.1080/02687030701612213

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, *32*(7), 1621–1626. https://doi.org/10.1161/01.STR.32.7.1621

Ramsberger, G., & Marie, B. (2007). Self-administered cued naming therapy: a single-participant investigation of a computer-based therapy program replicated in four cases. *American Journal of Speech-Language Pathology*, *16*(4), 343–358. https://doi.org/10.1044/1058-0360(2007/038)

Raymer, A. M., Kohen, F. P., & Saffell, D. (2006). Computerised training for impairments of word comprehension and retrieval in aphasia. *Aphasiology*, *20*(2-4), 257–268. https://doi.org/10.1080/02687030500473312

Robey, R. R. (1998). A Meta-Analysis of Clinical Outcomes in the Treatment of Aphasia. *Journal of Speech, Language, and Hearing Research*, *41*(1), 172–16. https://doi.org/10.1044/jslhr.4101.172

Rochon, E., Laird, L., Simic, T., Höbler, F., Marcotte, K., Bitan, T., et al. (2016). The Effect of Treatment Intensity on the Efficacy of Therapy for Anomia in Aphasia. Presented at the Canadian Stroke Congress, Quebec City, Canada. https://doi.org/10.1177/1747493016659793

Rose, M. L., Ferguson, A., Power, E., Togher, L., & Worrall, L. E. (2013). Aphasia rehabilitation in Australia: Current practices, challenges and future directions. *International Journal of Speech-Language Pathology*, *16*(2), 1–12. https://doi.org/10.3109/17549507.2013.794474

Sage, K., Snell, C., & Lambon Ralph, M. A. (2011). How intensive does anomia therapy for people with aphasia need to be? *Neuropsychological Rehabilitation*, *21*(1), 26–41. https://doi.org/10.1080/09602011.2010.528966

Schmitt, M. B., & Justice, L. M. (2012). Optimal intervention intensity for emergent literacy: What we know and need to learn. *International Journal of Speech-Language Pathology*, *14*(5), 451–455.

Stahl, B., Mohr, B., Büscher, V., Dreyer, F. R., Lucchese, G., & Pulvermüller, F. (2018). Efficacy of intensive aphasia therapy in patients with chronic stroke: a randomised controlled trial. *Journal of Neurology, Neurosurgery & Psychiatry*, *89*(6), 586–592. https://doi.org/10.1136/jnnp-2017-315962

Stroke Engine. (n.d.). PEDro score - Stroke Engine. Retrieved August 7, 2019, from https://www.strokengine.ca/en/glossary/pedro-score/

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: the 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation*, *23*(5), 619–638. https://doi.org/10.1080/09602011.2013.824383

Tate, R. L., Rosenkoetter, U., Wakim, D., Sigmundsdottir, L., Doubleday, J., Togher, L., et al. (2015). The Risk of Bias in N-of-1 Trials (RoBiNT) Scale: An expanded manual for the critical appraisal of single-case reports. Sydney, Australia: Author.

The Cochrane Collaboration. (2014). Review Manager (RevMan) [Computer Software]. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.

Togher, L. (2012). Challenges inherent in optimizing speech-language pathology outcomes: It's not just about counting the hours. *International Journal of Speech-Language Pathology*, *14*(5), 438–442. https://doi.org/10.3109/17549507.2012.689334

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.
https://doi.org/10.1080/17549500902726059

Wallace, S. J., Worrall, L. E., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., et al. (2016). Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disability and Rehabilitation*, *91*, 1–16. https://doi.org/10.1080/09638288.2016.1194899

Warren, S. F., Fey, M. E., & Yoder, P. J. (2007). Differential treatment intensity research: a missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(1), 70–77. https://doi.org/10.1002/mrdd.20139

Chapter 5 – Randomised Controlled Trial

Preface

The systematic review in Chapter Two demonstrated that there is limited high quality evidence on constraint or multimodal treatments, and that existing studies do not sufficiently explore outcomes beyond the level of impairment. It concluded with the need for a methodologically rigorous comparison of the treatment approaches. The systematic review of higher versus lower intensity aphasia therapy presented in Chapter Four showed that, despite plausible theoretical arguments for different schedules, data are not yet clear on which is superior. Given the limited feasibility of high intensity treatment in existing service models, establishing the efficacy of lower intensity treatments is important. Consequently, the following Chapter describes the method and results of a trial of CIAT Plus and M-MAT provided at a low-moderate intensity. Outcomes were measured at impairment, activity/participation and quality of life levels. Results of this trial can, in future, be compared to results of a large, powered trial of CIAT Plus and M-MAT conducted at high intensity. Such a comparison will provide preliminary evidence on the effect of altering intensity.

Study Design

Main trial

This study formed a nested substudy within the larger COMPARE clinical trial which was funded by the National Health and Medical Research Council, Australia (#1083010). COMPARE was a multi-site, single-blinded, randomised controlled trial with three arms — M-MAT, CIAT Plus and Usual Care (UC). Usual Care comprised care typically received by participants; for example, community aphasia groups or private therapy, or most often no intervention. Participants were placed in groups of two or three according to aphasia severity (mild, moderate, severe) and each group was then electronically randomised to one of the three arms. COMPARE examined CIAT Plus against M-MAT and both CIAT and M-MAT to the control arm, Usual Care (UC). CIAT Plus and M-MAT were both provided at a high intensity of 30 hours over two weeks (three hours per weekday). COMPARE is registered at The Australian and New Zealand Clinical Trials Register (ACTRN12615000618550) and the protocol has been published (Rose et al., 2019).

This study

This study, a sub-study of the main COMPARE randomised controlled trial, was a multisite, single-blinded, randomised controlled trial with two treatment arms — lower intensity M-MAT and lower intensity CIAT Plus. Following completion of their final Follow-Up assessment in the main trial, participants in the Usual Care arm were offered the opportunity to take part in lower intensity therapy (i.e., 30 hours of treatment over 5 weeks) as part of the substudy. Participants who consented to the substudy were re-randomised, in groups of two or three stratified by aphasia severity (based on Western Aphasia Battery - Revised Aphasia Quotient scores – WAB-R AQ), into the CIAT Plus or M-

MAT arm. Blinding of participants and therapists in behavioural interventions is not possible

but assessors were blinded to participant allocation. Figure 5-1 provides a schematic

overview of the trial process.

Figure 5-1 – Study timeline and relation to COMPARE main trial
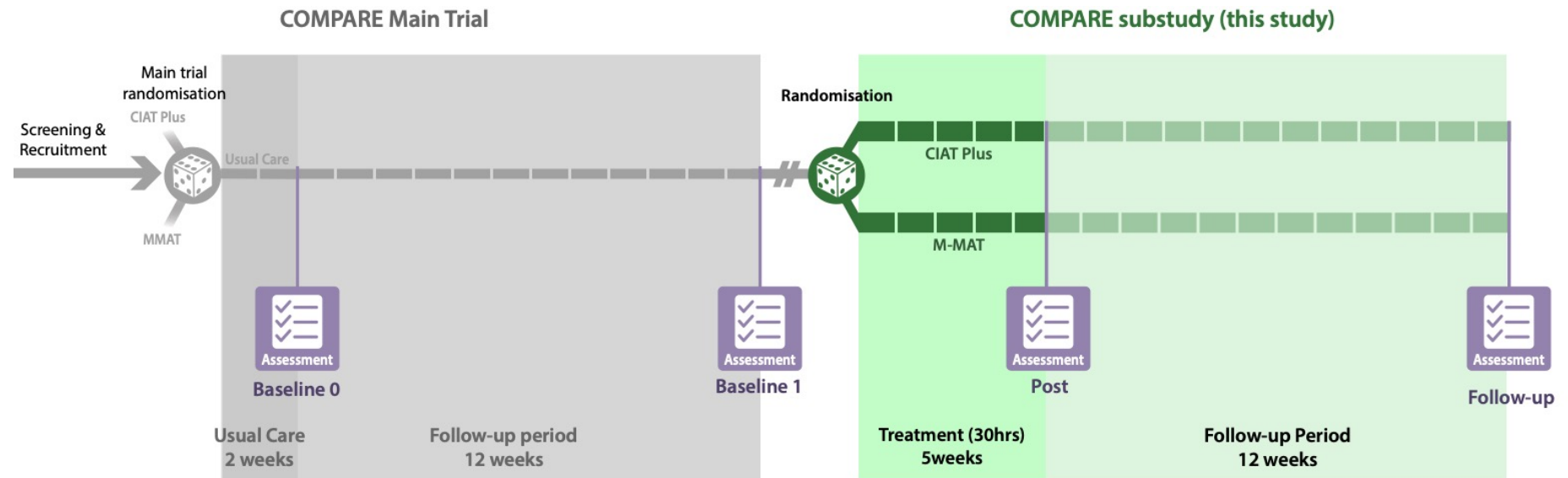
Aims and Hypotheses

Main Aim

There were two main aims of this research:

1. To demonstrate the efficacy of lower-intensity CIAT Plus and M-MAT at immediate and long term timepoints (12 weeks)

2. To compare the efficacy of CIAT Plus to M-MAT in a lower-intensity schedule

The independent variable was Treatment type (M-MAT or CIAT Plus), and the main dependent variable and primary outcome was aphasia severity (WAB-R AQ), with the following additional dependent variables as secondary outcomes:

Impairment level

- Confrontational naming (COMPARE naming battery)

- Connected Speech (Content Information Unit (CIU) count, CIUs per minute) (Nicholas & Brookshire, 1993)

Activity/Participation

- Multimodal communication (Scenario Test) (van der Meulen et al., 2010)

- Carer-rated communication (Communicative Effectiveness Index - CETI) (Lomas et al., 1989)

Quality of life

- Quality of life (Stroke and Aphasia Quality of Life Scale 39g - SAQOL-39) (Hilari et al., 2009)

Supplementary aim

The supplementary aim of this research was to explore the relationship between treatment response and baseline characteristics. For this analysis, the dependent variable was aphasia severity improvement (WAB-R AQ and naming battery) while the independent variables were:

- Baseline aphasia severity (WAB-R AQ) (Kertesz, 2007)

- Baseline confrontational naming (Naming Battery)

- Age

- Months post onset of aphasia

- Semantic impairment (Pyramids and Palm Trees test) (Howard & Patterson, 1992)

- Phonological impairment (phonemic errors on Naming Battery)

- Attention (Test of Everyday Attention - Elevator Counting) (Chen et al., 2013)

- Cognitive Flexibility (Test of Everyday Attention - Visual Elevator)

- Non-verbal reasoning (Raven's Progressive Matrices) (Basso et al., 1987)

- Auditory-verbal short-term memory (Picture Span Test - forwards) (DeDe et al., 2014)

- Auditory-verbal working memory (Picture Span Test - reverse)

Hypotheses

Main Aims

The following hypotheses were generated for this research for the two main

aims:

Primary Outcomes

HYPOTHESIS 1. CONSTRAINT VS MULTIMODAL

There will be no difference in aphasia severity improvement between lower-

intensity M-MAT and CIAT at immediately post treatment or 12-week Follow-up.

Rationale: A comparative systematic review with meta-analysis of the

treatment approaches did not find either to be superior based on limited evidence to

date (Pierce et al., 2017).

HYPOTHESIS 2. BASELINE TO POST INTERVENTION IMPROVEMENT

Compared to baselines, lower-intensity CIAT Plus and lower-intensity M-MAT

will each result in significantly reduced aphasia severity on the WAB-R AQ

immediately post treatment.

Rationale: In previous work, significant improvements on aphasia battery

tests have been shown in both CIAT (Meinzer et al., 2007; Pulvermüller et al., 2001)

and M-MAT (Attard et al., 2013; Rose et al., 2013) though changes were small or

inconsistent within some papers (Attard et al., 2013; Rose et al., 2013; Wilssens et

al., 2015).

HYPOTHESIS 3. BASELINE TO FOLLOW-UP IMPROVEMENT

Lower-intensity CIAT and lower-intensity M-MAT will each result in a long-term reduction in aphasia severity determined by significant improvement on the WAB-R AQ at 12-week Follow-Up compared to baselines.

Rationale: Few studies of CIAT or M-MAT have explored follow-up outcomes at a 12 week delay (Menahemi-Falkov et al., 2020), however one Phase I study found significantly improved WAB-R AQ scores at 12 weeks compared to baseline (+4.22, *p* = .002) after a crossover 48-hour dose of CIAT Plus and M-MAT (Rose et al., 2013).

HYPOTHESIS 4. MAINTENANCE OF GAINS

There will be no significant reduction in scores on the WAB-R AQ at 12-week Follow-Up compared to Post Intervention for either CIAT or M-MAT.

Rationale: As in Hypothesis 3, there is minimal investigation of outcomes more than one month post intervention for CIAT and M-MAT but one study reported no significant reduction in WAB-R AQ (-0.08, *p* = .414) between post-intervention and 12 week Follow-up (Rose et al., 2013).

Secondary outcomes

HYPOTHESIS 5. SECONDARY OUTCOMES

Compared to baseline, lower-intensity CIAT and lower-intensity M-MAT will result in significantly improved scores immediately post treatment and at 12-week Follow-Up for:

    a. Naming battery

    b. Connected speech (CIU Count and CIUs/min)

    c. CETI (activity/participation)

d. Scenario test (activity/participation)

e. SAQOL-39 (quality of life)

Rationales: A meta-analysis of constraint and multimodal interventions showed significant effects on confrontational naming (Pierce et al., 2017). The same review found that CETI scores typically show positive changes, although these outcomes were not meta-analysed. A phase I crossover study of M-MAT and CIAT Plus found improved scores on the SAQOL-39, with mixed results on the Scenario Test (Rose et al., 2013). An analysis of discourse-level outcomes from the same experiment found mixed but broadly positive results (Rose, Mok, et al., 2015)

Supplementary Aim

The following hypothesis was generated for the supplementary aim:

HYPOTHESIS 6: TREATMENT RESPONSE AND BASELINE CHARACTERISTICS

Baseline to Post Intervention response (WAB-R AQ change and Naming Battery change) for both treatment groups (M-MAT, CIAT) will be correlated with:

- baseline severity (WAB-R AQ and Naming Battery)

- scores on semantic impairment (Pyramids and Palm Trees)

- phonological impairment (Phonological errors on Naming Battery)

- non-verbal reasoning (Raven's Progressive Matrices)

- cognitive flexibility (Test of Everyday Attention - Visual Elevator)

- auditory-verbal short-term memory (Picture Span Test - Forwards)

- auditory-verbal working memory (Picture Span Test - Reverse)

- attention (Test of Everyday Attention - elevator counting)

- age

- months post onset

Rationale: Each of the domains hypothesised to predict treatment response have shown potential predictive value in impairment outcomes for aphasia therapy including semantic and phonological processing (Dignam et al., 2015; Lambon-Ralph et al., 2010), nonverbal reasoning (van de Sandt-Koenderman et al., 2008), cognitive flexibility (Lambon-Ralph et al., 2010), auditory-verbal memory (van de Sandt-Koenderman et al., 2008), attention (Lambon-Ralph et al., 2010), and age (Lahiri et al., 2020). Although time post onset has been found to predict response in the first six months compared to >6 months (RELEASE Collaboration, 2021), there has been no exploration of months post onset within this chronic phase.

Ethical approval

Written ethical approval was obtained by the La Trobe Human Research Ethics Committee (15-043). Approval was also obtained from the following committees:

- Alfred Health

- Austin Health

- East Metropolitan Health Service

- Eastern Health

- Gold Coast Health Human Research Ethics Committee

- Hunter New England Research Ethics & Governance

- Metro South Health Research Governance

- Monash Health Human Research Ethics Committee

- North Metropolitan Health Service

- Northern Sydney Local Health District

- Peninsula Health

- Royal Rehab

- South Eastern Sydney Local Health District

- South Metropolitan Health Service

- St Vincent's Hospital Melbourne

- St Vincent's Hospital Sydney Research Office

- Western Sydney Local Health District

Sample

For the main COMPARE trial, 216 participants were recruited. A sample size of 198 was derived statistically based on a power calculation to achieve 80% power at a level of 5% significance. An additional 18 participants were required due to the anticipated Design Effect of cluster randomisation (design effect 1.08 based on estimated Intraclass coefficient of 0.04).

The sample for this study was recruited from the Usual Care arm (UC) of the main COMPARE study which contained 70 Usual Care (control) participants. CIAT Plus and M-MAT low intensity arms could have comprised up to 35 people in each arm. However, the sample size of n = 70 represented a maximum possible number because participation in this sub-study was independent of participation in the main COMPARE trial and some participants declined later involvement. Where groups of three could not be formed, additional group members were sourced from people with aphasia who had previously completed the treatment. Such "replacement" participants did not complete assessment or contribute data to the trial.

Inclusion and exclusion criteria did not differ from the main COMPARE trial, which were as follows:

Inclusion Criteria

Participants:

1. Had a documented stroke resulting in aphasia at least 6 months prior to assessment at time of consent,

2. Had aphasia of any type (<93.8 WAB-R AQ),

3. Were fluent in English prior to stroke,

4. Were 18 years or older and able to give informed consent,

5. Were independent in toileting,

6. Were able and willing to attend/participate in assessments, 5-week treatment period, Post Intervention period and Follow-Up period,

7. Had a carer/significant other who was able and willing to participate in baseline, Post Intervention and Follow-Up assessments.

Exclusion criteria

Participants must not:

1. Have a non-stroke neurological event/diagnosis (head injury, neurosurgery, dementia, epilepsy, progressive neurological disorder)

2. Have severe apraxia of speech or dysarthria as detected on the Apraxia of Speech Rating Scale (Strand et al., 2014)

3. Have a current diagnosis of major clinical depression or other mental health condition that may affect involvement or adherence to the study protocol

4. Have uncorrected sensory loss (hearing/vision) preventing participation in communication assessments and treatments

5. Be unable to attend the 12-week Follow-Up assessments

6. Have any other serious medical condition prior to their stroke including malignancies, psychiatric, behavioural or drug-dependency problems, which are likely to influence the participant's ability to cooperate or in the opinion of the study investigator would prevent adherence to the protocol (Rose et al., 2019).

7. Be participating in any other therapy (including alternative therapies and clinical trials) or taking medications (including herbal preparations) that are not considered to be standard care for people with aphasia.

These criteria were confirmed on the telephone and at a screening visit, using self and carer-reported medical history, the Stroke Aphasia Depression Questionnaire (for depression), the Apraxia of Speech Rating Scale (for Apraxia of Speech), and the Western Aphasia Battery. Table 5-1 lists screening tools used.

Recruitment Sources

Potential participants were sought from a range of sources within Australia and New Zealand.

Health services

COMPARE had participating health services from across Australia and following site-specific ethics approval, medical records and databases were searched by local staff for potential participants and contacted with details of the trial.

Associations

The COMPARE trial was advertised by relevant associations and groups on their websites and social media. Participating organisations included the Australian Aphasia Association, the Stroke Foundation, http://www.aphasia.community/, Speech Pathology Australia, Speech Pathology in Email Chats, the Talkback Association (South Australia), Aphasia New Zealand and the New Zealand Speech Language Therapy Association.

General publicity

The COMPARE trial was publicised throughout the recruitment phase at relevant conferences, professional development events, through online groups, social media and via word of mouth. A webpage with details of the trial was set up for interested clinicians or people with aphasia.

Assessment and Intervention Timeline

Figure 5-1 illustrates how this study relates to the COMPARE main trial in terms of timing and assessments. The COMPARE trial (main study) spanned 14 weeks and this substudy followed, spanning an additional 17 weeks.

COMPARE main trial timeline

In the main trial, participants completed screening followed by two baseline assessment sessions. Participants' aphasia severity was determined based on WAB-R AQ (mild 93.7-66; moderate: 65-33; severe: 32-0). These were calculated by dividing the cutoff score indicating the presence of aphasia (<93.7) into three equal ranges as traditional severity ranges do not take into account that scores greater than 93.7 are within normal limits. Once there were sufficient participants of the same aphasia severity within a geographical area to form a group, the group was randomised into one of three treatment arms – CIAT Plus, M-MAT or UC. Randomisation was conducted via a computer algorithm from a central allocation system, using blocked randomisation within each severity. Groups then underwent a two week intervention phase. Groups randomised to UC continued with the care they were already receiving prior to the trial. For people living in the community with chronic aphasia, aphasia intervention is typically two hours or less per week (Verna et al., 2009). Post Intervention assessments were administered in the following week, and then following a twelve week Follow-up phase, a follow-up assessment was conducted. After their final main trial assessment, participants in the UC arm were invited to participate in the sub study. In many cases, not all participants from a UC group consented or were available for the substudy, resulting in gaps between completion of the main trial and randomisation into the substudy. The median gap was 8 weeks with the majority ranging from 6 to 18 weeks, save for one outlier at 95 weeks (see histogram in Appendix 4).

Substudy timeline

Two baselines were used for the present trial, allowing examination of stability of measures prior to intervention. If participants from UC consented to participate in the substudy, data from their main trial Post Intervention and Follow-up assessments doubled as Baseline 0 and Baseline 1 for the substudy, respectively. Some data from participants' initial main trial assessments were also utilised to avoid unnecessary reassessment, including data on demographics, handedness, stroke severity, apraxia of speech severity, attention and memory, and tests of eligibility (e.g., independence in toileting, medical history, depression).

Table 5-1 summarises when various assessments were conducted and their purpose, and each assessment is described in detail later in *Assessment Materials*.

Table 5-1 – Timing of Assessments and Purpose

Assessment	Purpose	Screening	Main trial Baselines x2	Baseline 0	Baseline 1	Post Intervention	12 Week Follow-up
Demographics	Participant description	X					
A Simplified Handedness Questionnaire	Participant description	X					
Modified Rankin Scale	Participant Description		X				
Medical History	Eligibility	X					
Stroke Aphasia Depression Questionnaire (SADQ-10)	Eligibility	X					
Apraxia of Speech Rating Scale (ASRS)	Eligibility	X					
Western Aphasia Battery – Aphasia Quotient (WAB-R AQ)	Eligibility **Outcome measure**	X		X	X	X	X
COMPARE Naming Battery Part 1 (All 180 items)	**Outcome measure**	X		X	X	X	X
Stroke and Aphasia Quality of Life Scale 39 (SAQOL-39)	**Outcome measure**			X	X	X	X
Connected speech measures	**Outcome measure**			X	X	X	X
Scenario Test	**Outcome measure**			X	X	X	X
Communicative Effectiveness Index (CETI)	**Outcome measure**			X	X	X	X
Pyramids & Palm Trees	Participant characteristic		X				
Test of Everyday Attention (TEA) subtests: Elevator counting; Visual elevator	Participant characteristic		X				
Picture Span Verbal Memory Test	Participant characteristic		X				
Raven's Progressive Matrices	Participant characteristic		X				

Baseline 0 and Baseline 1 Assessments

The two Baseline assessment included:

- Western Aphasia Battery Revised (WAB-R)

- Communicative Effectiveness Index (CETI)

- Naming Battery Part 1 (All items)

- Connected Speech Measures

- Scenario Test

- Stroke and Aphasia Quality of Life Scale 39 (SAQOL-39)

Randomisation

Using the same randomisation system as the main trial, participants agreeing to participate in the current study were re-randomised as groups into CIAT Plus or M-MAT. Simple randomisation was used, meaning imbalances in the number of groups in each arm was possible. Trial managers were notified of the randomisation result for each group via telephone using an external central allocation system.

Treatment phase

Participants attended two hours of treatment per day, three days a week for five weeks. The first intervention session (1 hour) commenced at 10am, followed by a 15-minute break and the second intervention session. The therapist then assigned a new home practice task for each participant. The total treatment time was 30 hours plus home practice tasks, to allow future comparison of this lower-intensity trial to the main, high intensity COMPARE trial. As the total hours were fixed, the

frequency of treatment per week was altered to produce a lower-intensity schedule.

The schedule of six hours per week was chosen for a number of reasons:

- Six hours per week falls between 'usual care' (<2 hours per week) and the high intensity of 15 hours per week (main trial)

- It allowed for three two-hour treatment days which was less than the three-hour treatment days provided in the main COMPARE trial

- Six hours per week doses have precedence in chronic aphasia treatment. Dignam et al. (2015) provided an Intensive Comprehensive Aphasia Program for chronic aphasia at six hours per week over eight weeks. Results were statistically significant at post intervention on the Boston Naming Test, the CETI, the Communication Confidence Rating Scale for Aphasia and the Assessment of Living with Aphasia.

Post Intervention and 12-week Follow-Up

The two assessment sessions (at Post Intervention and Follow-Up) following the treatment phase and follow-up phase mirrored the Post Intervention assessments of the main trial in both content and timing.

Assessment fidelity

All assessment and intervention sessions were video recorded with the participant's face in view. This allowed confirmation of adherence to the study protocol, off-line scoring of complex communication behaviours, and inter-rater reliability of assessment scoring. Each assessor's first administration of the Spontaneous Speech subtest of the WAB-R was reviewed by a fidelity monitor

employed by the trial to ensure reliability. If the monitor and assessor scores differed by >2 points, the assessor was provided with feedback and their subsequent assessment was also reviewed. In addition, all WAB-R AQ results where the score was ±5 points of a severity or eligibility cut-off (32, 65 and 93.7) were reviewed and re-scored in their entirety.

Blinding of assessors

Assessors were blinded to participant allocation. Assessors were not employed as therapists. All communication for the trial occurred via the trial managers and contact between therapists and assessors was not permitted. Assessments took place at participants' homes and assessors were provided with training to remind participants not to discuss any aspects of therapy or usual care received. Following each assessment visit, assessors completed a *Blinded Assessor Questionnaire* which ask the assessor whether they had been unblinded and whether they could guess which arm the participant was in.

Assessment Materials

The rationales for selection of assessments and their purpose are described further below.

Demographics form

The assessor collected information about participants including age, gender and time post stroke onset. These were reported as baseline characteristics.

A Simplified Handedness Questionnaire

This questionnaire confirmed the self-report of participants of handedness on a scale. Scores range from -1.00 (extreme left handed) to 1.00 (extreme right handed) (Bryden, 1982). Handedness was reported as a baseline characteristic.

Medical History form

A thorough medical history was taken of previous and ongoing conditions to confirm eligibility for the trial and to provide baseline information in the case of an adverse event.

Stroke Aphasia Depression Questionnaire (SADQ-10)

The SADQ-10 (Sutcliffe & Lincoln, 1998) is a depression screen developed for people with aphasia living in the community. It is completed by a carer or family member who has been able to observe the person with aphasia over the previous week. There are ten items in the SADQ-10 which are rated from Often to Never and the total score may range from 0 to 30. A score of 14 or higher is indicative of depression. The SADQ-10 has high internal consistency, sensitivity and specificity for detecting depression (Bennett et al., 2006). The SADQ-10 was used to screen for any active depressive disorders, which were an exclusion criterion for this trial.

Apraxia of Speech Rating Scale (ASRS)

The ASRS is a tool to differentially diagnose apraxia of speech against dysarthria and aphasia. It also provides information on severity and apraxia characteristics. Assessors reviewed video recordings of conversation, picture description, word and sentence repetition and rapid speech movements and rated the speaker on up to 16 items, scoring each from 0-4. People whose score indicated

severe apraxia or severe dysarthria were not eligible for this trial. Excellent validity and inter-rater (0.94) and intra-rater (0.91-0.98) reliability have been demonstrated for the ASRS (Strand et al., 2014).

Western Aphasia Battery - Revised

The Western Aphasia Battery - Revised (Kertesz, 2007) is an assessment of language impairment for aphasia. It is one of the most widely used assessments for aphasia (Hula et al., 2010) and was recommended as a core outcome for aphasia in a recent consensus project involving patients, carers, clinicians and researchers (Wallace et al., 2019). The WAB has a test-rest reliability of >0.90 (Shewan & Kertesz, 1980). The Aphasia Quotient (WAB-R AQ) is a global score of aphasia severity calculated from the four subscales: spontaneous speech, auditory comprehension, repetition and naming. A change of five points on the WAB-R AQ is considered a clinically meaningful change (Hula et al., 2010). Part 2 of the WAB-R includes Raven's Progressive Matrices, a test of non-verbal reasoning that was used as a baseline measure of cognitive ability and investigated as a treatment response predictor.

COMPARE Naming Battery

A naming battery was developed for the COMPARE trial with 100 commonly used nouns and 80 verbs, selected according to published data on word frequency, age of acquisition, syllable length and complexity. Each card had a photograph of the target and was the same size as a standard playing card. Name agreement was tested in non-aphasic individuals. There were also ten adjective cards which served as visual prompts for linguistic levels requiring adjectives. For example, with the

sentence, "I have a [adjective] [target]" where the target was "boot," the "yellow" adjective card might be used.

Participants were tested to examine confrontational naming ability by showing each card and allowing ten seconds to name the target. No cues were provided. Self-corrected responses within the ten second window were marked as correct. Other responses were recorded as no response, semantic error, phonological error or other error. A small number of alternative responses were marked as correct (e.g., jumper *or* sweater), if they were responses given during testing of non-aphasic individuals. The number of phonological errors made by participants on the naming battery was used to indicate the intactness of their phonology systems and was examined in relation to treatment response.

Initial scores on the naming battery at baseline were used to assign participants to one of three difficulty levels for treatment (Hard, Moderate, Easy), calculated using the following rules:

- If score on Easy items ≤40, prescribe Easy list
- If score on Easy items ≥40 and on Moderate items ≤40, prescribe Moderate list
- If score ≥40 for both Easy and Moderate, prescribe Hard list

Each difficulty level comprised a selection of 80 cards from the set of 180 – two sets of 16 verbs (32 in total) and three sets of 16 nouns (48 in total). The 80 "treated items" list therefore differed between groups according to assigned difficulty level, although there was some overlap in the items that occurred in each list.

Stroke and Aphasia Quality of Life Scale 39g (SAQOL-39)

The SAQOL-39 is a measure of health-related quality of life following stroke, specifically designed to be accessible for people with aphasia. The scale is conducted via interview with the participant and produces quality of life scores in three domains - Physical, Communication and Psychosocial. It has high internal consistency ($\propto$ = 0.95), test-retest reliability (0.96), convergent validity (r = 0.36–0.70), discriminant validity construct validity (r = 0.26) and acceptability for people with aphasia (Hilari et al., 2009).

Connected speech measures

Samples of connected speech were taken using a picture from the WAB-R and two pictures from Nicholas and Brookshire (1993). The WAB-R "Picnic scene" picture shows a range of people engaged in recreational activities by a lake. The pictures from Nicholas and Brookshire show (1) a birthday party where the family and guests are reacting to the dog having eaten the cake and (2) a scene where a man and a cat are both stuck in a tree, with firemen, a girl and a dog below the tree. The picture description tasks have good reliability and validity (MacWhinney et al., 2010) and measures used were the total count of Correct Information Units (CIUs) and CIUs per minute. A CIU is an utterance that is relevant and informative about the stimulus, but not necessarily grammatically accurate (Nicholas & Brookshire, 1993). Audio was recorded during assessments and transcribed by a third-party service. Speech pathology researchers and trained speech pathology students then analysed using SALT software (Miller & Iglesias, 2008) with manual correction and CIU identification.

Scenario Test

The Scenario Test measures the ability of people with aphasia to communicate using any verbal or non-verbal modalities in six social scenarios. Subjects are encouraged to indicate how they would communicate in the scenario using any suitable modality. Each item is rated from 0-3 based on success of informational exchange. The summary score can range from 0 to 54. The Scenario Test has very good psychometric properties including test-retest reliability (0.98), inter-rater reliability (0.86 – 1.00), convergent validity (0.5-0.85), and sensitivity to change (van der Meulen et al., 2010).

Communicative Effectiveness Index (CETI)

The CETI is a rating scale with 16 items which measure functional communication and participation in communicative activities. A carer or other family member marks visual analogue scales to indicate the ability of the person with aphasia to communicate successfully in each item. The ends of the scale range from "Not at all able" to "As able as before stroke". Overall score ranges from 0 to 100 and a change greater than 12 is considered the minimal clinically significant change. Test-retest reliability (ICC = 0.94), inter-rater reliability (ICC = 0.73) and validity have also been demonstrated (Lomas et al., 1989).

Pyramids & Palm Trees Test - 3 pictures version

The Pyramids & Palm Trees test (Howard & Patterson, 1992) is an assessment of semantic impairment. The participant is presented with a stimulus (e.g., pyramid) and two alternatives (e.g., palm tree and pine tree) and asked to choose the one most closely associated. The test can be administered using pictures or written

words. In this trial, the picture version was used to determine the intactness of participants' non-verbal semantic system for description of baseline characteristics and investigation of treatment response predictors. Scores range from 0 to 52 with a normal cut off at 49.

Test of Everyday Attention (TEA)

Two subtests of the TEA were used in this study, 1) Elevator Counting and 2) Visual Elevator. These are designed to assess sustained attention (1) and attention switching and cognitive flexibility (2). In the Elevator Counting task, participants are asked to count seven sequences of unevenly spaced auditory tones, ranging from 3 to 14 total tones in each sequence. The number of sequences correctly counted is scored (0-7). In the Visual Elevator task, participants attempt to track the "floor number" through increases and decreases as indicated by a sequence of up and down icons. Scoring is based on the number of correct responses (0-10). Results formed part of baseline description and the response prediction analysis. The TEA has acceptable test-retest reliability in chronic stroke patients (Chen et al., 2013).

Picture Span Verbal Memory Test

The Picture Span Verbal Memory Test examines both short-term and working auditory-verbal memory. A series of words (starting at two, increasing to six) are read aloud by the examiner and the participant points out the corresponding images from a page in order. For working memory, the participant must point out the images in reverse order. Research has found acceptable test-retest reliability (0.88-0.90), internal consistency (0.79-0.85), and construct validity (0.75-0.83) and the test was validated in controls and people with aphasia (DeDe et al., 2014). Auditory-

verbal memory was investigated as a predictor of treatment response because in both CIAT Plus and M-MAT, participants are guided to produce and/or repeat utterances of increasing length. The total number of correct items (out of 100) for forwards and for reverse were used in analysis.

Raven's Progressive Coloured Matrices

Raven's Progressive Coloured Matrices is a test of non-verbal reasoning. Subjects are asked to choose the missing coloured tile of a sequence from an array of six options. There are 36 items in the assessment which are progressively more difficult. One point is awarded for each correct response and an additional 1-point is awarded for completion in less than five minutes. The test has good internal consistency (0.85), and test-retest reliability ($\approx$0.8-0.9) (Burke, 1958) and is commonly used as an indicator of non-verbal reasoning in aphasia research. Normative data is available (Basso et al., 1987).

Modified Rankin Scale (MRS)

The Modified Rankin Scale is the most widely used measure of disability for people following stroke (Wilson et al., 2005) and has strong evidence of validity and reliability (Banks & Marotta, 2007). Individuals are ranked following a brief interview, with 0 signifying no symptoms and 5 signifying severe disability to that extent that the individual requires constant nursing care. The MRS was used at baseline for description of participant characteristics.

Fatigue and Distress Visual Analogue Scales

Before and after every assessment and intervention contact, a fatigue and distress measure was used to monitor participants' subjective fatigue and distress.

Two visual analogue scales were used. One end of the scale contains No fatigue or No distress and the other, Extreme fatigue or Extreme distress.

Intervention protocol

CIAT Plus protocol

CIAT Plus is a variant of the original constraint treatment for aphasia, CIAT (Pulvermüller et al., 2001). Language is stimulated and spoken communication practised via interactive group card games which use picture cards. There were six card games which all require participants to name or request cards in order to play - Fish, Bingo, Memory, Snap, Who Am I, and I Went Shopping. Three levels of cards (easy, moderate, hard) were available depending on the group's severity allocation, each with two sets of 16 verbs and three sets of 16 nouns. Thus, there were 80 items (48 nouns; 32 verbs) treated in each group, taken from the 180-item naming battery. Therapists alternated between nouns and verbs each hour. When at least 2/3 participants reached 80% accurate production on a particular noun or verb set or nine hours of treatment for that set had been provided, the next set of words was used.

Participant responses were shaped progressively through six linguistic levels. The lowest level requires a single word while the highest level requires a complex sentence with elements including subordinate clauses, adjectives, and prepositions. Table 5-2 details the linguistic levels for nouns and verbs.

Barriers were placed between participants to discourage the use of gestures or other non-verbal modalities to communicate, apart from games that required shared space; namely Snap, I Went Shopping and Memory. While participants were

not specifically discouraged from using other modalities (e.g., gesture, finger writing) to self-cue, they were reminded not to use these to communicate with other participants. Pens and writing paper were not available during treatment.

In CIAT Plus, when the participant was unable to produce the required utterance within approximately ten seconds, the therapist provided a phonemic cue for the target word. If the phonemic cue was effective, the participant was asked to repeat the target word three times. The phonemic cue step, effective or not, was always followed by the therapist showing the printed target word to the participant and reading it aloud. The participant was then asked to repeat the entire utterance (carrier phrase/sentence including the target word) three times. No other cueing was provided for word retrieval. Errors in the carrier phrase were prompted verbally (e.g., "You forgot to say *who* was dancing") and if necessary, a written sentence frame was used to demonstrate the desired sentence elements and structure (e.g., "I have: The [subject] is/was [target] [verb] the/a [object] ").

Table 5-2 – Target syntax for each linguistic level

	Level	Grammatical form	Example
Nouns	1	Noun	"Couch?"
	2	Carrier phrase + noun	"Do you have (a) couch?" or "I have (a) couch"
	3	Carrier phrase + adj + noun	"Do you have (a) red couch?" or "I have (a) red couch"
	4	Carrier phrase + adj + adj + noun	"Do you have (a) large, red couch?" or "I have (a) large, red couch"
	5	Carrier phrase + sub + to be + prep + adj + adj + noun	"Do you have (the) girl (is) on the large, red couch?" or "I have (the) girl (is) on the large, red couch"
	6	Carrier phrase + sub + verb + prep + adj + adj + noun	"Do you have (the) girl (is) sitting on the large, red couch?" or "I have (the) girl (is) sitting on the large, red couch"
Verbs	1	Verb	"Sweeping?"
	2	Carrier phrase + verb	"Do you have sweeping?" or "I have sweeping" "Do you have bouncing" or "I have bouncing"
	3	Carrier phrase + sub + verb	"Do you have (the) girl is sweeping?" or "I have (the) girl is sweeping" "Do you have (the) boy is bouncing?" or "I have (the) boy is bouncing?"
	4	Carrier phrase + sub + verb + object	"Do you have (the) girl is sweeping the floor?" or "I have (the) girl is sweeping the floor"
	5	Carrier phrase + sub + verb + object + prep phrase	"Do you have (the) girl is sweeping the floor with the broom?" or "I have (the) girl is sweeping the floor with the broom"
	6	Carrier phrase + sub + verb + object + prep phrase + conjunction + SVAdj	"Do you have (the) girl is sweeping the floor with the broom because it is dirty?" or "I have (the) girl is sweeping the floor with the broom because it is dirty"

Adj = adjective; sub = subject; prep = preposition; SVAdj = Subject, Verb, Adjectival phrase

CIAT Plus builds upon CIAT by assigning home tasks to improve carryover of language skills into real life. The therapist assigned each participant a 15-minute home task at the end of each intervention day. Tasks aimed to target the vocabulary and linguistic levels treated that day and were incorporated into the participants' existing plans for that day. For example, the participant may have been going to a supermarket or post office on the way home and the task may have included asking for items, requesting information about postage, etc. If the participant was going straight home the task might have involved telephoning or videoconferencing with a relative or friend or discussing a news item/social plan with their family member. Task assignments were recorded on a written log and sent home with the participant, and the outcomes discussed at the following intervention session.

During the intervention period participants were permitted to continue with their standard care (including speech therapy, physiotherapy, groups, etc.), although participating in the trial may have impacted participants' fatigue levels and time available for standard care. However, any treatments that were not considered standard care (as judged by trial coordinators) were not permitted, for example, alternative therapy, herbal preparations, or other clinical trials. Standard care was recorded in a diary provided by COMPARE during the interval between baselines. Diaries were not continued beyond the baseline period so as not to increase the burden of participation – by Baseline 1, participants or their carers had recorded 16 weeks of standard care.

M-MAT protocol

M-MAT uses the same structure as CIAT Plus in terms of card games, stimulus cards, and linguistic levels. Participants were also assigned home practice tasks as described in CIAT Plus. However, in contrast, M-MAT did not use visual barriers during any games. The cueing hierarchy also differed in M-MAT. If the participant was unable to produce a target utterance, the therapist prompted them to produce an iconic gesture to see whether they could successfully self-cue. If still unsuccessful, the participant (a) copied the gesture modelled by the therapist, (b) drew the target word and (c) read aloud and then copied out the written word. The target word was repeated aloud once with each of these steps. Following these cueing steps, the participant was then asked to repeat the entire utterance (carrier phrase/sentence including the target word) three times.

Table 5-3 - Cueing hierarchies for CIAT Plus and M-MAT

Step	Description	Correct	Incorrect
	CIAT Plus		
1	Participant names item and carrier phrase (e.g., "Couch"? or "Do you have a couch?")	✓ Move to next participant's turn	✗ Go to step 2
2	COMPARE Therapist provides phonemic cue (e.g., "It starts with /k/")	✓ Participant repeats the name three times (e.g., "couch, couch, couch") Go to step 3	✗ If the participant makes an error on the target word, do not request repetition; go to step 3

3	COMPARE Therapist provides orthographic cue (e.g., the written word "couch") + verbal name of item (e.g., "It's a couch....can you say couch?")	✓ Participant repeats the carrier phrase + name three times while the written word is in view (e.g., "Do you have couch, do you have couch, do you have couch") Move to next participant's turn	✖ If participant cannot repeat the word, acknowledge the difficulty, offer encouragement and move to the next participant's turn

M-MAT			
1	Participant names item and carrier phrase (e.g., "Couch"? or "Do you have a couch?")	✓ Move to next participant's turn	✖ Go to step 2
2	COMPARE Therapist asks participant to make an iconic gesture and name the item	✓ Repeat carrier phrase + name three times. Move to next participant's turn	✖ Go to step 3
3	COMPARE Therapist provides iconic gesture model and verbal name of item, and asks participant to copy gesture and repeat name of item	✓ Go to step 4	✖ COMPARE Therapist provides gesture refinement cues Go to step 4
4	COMPARE Therapists asks participant to draw item and name the item	✓ Go to step 5	✖ COMPARE Therapist provides drawing refinement cues Go to step 5
5	COMPARE Therapist asks participant to read, write and name the item	✓ Go to step 6	✖ COMPARE Therapist provides the item name in writing for participant to copy Go to step 6
6	COMPARE Therapist asks participant to repeat the carrier phrase + name of item 3 times with written word and drawing in view	✓ Move to next participant's turn	✖ If participant cannot repeat the word, acknowledge the difficulty, offer encouragement and move to the next participant's turn

Treatment Adherence

All therapists for the trial completed face to face or online mandatory training for one of the interventions only. Practical simulations of the treatments were offered as part of the training to ensure therapists fully understood the intervention. Differing treatment arms were not collocated within a building at the same time, to avoid therapists and participants discussing treatments and potential contamination across study arms.

All treatment sessions were videorecorded in full. To confirm treatment adherence, the integrity monitor — a speech pathologist employed by COMPARE — reviewed the full recording of each therapist's first day of treatment and a random 15 minute quarter of each session from the sixth day of treatment. A nine-point checklist was used to quantify adherence to key aspects of each therapy. The monitor provided feedback to the therapists on any deviations from the M-MAT or CIAT Plus intervention protocols within 24 hours of the treatment.

Therapists entered therapy start and stop times into the database each day, and the integrity monitor checked these times to confirm that the total duration of intervention provided met the time specified in the protocol.

Analysis

Baseline characteristics

The characteristics of participants at baseline were taken from the main trial screening and baseline assessments, and summarised using means and standard deviations or proportions as appropriate. Characteristics included demographics, baseline severity and measures of cognition. Statistical baseline comparisons were

not used as they are strongly discouraged by leading journals and the CONSORT guidelines for reporting RCTs (de Boer et al., 2014). Instead, the standardised mean difference for continuous variables was calculated with any difference greater than 0.1 considered to show possible imbalance (Schober & Vetter, 2019). Corrections for imbalances were to be considered depending on the final sample size and magnitude of imbalance.

Main aims – Primary and secondary outcomes

For hypotheses relating to time and group across outcome measures, mixed repeated measures ANOVAs were conducted to detect significant within- and between-group change across time. Bonferroni correction was applied within each ANOVA conducted.

In order to meet assumptions of the mixed ANOVA:

- Absence of outliers within each cell was confirmed using Studentised residuals of $>\pm3$

- Normality within each cell was checked using the Shapiro-Wilk test and QQ plots

- Homogeneity of variance was checked using Levene's test

- Sphericity was checked using Mauchly's (sig. >.05). In the case of low sphericity, correction to the significance level of the ANOVA interactions was applied using Greenhouse-Geisser.

If results of the ANOVA indicated a statistically significant interaction ($p <$ .05), simple main effects were examined using univariate ANOVAs. Effect sizes

(partial Eta squared) were reported for each comparison. For a statistically significant

main effect with no significant interaction, the pairwise comparisons were reported.

For the naming battery, the ANOVA was conducted separately for treated

items (n = 80), and for untreated items (n = 100) as a control set, with the prediction

that untreated items would show improvement.

Supplementary Aim

To investigate the relationship between baseline characteristics and

improvement in the WAB-R AQ and naming battery, correlations were conducted

using Pearson's r or Spearman's rho depending on whether the data was normally

distributed. Improvement was defined as change in scores from Baseline 1 to Post

Intervention.

Results

Participant Flow

Seventy of 216 participants in the main COMPARE trial were randomised to the Usual Care arm between July 2016 and March 2020. Of these 70, 54 consented to participate in the present substudy. Figure 5-2 displays the CONSORT Flow Diagram and participant flow. The substudy was ceased early (in April 2020) when COVID-19 infection rates in Australia indicated possible risk to participants if they attended group therapy. Fourteen participants were therefore unable to participate (cancelled prior to randomisation). Four participants did not have sufficient nearby participants to form a treatment group by the cessation of the trial, two withdrew due to health concerns unrelated to the trial and one experienced a Serious Adverse Event. This event was unrelated to the trial, as verified by the data safety management board.

Thirty-three participants were randomised to CIAT Plus ($n = 16$) and M-MAT ($n = 17$). A total of 16/17 M-MAT participants and 10/16 CIAT Plus participants received the full allocated intervention. One M-MAT participant was unexpectedly admitted to hospital for an adverse event unrelated to the study and completed only 21/30 hours of intervention. One CIAT Plus group ($n = 3$) was ceased after two days (6/30 hours) due to COVID-19 precautions and in another group, two participants withdrew due to travel time and an interpersonal conflict respectively, meaning the remaining participant could not continue. Data from ten CIAT Plus and 16 M-MAT participants were analysed, although participants with missing datapoints were unable to be included in the ANOVAs. Intention To Treat analysis and imputation of missing data were not conducted due to the small sample size. The Follow-Up

assessments for one M-MAT group were converted to online format (via videocall) due to COVID-19 precautions. One participant declined to participate in this online assessment and therefore did not have data at the 12 week Follow-Up point. The WAB-R AQ is equivalent in videoconference and in-person administration (Dekhtyar et al., 2020), as is the SAQOL-39g (Caute et al., 2012). The remaining measures have not had online and face-to-face administration compared in research but the conversion was required for safety of participants and scores are not hypothesised to differ dramatically in this format.

Figure 5-2 – Participant flow through trial
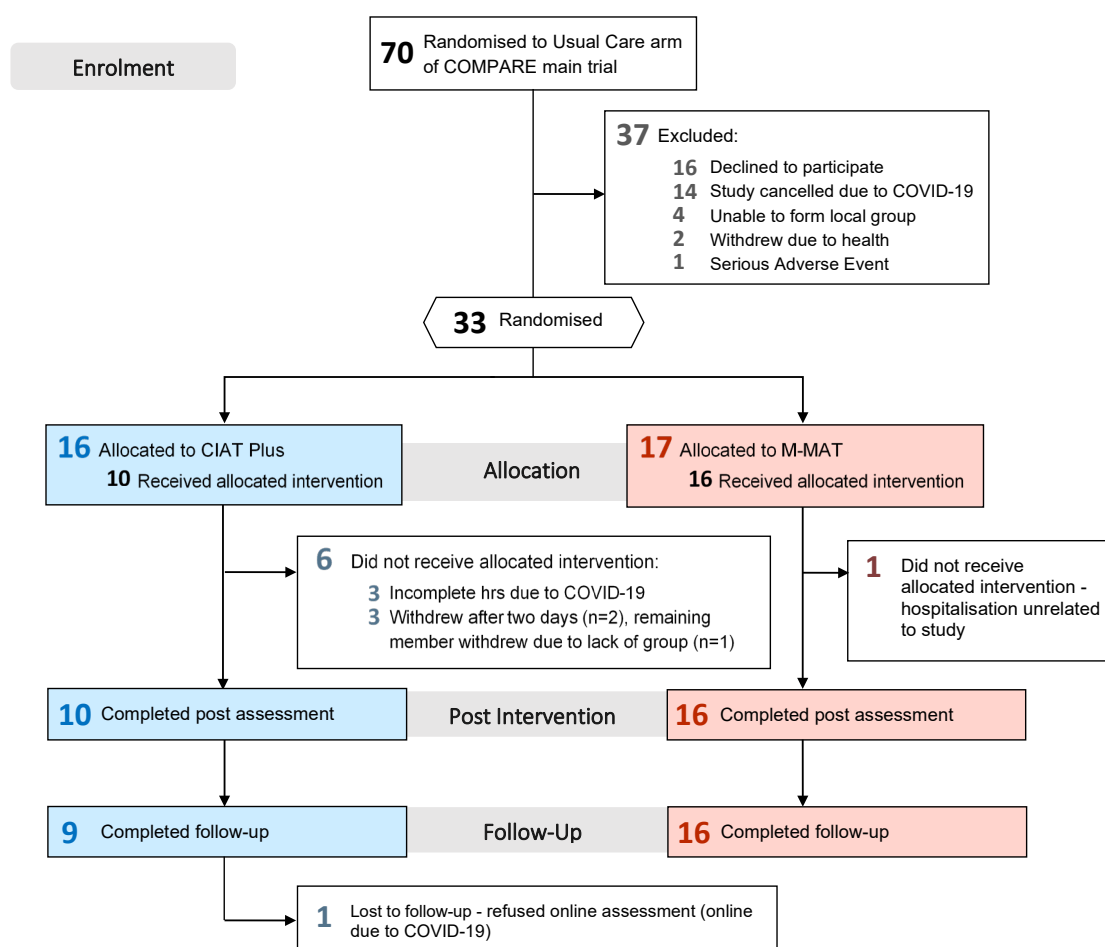
Baseline Characteristics

Table 5-4 shows participant characteristics at Baseline 0. The mean age of

participants was 64 years (SD = 11.5) with a range of 38 to 86 years. Six of 26 (23%)

were female. Participants were a mean of 44.2 months post stroke (SD = 27.6)

ranging from 9 months to 8;10 years. The mean WAB-R AQ was 68.6 (SD = 19).

Sixteen participants had mild aphasia, nine had moderate aphasia and one person

had severe aphasia. Analysis of baseline characteristics that were continuous and

normally distributed showed that the M-MAT and CIAT Plus arms were comparable

in terms of aphasia severity, non-verbal reasoning and short term memory, while

there were differences above >0.1 SMD in age (mean difference 6.1 years),

education (0.8 years) and several cognitive and linguistic measures. However,

inspection of the raw mean differences did not suggest highly imbalanced groups

(6.1 years in age, 0.8 years education, 0.8/54 Pyramids & Palm Trees score, 0.4

Elevator counting, 0.4 Visual Elevator, 2.2 total items on Picture Span Verbal

Memory). Appendix 4 contains visual representations of each characteristic in

beeswarm plots.

Table 5-4 – Baseline 0 Participant Characteristics

	M-MAT (*n = 16*)	CIAT Plus (*n = 10*)	SMD
Age	66.1 (11.2)	60.5 (11.5)	0.491*
Female:Male (% female) *Nil non-binary/non-disclosed reported*	5:11 (31%)	1:9 (10%)	
Years of education	15.4 (4.5)	14.7 (3.8)	0.176*
Handedness			
Right *n* (%)	15 (94%)	8 (80%)	
Left *n* (%)	1 (6%)	1 (10%)	
No preference *n* (%)	-	1 (10%)	
Stroke type			
Ischaemic	9 (56%)	8 (80%)	
Haemorrhagic	6 (38%)	2 (20%)	
Unknown	1 (6%)	-	

Modified Rankin Scale Score			
Low (Scored 0-2)	6	4	
High (Scored 3-6)	10	6	
Months Post Onset Median (IQR)	30.5 (35.5)	46.5 (47.5)	
WAB-R AQ Mean (SD)	69.1 (21.5)[a]	67.9 (15.3)	0.065
Mild n (%)	12 (66%)	4 (40%)	
Moderate n (%)	3 (19%)	6 (60%)	
Severe n (%)	1 (6%)	-	
Apraxia Severity Rating Scale			
No Apraxia	7 (44%)	3 (30%)	
Mild	3 (19%)	1 (10%)	
Mild-Moderate	3 (19%)	4 (40%)	
Moderate	3 (19%)	2 (20%)	
Pyramids and Palm Trees	48.0 (4.0)	48.8 (2.3)	0.235*
Attention (Elevator Counting)	5.8 (1.8)	6.2 (1.1)	0.282*
Cognitive flexibility (Visual Elevator)	5.0 (2.8)	5.4 (2.5)	0.144*
Short-term Memory (Forward span – total items recalled on Picture Span Verbal Memory)	42.9 (21.4)	39.7 (18.4)	0.159*
Working Memory (Reverse Span - total items recalled on Picture Span Verbal Memory)	32.3 (14.8)	32.1 (17.0)	0.011
Non-verbal reasoning (Raven's Progressive Matrices)	29.3 (6.7)	29.1 (5.7)	0.024

Note: Percentages rounded and may not total 100%.
[a]*One WAB-R AQ not available at Baseline 0, instead extracted from Baseline 1 for this table.*

Participants who withdrew

Baseline characteristics of the six participants who did not complete intervention are visualised in Appendix 4 alongside data for participants who completed intervention. Age, years of education, months post onset and apraxia of speech were similar between included and withdrawn, while the participants who withdrew had a higher proportion of females and a higher mean WAB-R AQ score.

Intervention Data and Treatment Fidelity

Table 5-5 contains data on intervention provided to participants and results of fidelity checks. The M-MAT arm had a higher proportion of participants assigned to the Hard and Mild word sets compared to CIAT Plus while CIAT Plus had more

participants assigned to the Moderate word set. Although assignment of a word set was based on initial naming battery scores, these proportions also correspond to the distribution of severity in each arm. Participants received a mean of 28.7 hours of intervention (SD 1.6) with similar total dose (mean 0.6hr difference) and average session duration (mean 1min difference) between arms. In both arms, the majority of participants progressed three levels of grammatical complexity by the end of intervention (M-MAT 56%, CIAT Plus 60%). Therapy integrity monitoring was conducted for 10/10 initial treatment days and 9/10 of the sixth treatment days. Of the monitored sessions, one CIAT Plus treatment day was non-adherent to the therapy protocol – insufficient phonemic and orthographic cues were provided. Feedback was provided to the therapist within 48 hours and the following therapy integrity check was 100% adherent. All M-MAT treatment sessions were adherent to the protocol.

Blinding

During baseline assessments, two assessors indicated that they had become unblinded (i.e., could guess that the participants were in the usual care arm) due to comments from the carer. Baseline assessments occurred prior to randomisation into a stubstudy treatment. Different assessors were used for the substudy for these participants. There was no unblinding at the post intervention assessment. At the follow up assessment, two unblinded assessments (both M-MAT) occurred where a trial coordinator completed assessments over zoom due to COVID-19. This was approved by the management committee.

Usual care during baseline period

The majority of participants did not participate in speech pathology-related usual care during the baseline period. The median number of hours of aphasia groups or speech pathology intervention was 0 (IQR 0 to 1.2) for CIAT Plus and 0 (IQR 0 to 3.0) for M-MAT.

Table 5-5 – Intervention data

	Total	M-MAT	CIAT Plus
Intervention hours Mean (SD)	28.7 (1.6)	28.5 (1.8)	29.1 (1.4)
Length of sessions Mean mins (SD)	58 (3)	57 (3)	58 (3)
Stimulus Set			
Easy	5	4 (25%)	1 (10%)
Moderate	6	2 (12.5%)	4 (40%)
Hard	15	10 (62.5%)	5 (50%)
Grammatical Levels progressed by final session			
0	4 (15.4%)	3 (19%)	1 (10%)
1	3 (11.5%)	0	3 (30%)
2	2 (7.7%)	2 (12.5%)	0
3	15 (57.7%)	9 (56%)	6 (60%)
4	3 (11.5%)	2 (12.5%)	1 (10%)
Intervention checks / Planned checks	38/40	22/24	16/16
% of checked sessions compliant	95% (36/38)	100% (22/22)	88% (14/16)

Main Aim - comparing CIAT Plus to M-MAT and demonstrating efficacy

Full details of statistical analysis are found in Appendix 4 of this chapter.
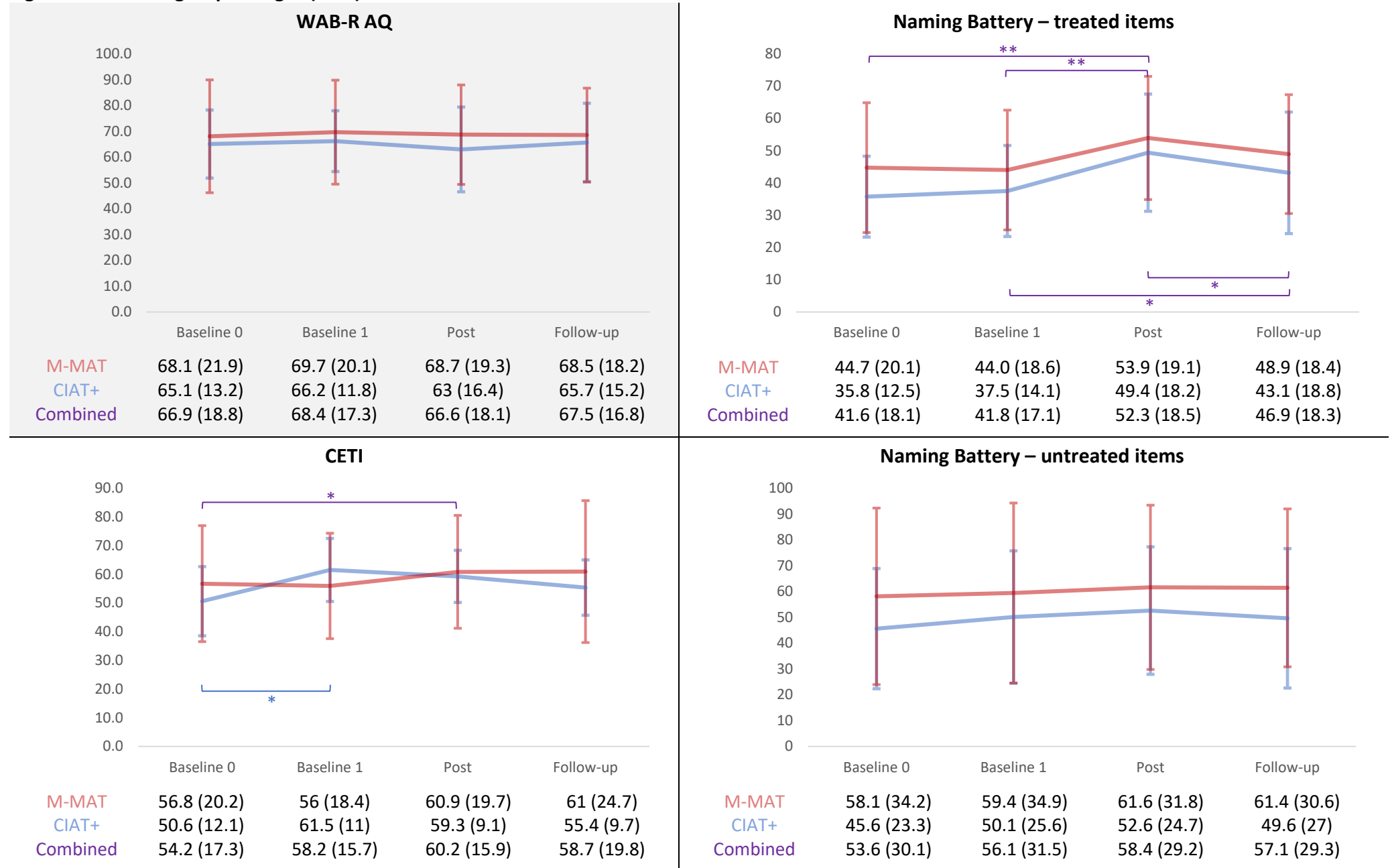
Figure 5-3 shows mean scores by group and time for all outcome measures, with

WAB-R AQ in the first panel. Figure 5-4 shows individual scores across time for each

outcome.
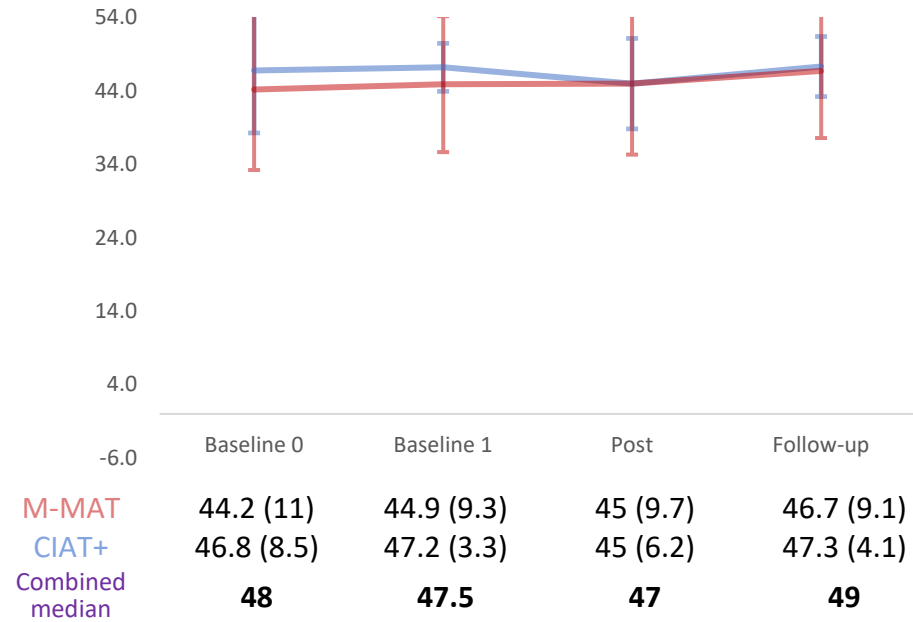
Primary Outcome - WAB-R AQ

There was no significant main effect of time on WAB-R AQ scores, $F(3, 66) =$

0.917, $p =0.44$, partial $\eta^2 = .040$; no main effect of group, $F(1, 22) = .255$, $p = .618$,

partial $\eta^2 = .011$; and no interaction between group and time, WAB-R AQ, $F(3, 66) =$

0.523, $p = .67$, partial $\eta^2 = .023$. Hypothesis 1 was therefore supported in that there

was no difference between CIAT-Plus and M-MAT at post-treatment or 12-week

Follow-Up. However, hypothesis 2 — that both treatments would result in >5 point

improvement at Post Treatment — was refuted, with a non-significant mean

decrease in WAB-R AQ scores from Baseline 1 to Post Intervention (M-MAT -1.0, CIAT

Plus -3.2). Hypothesis 3 — that both treatments would maintain improvement at

Follow-Up — was also rejected as there was no significant difference between

Follow-Up and either Baseline. Lastly, Hypothesis 4 predicted that there would be no

significant reduction in scores at Follow-Up compared to Post Intervention. While

there was no significant reduction in scores (M-MAT -1.2, CIAT Plus -0.5) between

these timepoints, there was no gain to maintain and so this Hypothesis was also
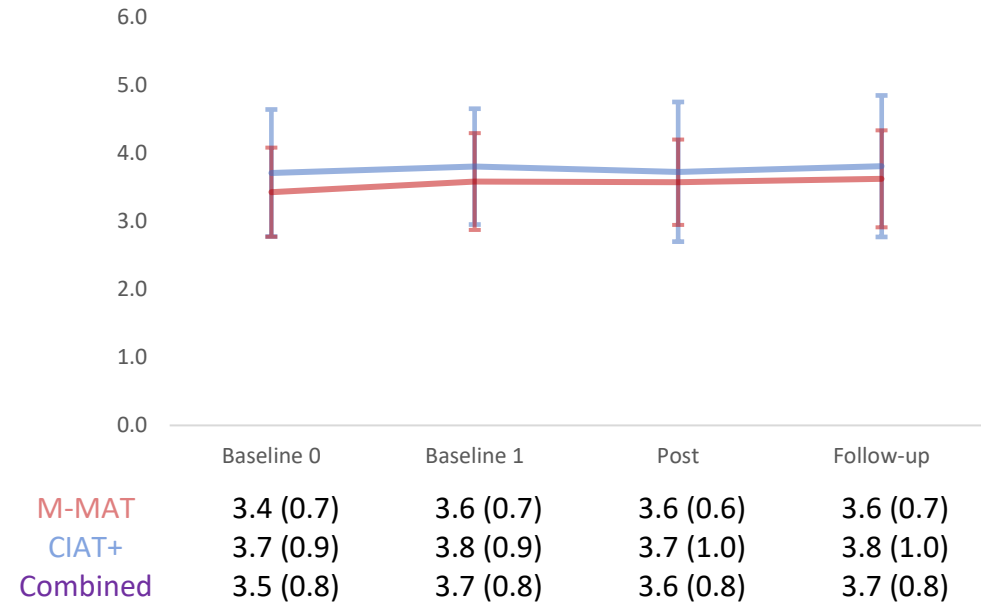
rejected.

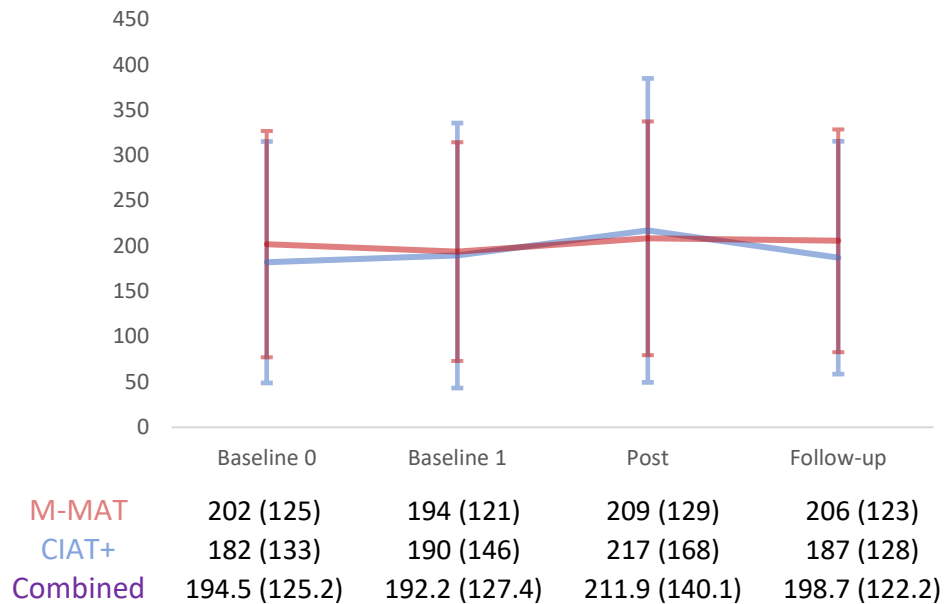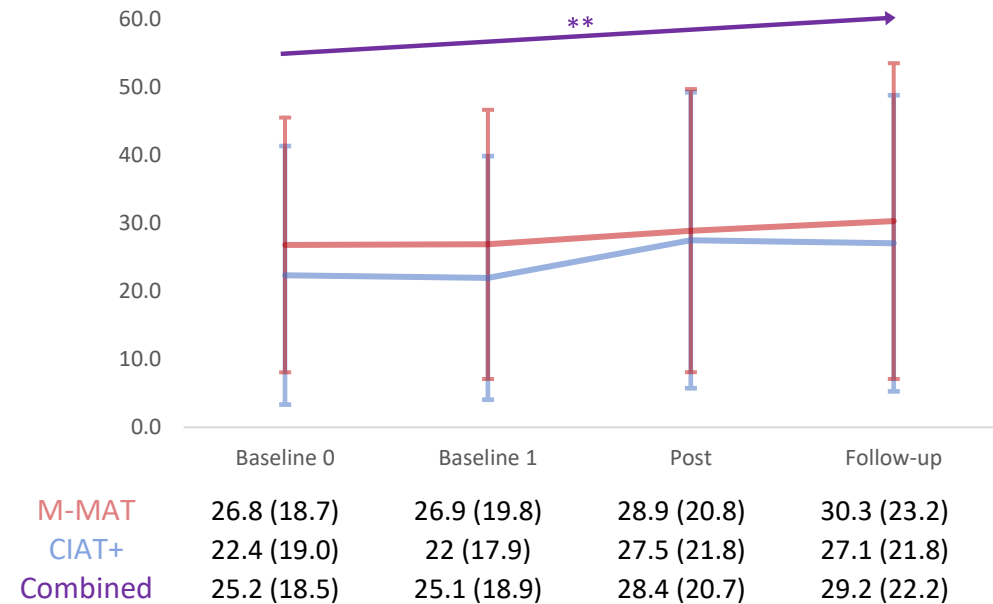**Figure 5-3 – Mean group changes (± SD) over time for CIAT Plus and M-MAT**

WAB-R AQ

	Baseline 0	Baseline 1	Post	Follow-up
M-MAT	68.1 (21.9)	69.7 (20.1)	68.7 (19.3)	68.5 (18.2)
CIAT+	65.1 (13.2)	66.2 (11.8)	63 (16.4)	65.7 (15.2)
Combined	66.9 (18.8)	68.4 (17.3)	66.6 (18.1)	67.5 (16.8)

Naming Battery – treated items

	Baseline 0	Baseline 1	Post	Follow-up
M-MAT	44.7 (20.1)	44.0 (18.6)	53.9 (19.1)	48.9 (18.4)
CIAT+	35.8 (12.5)	37.5 (14.1)	49.4 (18.2)	43.1 (18.8)
Combined	41.6 (18.1)	41.8 (17.1)	52.3 (18.5)	46.9 (18.3)

CETI

	Baseline 0	Baseline 1	Post	Follow-up
M-MAT	56.8 (20.2)	56 (18.4)	60.9 (19.7)	61 (24.7)
CIAT+	50.6 (12.1)	61.5 (11)	59.3 (9.1)	55.4 (9.7)
Combined	54.2 (17.3)	58.2 (15.7)	60.2 (15.9)	58.7 (19.8)

Naming Battery – untreated items

	Baseline 0	Baseline 1	Post	Follow-up
M-MAT	58.1 (34.2)	59.4 (34.9)	61.6 (31.8)	61.4 (30.6)
CIAT+	45.6 (23.3)	50.1 (25.6)	52.6 (24.7)	49.6 (27)
Combined	53.6 (30.1)	56.1 (31.5)	58.4 (29.2)	57.1 (29.3)

Scenario Test

	Baseline 0	Baseline 1	Post	Follow-up
M-MAT	44.2 (11)	44.9 (9.3)	45 (9.7)	46.7 (9.1)
CIAT+	46.8 (8.5)	47.2 (3.3)	45 (6.2)	47.3 (4.1)
Combined median	**48**	**47.5**	**47**	**49**

SAQOL

	Baseline 0	Baseline 1	Post	Follow-up
M-MAT	3.4 (0.7)	3.6 (0.7)	3.6 (0.6)	3.6 (0.7)
CIAT+	3.7 (0.9)	3.8 (0.9)	3.7 (1.0)	3.8 (1.0)
Combined	3.5 (0.8)	3.7 (0.8)	3.6 (0.8)	3.7 (0.8)

CIU count*

	Baseline 0	Baseline 1	Post	Follow-up
M-MAT	202 (125)	194 (121)	209 (129)	206 (123)
CIAT+	182 (133)	190 (146)	217 (168)	187 (128)
Combined	194.5 (125.2)	192.2 (127.4)	211.9 (140.1)	198.7 (122.2)

CIUs per minute

	Baseline 0	Baseline 1	Post	Follow-up
M-MAT	26.8 (18.7)	26.9 (19.8)	28.9 (20.8)	30.3 (23.2)
CIAT+	22.4 (19.0)	22 (17.9)	27.5 (21.8)	27.1 (21.8)
Combined	25.2 (18.5)	25.1 (18.9)	28.4 (20.7)	29.2 (22.2)

*One outlier removed for this figure and for the analysis. Sensitivity test indicated no impact of removal.

Note: Purple marks indicate significance of combined group data. $* = p < .05$; $** = p < .01$. Bonferroni correction applied within ANOVAs for each outcome measure.

Figure 5-4 – Individual changes over time
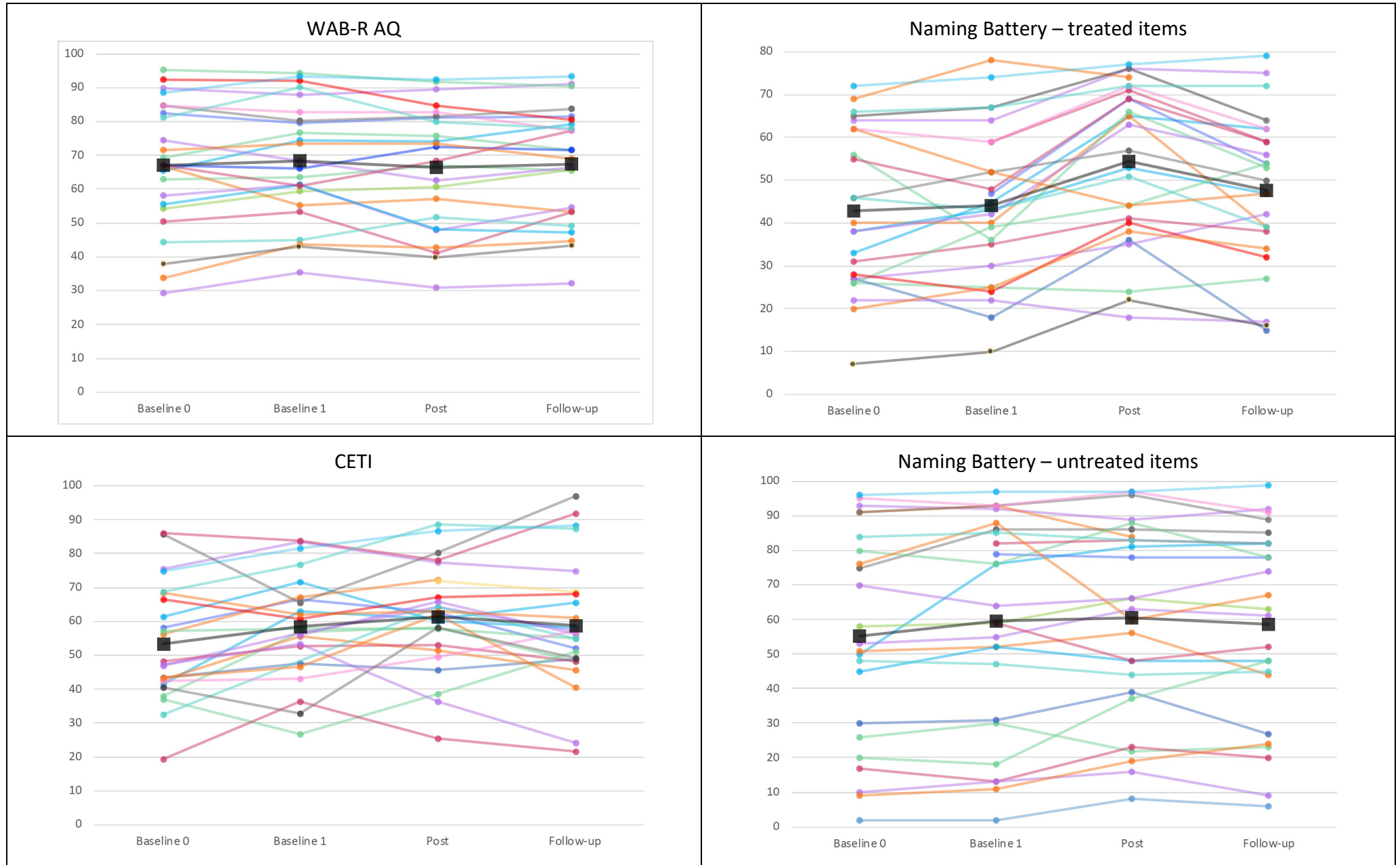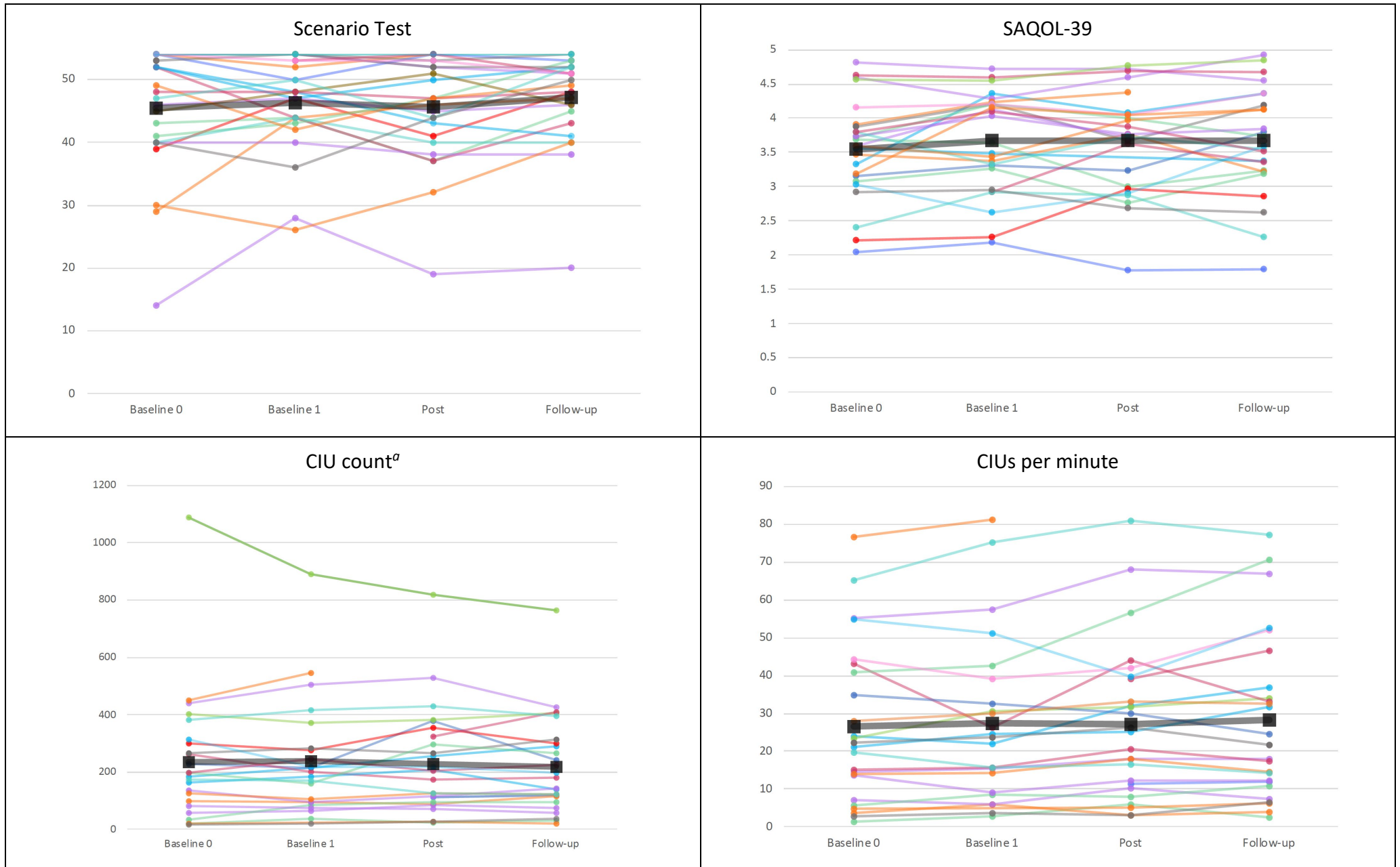
Note: Lines colours are for clarity and do not represent the same participant across figures. Participants with missing datapoints included.
[a]Includes outlier excluded from data analysis

Secondary Outcomes

Naming Battery

A significant main effect of time was found for treated items $F(2.3, 48.6) = 14.392$, $p < .001$, partial $\eta^2 = .407$. This effect size of 0.407 is considered large according to Cohen's benchmarks (large >0.1379) for Partial Eta Squared (Richardson, 2011). Post hoc analysis revealed significant differences between Post Intervention and each Baseline (both p < .001), Post Intervention and Follow-Up ($p = .040$), and Baseline 1 and Follow-Up ($p = .021$). Table 5-6 shows mean difference in raw naming scores based on estimated marginal means.

There was no significant main effect of group, $F(1, 21) = .732$, $p = .402$, partial $\eta^2 = .034$; and no significant interaction between group and time, $F(2.3, 48.6) = .438$, $p = .667$, partial $\eta^2 = .020$.

An outlier in CIAT Plus at Baseline 0 violated assumptions of normality and a second analysis without this participant was also conducted (see Appendix 4 for detailed results). The significant differences between Post Intervention and other timepoints remained, but Follow-Up was no longer significantly different to Baseline 1 ($p = .054$). Group and interaction effects were unchanged.

In untreated items of the naming battery, there was no significant main effect of time, $F(2.086, 41.722) = 2.560$, $p = .063$, partial $\eta^2 = .114$; nor a main effect of group, $F(1, 20) = 0.645$, $p = .431$, partial $\eta^2 = .031$; nor any significant interaction between time and group, $F(2.086, 41.722) = .411$, $p = .674$, partial $\eta^2 = .20$.

Hypothesis 5a was therefore supported for the naming battery — that CIAT Plus and M-MAT would result in improved scores at Post Treatment and Follow-Up.

However, it was also noted that there was a significant decrease in naming scores between Post Treatment and Follow-Up (M = -5.6 words).

Table 5-6 – Pairwise comparisons for Naming Battery Treated Items

Time	Comparison	Mean Difference (95% CI*)	Sig.[a]
	B1	-0.508 (-5.102 to 4.085)	1.000
Baseline 0	Post	**-11.413 (-17.964 to -4.861)**	**0.000****
	F/U	-5.788 (-12.867 to 1.292)	0.161
	B0	0.508 (-4.085 to 5.102)	1.000
Baseline 1	Post	**-10.904 (-16.929 to -4.880)**	**0.000****
	F/U	**-5.279 (-9.944 to -.614)**	**0.021***
	B0	**11.413 (4.861 to 17.964)**	**0.000****
Post Intervention	B1	**10.904 (4.880 to 16.929)**	**0.000****
	F/U	**5.625 (0.188 to 11.062)**	**0.040***
	B0	5.788 (-1.292 to 12.867)	0.161
Follow-Up	B1	**5.279 (0.614 to 9.944)**	**0.021***
	Post	**-5.625 (-11.062 to -.188)**	**0.040***

Based on estimated marginal means
[a]Adjustment for multiple comparisons: Bonferroni.
* $p < .05$
** $p < .01$

Connected Speech - CIU count and CIUs/minute

Outlier data was identified for one participant in the CIU count outcome measure which had a studentised residual value >3 for each timepoint. The repeated measures ANOVA was conducted with and without the participant with no impact on the result (see Appendix 4 for outcomes of both analyses). There was no significant main effect of time, $F(3, 57) = 1.678$, $p = .182$, partial $\eta^2 = .081$ or group, $F(1, 19) =$

0.022, *p* = .884, partial η² = .001. There was also no significant interaction between time

and group, *F*(3, 57) = 0.769, *p* = .516, partial η² = .039.

For CIUs per minute, the main effect of time was statistically significant, *F*(3, 60)

= 4.492, *p* = .007, partial η² = .183; however, there were no statistically significant

differences between timepoints on post hoc pairwise comparisons (Table 5-7). Mean

raw sequential changes in mean CIUs per minute across timepoints were -0.1, +3.3 and

+0.7 respectively. The main effect of group was not significant, *F*(1, 20) = .156, *p* = .697,

partial η² = .008; and there was no statistically significant interaction between group

and time, *F*(3, 60) = .529, *p* = .664, partial η² = .026.

Overall, there was insufficient evidence of improvement to support Hypothesis

5b for connected speech, which predicted significantly improved CIUs and CIUs per

minute at Post Treatment and Follow-Up.

Table 5-7 – Pairwise comparisons for CIUs per minute

Time	Comparison	Mean Difference (95% CI*)	Sig.[a]
Baseline 0	B1	0.147 (-3.299 to 3.592)	1.000
	Post	-3.632 (-8.097 to 0.832)	0.164
	F/U	-4.112 (-9.924 to 1.700)	0.309
Baseline 1	B0	-0.147 (-3.592 to 3.299)	1.000
	Post	-3.779 (-7.610 to 0.052)	0.055
	F/U	-4.259 (-9.119 to 0.601)	0.111
Post Intervention	B0	3.632 (-0.832 to 8.097)	0.164
	B1	3.779 (-0.052 to 7.610)	0.055
	F/U	-0.480 (-4.463 to 3.503)	1.000
Follow-Up	B0	4.112 (-1.700 to 9.924)	0.309
	B1	4.259 (-0.601 to 9.119)	0.111
	Post	0.480 (-3.503 to 4.463)	1.000

Based on estimated marginal means
[a]Adjustment for multiple comparisons: Bonferroni.

CETI

There was a statistically significant interaction between group and time, $F(3, 60)$ = 3.006, p = .037, partial η^2 = .131. Post hoc analysis using Univariate ANOVAs found no significant between-group differences at any timepoint. The simple main effect of time in CIAT Plus was significant, $F(3, 24)$ = 8.692, $p < .001$, partial η^2 = .521; and pairwise comparison revealed a statistically significant difference between Baseline 0 and Baseline 1 in the CIAT group (mean improvement 10.9, $p = .004$). In the M-MAT group, the simple main effect of time was not significant, $F(3, 36)$ = 1.305, $p < .288$, partial η^2 = .098.

The main effect of time for CETI was significant, $F(3, 60)$ = 3.173, $p = .031$, partial η^2 = .137. This value of Partial Eta Squared corresponds to a large effect size as per Cohen's values (Richardson, 2011). Post hoc comparisons found a significant difference between Baseline 0 and Post Intervention across treatments (Mean improvement +6.4, $p = .036$) – see Table 5-8.

There was no significant main effect of group $F(1, 20)$ = .074, $p = .788$, partial η^2 = .004.

Results do not unequivocally support Hypothesis 5c as measured by the CETI (i.e., significantly better functional communication at Post Intervention and Follow-Up). Post Intervention was significantly improved from Baseline 0 but not Baseline 1. Follow-Up scores were not significantly different.

Table 5-8 – Pairwise comparisons for CETI

Time	Comparison	Mean Difference (95% CI*)	Sig.[a]
Baseline 0	B1	-5.058 (-10.341 to 0.225)	0.066
	Post	**-6.389 (-12.476 to -0.301)**	**0.036***
	F/U	-4.482 (-11.172 to 2.208)	0.384
Baseline 1	B0	5.058 (-0.225 to 10.341)	0.066
	Post	-1.331 (-7.439 to 4.778)	1.000
	F/U	0.576 (-7.748 to 8.901)	1.000
Post Intervention	B0	**6.389 (0.301 to 12.476)**	**0.036***
	B1	1.331 (-4.778 to 7.439)	1.000
	F/U	1.907 (-3.815 to 7.629)	1.000
Follow-Up	B0	4.482 (-2.208 to 11.172)	0.384
	B1	-0.576 (-8.901 to 7.748)	1.000
	Post	-1.907 (-7.629 to 3.815)	1.000

Based on estimated marginal means
[a]Adjustment for multiple comparisons: Bonferroni.
* $p < .05$
** $p < .01$

Scenario Test

Data for the Scenario Test failed multiple assumptions of the repeated measures mixed ANOVA including parametric distribution, absence of outliers and equality of variances. The Kruskall Wallace H test was therefore applied. As there were no obvious differences in scores between groups at any timepoint (as per chart and scores in Figure 5-3) and given the absence of between-group effects in all other outcomes, a Friedman test was applied across time without between-group comparisons. Scenario Test scores were not significantly different at the four timepoints, $\chi^2(3) = 4.327$, $p = .228$. Hypothesis 5d was therefore rejected for the Scenario Test.

SAQOL-39

Results of analysis for the SAQOL-39 did not find a significant main effect of time, $F(3, 60) = 1.091$, $p = .360$, partial $\eta^2 = .052$; of group, $F(1, 20) = .398$, $p = .536$, partial $\eta^2 = .019$; or any significant interaction between group and time $F(3, 60) = .194$, $p = .900$, partial $\eta^2 = .010$. Hypothesis 5e was therefore rejected for the SAQOL-39.

Supplementary aim - exploration of baseline characteristics and treatment response

The change in WAB-R AQ (B1 to Post Intervention change) and several other baseline measures were not normally distributed, therefore Spearman's Rank Correlation was applied to explore relationships between participant characteristics previously demonstrated to correlate with improvement and change in scores. As displayed in Table 5-9, no measures were significantly correlated with change in either WAB-R AQ or naming battery (treated items).

While individual WAB-R AQ scores did increase and decrease between Baseline 1 and Post Intervention, Post Hoc analysis found no significant correlation between these changes and change in Treated Items Naming Battery ($r_s = .22$, $p = .27$). Further, the change in Treated Items Naming Battery was not correlated with initial naming severity/naming battery scores at Baseline 0 ($r_p = -.11$, $p = .615$).

Table 5-9 – Spearman's rank correlation: Participant characteristics vs. improvement from B1 to Post Intervention

Characteristic	r_s	p	S
WAB-R AQ			
Age	0.198	0.331	2344.7
Months Post Onset	-0.001	0.995	2929
Baseline severity	0.098	0.635	2639.6
Semantic impairment	0.139	0.498	2517.7
Phonological impairment	0.100	0.628	2633.2
Attention	0.307	0.135	1801.3
Cognitive Flexibility	0.304	0.193	926.24
Short Term Memory	0.014	0.946	2884.4
Working Memory	0.281	0.194	1455.7
Naming battery – treated items			
Age	-0.278	0.169	3738.5
Months Post Onset	-0.332	0.097	3897.3
Baseline severity	-0.041	0.844	3043.9
Semantic impairment	0.082	0.690	2684.7
Phonological impairment	0.001	0.995	2921
Attention	-0.109	0.605	2882.6
Cognitive Flexibility	0.264	0.260	978.39
Short Term Memory	-0.026	0.901	2999.8

Discussion

CIAT Plus and M-MAT – Between group comparisons

We hypothesised no significant differences between CIAT Plus and M-MAT at Post Intervention or Follow-Up timepoints. For the primary outcome (WAB-R AQ), there was no significant difference between the treatment groups across time, although this outcome did not show improvement across time in either group. Nor were there any

significant between-group differences on the Naming Battery treated items, despite significant improvement within both groups. Our hypothesis was based on our previous meta-analysis that found insufficient data to favour either approach over the other (Pierce et al., 2017) and the current results are consistent with this review.

There were also no main effects of group for secondary outcome measures, and no interactions between group and time with the exception of the CETI, where an interaction was found from Baseline 0 to Baseline 1 within CIAT Plus but not M-MAT. Such a change between baselines is presumably measurement error – it is difficult to hypothesise a feasible explanation for a true change in the CIAT Plus group during the baseline period which would not also apply to the M-MAT group. In fact, median speech pathology-related usual care was lower for CIAT Plus participants (1.2) than M-MAT (3.0). The clinically significant difference for the CETI is 12 (Lomas et al., 1989) while the mean change in the current study was 10.9. However, a p value of .004 means a change of this magnitude or greater should only occur by chance 1 in 250 times. Although the CETI was shown to be valid and reliable in Lomas et al. (1989), their sample included just 11 participants with acute-subacute aphasia and 11 with chronic aphasia. Further examination of the psychometrics properties of the CETI is required. In addition, it is a subjective rating scale and the small sample size of the CIAT Plus group ($n = 9$ for this outcome) may have affected results.

Overall, these results demonstrate that neither CIAT Plus nor M-MAT is superior when delivered at an intensity of 6 hours/week × 5 weeks. Instead, our results indicate that neither constraint nor multimodal approaches should be avoided or mandated in clinical practice. In previous research, some people have indicated that they prefer one of these treatments over the other. A participant in the Phase I crossover study by

Attard, Rose & Lanyon (2013) (*n* = 2) indicated a preference for MMAT over CIAT Plus, while in the Rose et al. (2013) phase 1 study (*n* = 11), six participants preferred M-MAT, three preferred CIAT Plus and two had no preference. One of these participants found the visual barriers of CIAT Plus disconcerting ("I am alone on the barrier," p963) while another liked the barriers, suggesting they reduced his performance pressure. Results from the current study suggests that individual preferences should therefore be considered in choosing these treatments without concern over efficacy.

Changes over time – within-group comparisons

Treated Naming Battery items improved following treatments. CIUs per minute improved across time overall, though no specific timepoint comparisons were statistically significant, and the CETI improved from Baseline 0 to Post Intervention. However, the WAB-R AQ scores did not change over time, nor did the CIU count, Scenario Test, SAQOL-39 or untreated Naming Battery items.

The improvement in treated items is consistent with our 2017 meta-analysis (Pierce et al., 2017) where 14/15 of the high rigour Single Case Experimental Designs we analysed used treated items as the primary outcome measure and all found positive effects. Improvement on treated items for naming is also consistent with interventions outside of constraint and multimodal treatments. Wisenburn and Mahoney (2009) found a large overall effect size for naming after reviewing 44 treatment studies. Thomas et al. (2020) also reviewed 48 naming intervention studies and found that all improved treated words, with a median of 33.5% or 16.4 word improvement compared to baseline (range 9 to 76%, 2.5 to 45.5 words), though studies included participants

with acute through to chronic aphasia. Median improvement in our trial was 13.75% or 11 items of the 80 treated items (range -10 to 37.5%, -8 to 30 words).

CIUs per minute improved over time with no specific timepoints reaching significance. However, the CIU *count* was stable over time, suggesting participants were able to describe the picture stimuli using a similar amount of information over a shorter duration. Importantly, the picture descriptions were standardised tasks rather than stimuli designed to elicit use of naming battery items, though there may have been some treated items present by coincidence, e.g., Fall, Give. Participants may have produced sentences more efficiently due to briefer word retrieval pauses and/or due to their practice of increasingly complex sentences in both treatments. Carrier sentences in both treatments included SVO and OVS structures which could have reduced sentence generation time. The effect of time on CIUs per minute might also be explained by increasingly familiarity with the picture description tasks; however, inspection of the mean of all participants (after pairwise exclusion for missing data) shows changes of -0.1 CIUs/min between baselines, +3.3 after intervention, and +0.7 at Follow-Up, demonstrating that the bulk of improvement occurred after intervention. These changes are raw means that are accompanied by high variance and are not statistically significant. The larger sample of the main COMPARE trial may be able to investigate impacts of intervention on CIUs more clearly.

The CETI, a measure of communication activity and participation, improved significantly from Baseline 0 to Post Intervention which suggests that the participants' carers noted a difference in these domains. However, the raw change was +6 which is less than the clinically significant difference of 12 (Lomas et al., 1989), and the

significant difference between baselines within the CIAT Plus group calls into question the validity of these results based on our small sample.

At the group level, the primary outcome – the WAB-R AQ – failed to change over time. The simplest explanation for the lack of effect on WAB-R AQ in the present study is that, while treatments improved naming of treated items, and possibly connected speech efficiency (CIUs per minute), they did not impact overall underlying aphasia severity. Individual score changes for treated items of the Naming Battery did not correlate with changes in WAB-R AQ, supporting this explanation. Similarly, in their large trial, Fleming et al. (2020) compared software-based treatment for auditory comprehension to a control group in a Randomised Controlled Trial and found that participants improved on trained items but not auditory comprehension subtests of the Comprehensive Aphasia Test.

However, this is in contrast to our previous review (Pierce et al., 2017) where three of four high rigour group studies we reviewed also measured aphasia severity as a primary outcome. The three studies used the Aachen Aphasia Test, though Pulvermüller et al. (2001) used only four of five subtests. Two of the three (Meinzer et al., 2007; Pulvermüller et al., 2001) found improvement on the Aachen scores following CIAT and the other (Wilssens et al., 2015) reported that 4/5 subtests significantly improved. Meinzer et al. and Pulvermüller et al. used marginally smaller samples than the present trial. Mean participant age was also younger (Meinzer et al. 56 years, Pulvermüller et al. 55 years) compared to this trial (61 years for CIAT Plus). More recently, Stahl et al. (2018) found significant improvements on an abbreviated AAT (4/5 subtests) following both 24 and 48 hours of a constraint-based treatment. The schedule for the arm receiving 24 hours of treatment matched the present trial, with two hours x three days

each week, though at a lower total dose (24 vs. 30 hours). Our results do not support

such a change in severity following CIAT Plus (or M-MAT), although the WAB-R and the

Aachen tests may not be fully equivalent. Participant characteristics may also explain

the differing outcomes. Participants in Stahl et al. had a normal mean visual short term

memory; in the current study, visual short term memory was not assessed but mean

scores on the Picture Span Verbal Memory (41.7, forwards) were well below normal

scores (83.2) (DeDe et al., 2014). Visual and auditory short term memory are thought to

closely interact (Baddeley, 2003) and so participants in our study may have also had

poorer visual short term memory, as one example.

Taken alongside the absence of change in most other outcome measures, these

results suggest that improved confrontational naming does not transfer to

activity/participation or quality of life following 30 hours of CIAT Plus or M-MAT

distributed over 5 weeks. In our review of constraint and multimodal treatments, only a

small proportion of eligible studies included outcome measures for

activity/participation and quality of life (Pierce et al., 2017). Those that used the CETI

(3/4 CIAT and 4/6 multimodal) mostly reported small mean increases less than the

clinically significant difference of 12. Statistically significant improvements were found

in two CIAT studies on the Communicative Activity Log (Pulvermüller et al., 2001) and

the Amsterdam–Nijmegen Everyday Language Test (Wilssens et al., 2015) while the

Scenario Test changes were generally small. A recent large RCT (Palmer et al., 2019)

found significant improvements in naming for software-based treatment compared to

attention control and usual care (both $p < 0.0001$) yet no differences between groups in

functional communication (Therapy Outcome Measures), untreated words, use of

treated words in conversation, and self-ratings of participation and quality of life.

However, Breitenstein et al. (2017) used an activity/participation measure (Amsterdam–Nijmegen Everyday Language Test A-scale) as the primary outcome in their large RCT of standard treatment compared to waitlist, and found that the treatment group significantly improved compared to controls. Effective aphasia treatments must impact performance outside the clinic environment but the evidence that such generalisation occurs is not yet conclusive in high quality trials.

An alternative explanation for the lack of change is that providing treatment for two hours × three days per week × five weeks (30 hours total), although sufficient to generate change in naming, was insufficiently stimulating to the language system and thus did not improve general language function, activity/participation or quality of life. A higher dose, or higher intensity of the same dose, may be required to achieve this. However, the failure of generalisation to general function in some other trials (e.g. Palmer et al., 2019) suggests that scheduling was not the key problem. In addition, one study provided constraint-based treatment at a schedule matching ours but a lower total dose (24 vs 30 hours), yet found improvement on impairment-based measures (Stahl et al., 2018). Comparison of our results to the main COMPARE trial ($n = 216$), which provided the same dose of identical treatment over just two weeks, will allow testing of the impact of intensity, when results are released.

Measurement error?

A change of >5 points on the WAB-R AQ is considered clinically meaningful (Hula et al., 2010), and >8.26 is the Minimal Detectable Change at 90% confidence interval (Breitenstein et al., 2021). In this study, the mean changes in the WAB-R AQ between timepoints were +1.4 (SD 5.2), -1.7 (5.7) and +0.9 (4.4) respectively. However,

examination of Figure 5-4 shows positive and negative individual changes that exceed the change thresholds, even between Baselines when no change would be expected: Individual changes ranged from -11.4 through to +10 between Baselines, -13.4 to +7.2 from B1 to Post Intervention, and -5.4 to +12.4 Post to Follow-Up. These changes are difficult to explain, particularly the negative changes and changes between Baselines. Our protocol included monitoring of assessment fidelity, with video review of any new assessor administering the WAB as well as any assessments where the WAB-R AQ was close to severity group (mild, moderate, severe) cutoff scores – a highly rigorous monitoring protocol. Where possible, the same assessor was used for a participant at each timepoint. Erroneous administration or scoring of the WAB is therefore unlikely to be responsible for the individual variability in scores. In addition, the Pearson's correlation coefficient of Baseline 0 to Baseline 1 scores was .96 (95% CI .91 to .98) demonstrating high reliability of the WAB. Rather than measurement error or bias, the 'bounces' in WAB-R AQ scores are likely explained by intraindividual variability. While this variability is frequently encountered in clinical practice, it has not been widely considered in aphasia research, though it is beginning to be examined further (Duncan et al., 2016). Internal variables such as fatigue, performance anxiety, mood and overall alertness may form part of the variability and these need to be considered in sample size calculations of future trials.

Maintenance of gains

Maintenance data is reported in less than half of aphasia research studies, with even fewer studies reporting data at 12 or more weeks after intervention (Menahemi-Falkov et al., 2020). The Follow-Up timepoint in this study was 12 weeks after

intervention. For the Naming Battery treated items, combined Follow-Up data was significantly less than at Post Intervention but significantly higher than Baseline 1. In other words, by 12 weeks, gains were already diminished though not back to baseline. A review of 44 word-finding treatments in aphasia found large effects for the first two months after therapy and a "lingering" but diminishing effect at three months (Wisenburn & Mahoney, 2009). Examining the raw data (Table 5-6 – number of named items), our data shows a mean gain of ≈11 items at Post Intervention, then a loss of about half of these gains (5.6 items) by Follow-Up. Our data therefore further supports the need to actively maintain gains from intensive treatment; for example, facilitating self-practice, additional therapy, or group intervention (Menahemi-Falkov et al., 2020). Comparison of our data with the forthcoming COMPARE main trial data at Follow-Up will also be informative – according to theories derived from the field of cognitive psychology, the more distributed schedule of this substudy should result in superior long term retention to the massed schedule of the main trial (Cepeda et al., 2006; Pierce et al., 2020).

As with the majority of results in the present study, individual levels of maintenance varied significantly (Figure 5-4). Individual analysis of individual trends has been recommended to allow a more nuanced picture of maintenance and examination of individual predictors (Menahemi-Falkov et al., 2020).

Supplementary aim: Treatment response and baseline characteristics

The supplementary aim of this study was to examine the relationship of treatment response to characteristics previously identified as predictors. None of the participants characteristics measured at baseline were significantly correlated with

changes on the WAB-R AQ or treated items of the Naming Battery. This result is interesting in the context of the variable results between participants. Change in treated items (Baseline 1 to Post Intervention) ranged from -8 to +30 items or -10% to +37.5% of the treated items. Some participants responded well to the intervention while others did not respond or performed worse – five participants showed negative changes in naming. While some of these changes may be attributable to previously discussed intra-individual variability, past research on treatment response would suggest that a detectable proportion of change was not random but predicted by the participant characteristics we examined (Dignam et al., 2015; Lambon-Ralph et al., 2010). Our correlations did not find any such relationships in this trial. However, while correlations are an exploratory approach, mixed effects modelling would have been an ideal method to elucidate the predictive value of each factor, including intraindividual variability as random error. This was not undertaken in this study because our sample was not recruited based on power calculations and was underpowered for a mixed effects approach. The results of the main COMPARE trial may be better powered to detect any relationships.

Limitations of this research

COVID-19 resulted in early termination of this trial and the loss of 14 participants who had consented to participate. More participants than anticipated also declined participation in the substudy ($n = 16$), possibly due to the burden of assessments – by the time of consenting to the substudy participants had typically completed 2-3 assessments for the COMPARE main trial at several hours each. An additional seven participants commenced but did not receive a complete treatment due to COVID ($n =$

3), withdrawal (n = 3) and an adverse event (n = 1). The resulting sample size was therefore smaller than planned and may have been underpowered to detect some within- and between-group differences and predictors of treatment response. Further, ANOVA requires casewise deletion where any data point is missing for an outcome, meaning that a participant's data cannot be included for analysis of that outcome. This further reduced sample size for some analyses, though no more than four participants were dropped for any one outcome. There were significant baseline differences between groups according to a standardised mean difference > 0.1, most notably age. However, the raw difference in means for most characteristics do not point to marked differences in outcomes (e.g., 6 years mean difference in age, 0.8 points mean difference in Pyramids & Palm Trees score).

Intention To Treat analysis was considered but not utilised. Intention To Treat is a preferred strategy for analysis but is not explicitly required by the CONSORT guidelines; instead, a clear description of who was included in the analysis should be reported (Moher et al., 2010). In four of seven participants who did not complete the minimum number of treatment hours, the causes were events unrelated to treatment – an adverse event with no connection to the trial, and trial cessation due to COVID-19. Two of the remaining withdrew after only two days of treatment due to travel time and an interpersonal conflict. The third group member was unable to continue treatment without a group. Inclusion of these participants who completed just 4/30 hours of the protocol would have diluted the results, but it is acknowledged that the reasons for withdrawal have been also influenced by the therapeutic atmosphere. Further qualitative information from these participants would be valuable.

Finally, despite recruitment being open to people with severe aphasia, only one participant in the study had severe aphasia (WAB-R AQ <33). Severity of aphasia is a negative predictive factor in long-term prognosis (Plowman et al., 2012) and therefore more data on treatment outcomes for this group is needed. Future research should consider how to increase recruitment and inclusion of people living with severe aphasia.

Conclusions

In summary, the key results of this study were:

− No significant differences between CIAT Plus and M-MAT for any outcomes

− The primary outcome, WAB-R AQ, did not significantly change over time

− Treated items of the naming battery did significantly change over time, with significant improvement following intervention. At 12 week Follow-up, treated items were significantly above baseline but significantly reduced from the post intervention score

− CETI scores improved from Baseline 0 to Post Intervention, but also between baselines for CIAT Plus

− CIUs per minute improved over time, possibly due to intervention, though pairwise comparisons were non-significant

− Other outcomes did not change over time

− Neither WAB-R AQ changes nor changes in treated items of the naming battery significantly correlated with baseline characteristics

This pilot RCT demonstrated that naming of treated items improves following low-moderate intensity group CIAT Plus or M-MAT treatment, and that the gains are

still partially maintained after twelve weeks. The efficiency of connected speech might also improve, but further generalisation to untreated items, general impairment measures and activity/participation and quality of life outcomes is not apparent. In addition, outcomes for CIAT Plus and M-MAT are equivalent at this schedule.

References

Attard, M. C., Rose, M. L., & Lanyon, L. E. (2013). The comparative effects of Multi-Modality Aphasia Therapy and Constraint-Induced Aphasia Therapy-Plus for severe chronic Broca's aphasia: An in-depth pilot study. *Aphasiology*, *27*(1), 80–111. https://doi.org/10.1080/02687038.2012.725242

Banks, J. L., & Marotta, C. A. (2007). Outcomes validity and reliability of the modified Rankin scale: Implications for stroke clinical trials: A literature review and synthesis. *Stroke; a Journal of Cerebral Circulation*, *38*(3), 1091–1096. https://doi.org/10.1161/01.str.0000258355.23810.c6

Basso, A., Capitani, E., & Laiacona, M. (1987). Raven's coloured progressive matrices: normative values on 305 adult normal controls. *Functional Neurology*, *2*(2), 189–194.

Bennett, H. E., Thomas, S. A., Austen, R., Morris, A. M. S., & Lincoln, N. B. (2006). Validation of screening measures for assessing mood in stroke patients. *British Journal of Clinical Psychology*, *45*(3), 367–376. https://doi.org/10.1348/014466505x58277

Breitenstein, C., Bowen, A., Brady, M., Hilari, K., Hillis, A. E., Kiran, S., Menahemi-Falkov, M., Rose, M. L., Szaflarski, J., Tippett, D. C., Visch-Brink, E., Wallace, S. J., & Willmes, K. (2021). *Operationalising treatment success in aphasia rehabilitation [Submitted manuscript]*.

Breitenstein, C., Grewe, T., Flöel, A., Ziegler, W., Springer, L., Martus, P., Huber, W., Willmes, K., Ringelstein, E. B., Haeusler, K. G., Abel, S., Glindemann, R., Domahs, F., Regenbrecht, F., Schlenck, K.-J., Thomas, M., Obrig, H., Langen, E. de, Rocker, R., … Bamborschke, S. (2017). Intensive speech and language therapy in patients with chronic aphasia after stroke: A randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. *The Lancet*, *389*(10078), 1528–1538. https://doi.org/10.1016/s0140-6736(17)30067-3

Burke, H. R. (1958). Raven's Progressive Matrices: A Review and Critical Evaluation. *The Journal of Genetic Psychology, 93*(2), 199–228. https://doi.org/10.1080/00221325.1958.10532420

Caute, A., Northcott, S., Clarkson, L., Pring, T., & Hilari, K. (2012). Does mode of administration affect health-related quality-of-life outcomes after stroke? *International Journal of Speech-Language Pathology*, *14*(4), 329–337. https://doi.org/10.3109/17549507.2012.663789

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Chen, H.-C., Koh, C.-L., Hsieh, C.-L., & Hsueh, I.-P. (2013). Test of Everyday Attention in patients with chronic stroke: Test-retest reliability and practice effects. *Brain Injury*, *27*(10), 1148–1154. https://doi.org/10.3109/02699052.2013.775483

de Boer, M. R., Waterlander, W. E., Kuijper, L. D. J., Steenhuis, I. H. M., & Twisk, J. W. R. (2014). Testing for baseline differences in randomized controlled trials: an

unhealthy research behavior that is hard to eradicate. *International Journal of Behavioral Nutrition and Physical Activity*, *12*(1), 4–4. https://doi.org/10.1186/s12966-015-0162-z

DeDe, G., Ricca, M., Knilans, J., & Trubl, B. (2014). Construct validity and reliability of working memory tasks for people with aphasia. *Aphasiology*, *28*(6), 692–712. https://doi.org/10.1080/02687038.2014.895973

Dekhtyar, M., Braun, E. J., Billot, A., Foo, L. & Kiran, S. (2020). Videoconference Administration of the Western Aphasia Battery–Revised: Feasibility and Validity. *American Journal of Speech-language Pathology 29*(2), 673–687.

Dignam, J. K., Copland, D., McKinnon, E., Burfein, P., O'Brien, K., Farrell, A., & Rodriguez, A. (2015). Intensive versus distributed aphasia therapy a nonrandomized, parallel-group, dosage-controlled study. *Stroke*, *46*(8), 2206–2211. https://doi.org/10.1161/strokeaha.115.009522/-/dc1

Duncan, E. S., Schmah, T., & Small, S. L. (2016). Performance Variability as a Predictor of Response to Aphasia Treatment. *Neurorehabilitation and Neural Repair*, *30*(9), 876–882. https://doi.org/10.1177/1545968316642522

Fleming, V., Brownsett, S., Krason, A., Maegli, M. A., Coley-Fisher, H., Ong, Y.-H., Nardo, D., Leach, R., Howard, D., Robson, H., Warburton, E., Ashburner, J., Price, C. J., Crinion, J. T., & Leff, A. P. (2020). Efficacy of spoken word comprehension therapy in patients with chronic aphasia: a cross-over randomised controlled trial with structural imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, *89*, jnnp-2020-324256-7. https://doi.org/10.1136/jnnp-2020-324256

Hilari, K., Lamping, D. L., Smith, S. C., Northcott, S., Lamb, A., & Marshall, J. (2009). Psychometric properties of the Stroke and Aphasia Quality of Life Scale (SAQOL-39) in a generic stroke population. *Clinical Rehabilitation*, *23*(6), 544–557. https://doi.org/10.1177/0269215508101729

Howard, D., & Patterson, K. E. (1992). *The Pyramids and Palm Trees Test: A test of semantic access from words and pictures*. Thames Valley Test Company.

Hula, W., Donovan, N. J., Kendall, D. L., & Gonzalez-Rothi, L. J. (2010). Item response theory analysis of the Western Aphasia Battery. *Aphasiology*, *24*(11), 1326–1341. https://doi.org/10.1080/02687030903422502

Kertesz, A. (2007). *Western Aphasia Battery (Revised)*. PsychCorp.

Lahiri, D., Dubey, S., Ardila, A., Sanyal, D., & Ray, B. K. (2020). Determinants of aphasia recovery: exploratory decision tree analysis. *Language, Cognition and Neuroscience*, *36*(1), 1–8. https://doi.org/10.1080/23273798.2020.1777314

Lambon-Ralph, M. A., Snell, C., Fillingham, J. K., Conroy, P., & Sage, K. (2010). Predicting the outcome of anomia therapy for people with aphasia post CVA: both language and cognitive status are key predictors. *Neuropsychological Rehabilitation*, *20*(2), 289–305. https://doi.org/10.1080/09602010903237875

Lomas, J., Pickard, L., Bester, S., & Elbard, H. (1989). The Communicative Effectiveness Index: Development and Psychometric Evaluation of a Functional

Communication Measure for Adult Aphasia. *Journal of Speech and Hearing Disorders*, *54*, 113–124.

MacWhinney, B., Fromm, D., Holland, A. L., Forbes, M., & Wright, H. (2010). Automated analysis of the Cinderella story. *Aphasiology*, *24*(6–8), 856–868. https://doi.org/10.1080/02687030903452632

Meinzer, M., Streiftau, S., & Rockstroh, B. (2007). Intensive language training in the rehabilitation of chronic aphasia: Efficient training by laypersons. *Journal of the International Neuropsychological Society*, *13*(5), 846–853. https://doi.org/10.1017/s1355617707071111

Menahemi-Falkov, M., Breitenstein, C., Pierce, J. E., Hill, A. J., O'Halloran, R., & Rose, M. L. (2020). *A systematic review of maintenance following intensive therapy programs in chronic post-stroke aphasia: Importance of individual response analysis [Manuscript submitted for publication]*.

Miller, J. F., & Iglesias, A. (2008). *Systematic Analysis of Language Transcripts (SALT), English & Spanish [Computer Software]* (Version 9) [Computer software].

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M., & Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*, *340*(mar23 1), c869–c869. https://doi.org/10.1136/bmj.c869

Nicholas, L. E., & Brookshire, R. H. (1993). A System for Quantifying the Informativeness and Efficiency of the Connected Speech of Adults With Aphasia. *Journal of Speech, Language, and Hearing Research*, *36*(2), 338–350. https://doi.org/10.1044/jshr.3602.338

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M., Bowen, A., Latimer, N., Julious, S., Cross, E., Alshreef, A., Harrison, M., Bradley, E., Witts, H., & Chater, T. (2019). Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): a multicentre, single-blinded, randomised controlled trial. *The Lancet. Neurology*, *18*(9), 821–833. https://doi.org/10.1016/s1474-4422(19)30192-9

Pierce, J. E., Menahemi-Falkov, M., O'Halloran, R., Togher, L., & Rose, M. L. (2017). Constraint and multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation*, *29*(7), 1005–1041. https://doi.org/10.1080/09602011.2017.1365730

Pierce, J. E., O'Halloran, R., Menahemi-Falkov, M., Togher, L., & Rose, M. L. (2020). Comparing higher and lower weekly treatment intensity for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation*, *34*(2), 1–25. https://doi.org/10.1080/09602011.2020.1768127

Plowman, E., Hentz, B., & Ellis, C. (2012). Post-stroke aphasia prognosis: a review of patient-related and stroke-related factors. *Journal of Evaluation in Clinical Practice*, *18*(3), 689–694. https://doi.org/10.1111/j.1365-2753.2011.01650.x

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, *32*(7), 1621–1626. https://doi.org/10.1161/01.str.32.7.1621

RELEASE Collaboration (2020). Predictors of Poststroke Aphasia Recovery. *Stroke, 52*, 1-10. https://doi.org/10.1161/strokeaha.120.031162

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Rose, M. L., Attard, M. C., Mok, Z., Lanyon, L. E., & Foster, A. M. (2013). Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: A phase 1 study. *Aphasiology*, *27*(8), 938–971. https://doi.org/10.1080/02687038.2013.810329

Rose, M. L., Mok, Z., Carragher, M., Katthagen, S., & Attard, M. C. (2015). Comparing multi-modality and constraint-induced treatment for aphasia: a preliminary investigation of generalisation to discourse. *Aphasiology*, 1–21. https://doi.org/10.1080/02687038.2015.1100706

Rose, M. L., Copland, D., Nickels, L., Togher, L., Meinzer, M., Rai, T., Cadilhac, D. A., Kim, J., Foster, A., Carragher, M., Hurley, M., & Godecke, E. (2019). Constraint-induced or multi-modal personalized aphasia rehabilitation (COMPARE): A randomized controlled trial for stroke-related chronic aphasia. *International Journal of Stroke*, *14*(9), 972–976. https://doi.org/10.1177/1747493019870401

Schober, P., & Vetter, T. R. (2019). Correct Baseline Comparisons in a Randomized Trial. *Anesthesia and Analgesia*, *129*(3), 639. https://doi.org/10.1213/ane.0000000000004211

Shewan, C. M., & Kertesz, A. (1980). Reliability and Validity Characteristics of the Western Aphasia Battery (WAB). *Journal of Speech and Hearing Disorders, 45*(3), 308–324. https://doi.org/10.1044/jshd.4503.308

Stahl, B., Mohr, B., Büscher, V., Dreyer, F. R., Lucchese, G., & Pulvermüller, F. (2018). Efficacy of intensive aphasia therapy in patients with chronic stroke: a randomised controlled trial. *Journal of Neurology, Neurosurgery & Psychiatry*, *89*(6), 586–592. https://doi.org/10.1136/jnnp-2017-315962

Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The Apraxia of Speech Rating Scale: a tool for diagnosis and description of apraxia of speech. *Journal of Communication Disorders*, *51*(2014), 43–50. https://doi.org/10.1016/j.jcomdis.2014.06.008

Sutcliffe, L. M., & Lincoln, N. B. (1998). The assessment of depression in aphasic stroke patients: the development of the Stroke Aphasic Depression Questionnaire. *Clinical Rehabilitation*, *12*(6), 506–513. https://doi.org/10.1191/026921598672167702

Thomas, L., Lander, L., Cox, N., & Romani, C. (2020). Speech and language therapy for aphasia: parameters and outcomes. *Aphasiology*, *34*(5), 1–40. https://doi.org/10.1080/02687038.2020.1712588

van de Sandt-Koenderman, W. M. E., van Harskamp, F., Duivenvoorden, H. J., Remerie, S. C., van der Voort-Klees, Y. A., Wielaert, S. M., Ribbers, G. M., & Visch-Brink, E. G. (2008). MAAS (Multi-axial Aphasia System): realistic goal setting in aphasia rehabilitation. *International Journal of Rehabilitation Research*, *31*(4), 314–320. https://doi.org/10.1097/mrr.0b013e3282fc0f23

van der Meulen, I., van de Sandt-Koenderman, W. M. E., Duivenvoorden, H. J., & Ribbers, G. M. (2010). Measuring verbal and non-verbal communication in aphasia: reliability, validity, and sensitivity to change of the Scenario Test. *International Journal of Language & Communication Disorders*, *45*(4), 424–435. https://doi.org/10.3109/13682820903111952

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. https://doi.org/10.1080/17549500902726059

Wallace, S. J., Worrall, L. E., Rose, T., Dorze, G. L., Breitenstein, C., Hilari, K., Babbitt, E., Bose, A., Brady, M., Cherney, L. R., Copland, D., Cruice, M., Enderby, P., Hersh, D., Howe, T., Kelly, H., Kiran, S., Laska, A. C., Marshall, J., … Webster, J. (2019). A core outcome set for aphasia treatment research: The ROMA consensus statement. *International Journal of Stroke*, *14*(2), 180–185. https://doi.org/10.1177/1747493018806200

Wilson, J. T. L., Hareendran, A., Hendry, A., Potter, J., Bone, I., & Muir, K. W. (2005). Reliability of the modified Rankin Scale across multiple raters: benefits of a structured interview. *Stroke; a Journal of Cerebral Circulation*, *36*(4), 777–781. https://doi.org/10.1161/01.str.0000157596.13234.95

Wilssens, I., Vandenborre, D., van Dun, K., Verhoeven, J., Visch-Brink, E., & Marien, P. (2015). Constraint-induced aphasia therapy versus intensive semantic treatment in fluent aphasia. *American Journal of Speech-Language Pathology*, *24*(2), 281–294. https://doi.org/10.1044/2015_ajslp-14-0018

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology*, *23*(11), 1338–1352. https://doi.org/10.1080/02687030902732745

Chapter 6 – Integrative Summary

Overview

This chapter summarises the findings from the work in this thesis and considers their implications for future aphasia research and clinical practice. There were two important aspects of aphasia intervention that were the primary topics of this thesis – comparative effectiveness of treatments (specifically, constraint and multimodal approaches) and the intensity of treatment schedules. These topics were explored within chronic aphasia. Both aspects have been insufficiently studied to date and require more data to guide intervention (Brady et al., 2016). While no one program of work can solve an area of research, this thesis has contributed important findings to both aspects and in the process, highlighted a number of additional issues that will impact future research and practice.

Topic A – Constraint-induced and multimodal interventions compared

Constraint-induced and multimodal approaches were chosen for comparison as prominent and contrasting treatment approaches, within the contexts of rapid growth (constraint) and a long history (multimodal). As outlined in Chapter 1, multimodal approaches have a significant place in the history of aphasia intervention and, for some modalities, an implicit acceptance in much of clinical practice. Research and clinical implementation of constraint approaches have grown in the two decades since they were first applied to aphasia after being translated from interventions for motor impairments.

Chapter 2, a published review, systematically reviewed and analysed outcomes for published research on constraint and multimodal approaches. In the rationale for that review, we stated:

"If patients clearly respond better to being constrained to the verbal modality,

then use of multimodal cues should be re-examined. Conversely, if patients

improve more with multimodal cues, the contribution of verbal constraint in the

CIAT protocol may be questioned." (Pierce et al., 2017, p. 1009)

This logic served as the rationale for the systematic review as well as the study outlined in Chapter 5. However, it failed to outline a third possible outcome – that neither approach is found to have superior outcomes. The results of the meta-analysis of Single Case Experimental Designs did not find strong evidence that one approach better improved naming of treated items, while there was insufficient high-quality data to accurately compare other impairment, activity/participation or quality of life outcomes. The results of the RCT in Chapter 5 also found neither treatment to be superior, with comparison of CIAT Plus and M-MAT revealing strikingly similar outcomes – in fact, there were no significant between-group differences for any outcomes or timepoints. Taken together, results from Chapter 2 and Chapter 5 show that there is no evidence for constraint or multimodal treatment being superior.

As described in Chapter 5, participant experiences of these treatments vary. In two previous studies, participants underwent both CIAT Plus and M-MAT in a crossover fashion and some reported a preference for one treatment over the other (Attard et al., 2013; Rose et al., 2013). Like or dislike of visual barriers in CIAT Plus influenced some decisions, others liked the opportunity M-MAT provided to use non-speech modalities, and some reported that they appreciated the increased speech practice in CIAT Plus. Maher et al. (2006) reported a comparison of a constraint treatment (CILT) to Promoting Aphasic Communicative Effectiveness (PACE) in which 2/4 CILT participants

were frustrated at being unable to write and 2/5 PACE participants were opposed to using non-speech modalities, with one confining his responses to speech only. In my own experience of conducting CIAT Plus within this trial, one participant repeatedly requested to use pencil and paper as this was a method of self-cueing that he regularly used in communication. These examples are evidence that some people with aphasia will have strong preferences for or against multimodal treatments, and for or against constraint. Based on the findings in this thesis, one important clinical implication is that speech pathologists need not be concerned that one treatment is inferior to the other. Instead, they may consider other factors if choosing one of these treatment approaches for chronic aphasia; for example, patient preferences, or any natural tendency to use non-speech modalities. Formal qualitative research into participant experiences of both treatment approaches have not been conducted but could provide further guidance for selecting a treatment.

In the introduction (Chapter 1) it was argued that each approach can be framed as a different understanding of language representation in the brain; broadly, that multimodal approaches suggest synergy between speech and other communication networks, while constraint approaches imply that use of non-speech networks interfere with neuroplastic change. Our results suggest that neither model is adequately descriptive of language representation and recovery, or that both are somehow true.

In summary, the work in this thesis has shown that there is no evidence for superior efficacy of either CIAT Plus or M-MAT.

The potency of individual components of intervention are not yet known

"These similar results in direct comparisons [of CIAT Plus and M-MAT] are puzzling given the contrasting nature of the cues provided." (Pierce et al., 2017, p. 1010)

Our scoping review of the definitions of "multimodal therapy" (Chapter 3) described interventions in terms of whether non-speech modalities were employed in (a) stimuli, (b) participant response or (c) cueing (Pierce et al., 2019). These three aspects of treatment would differentiate a large number of interventions in aphasia, yet it has not been shown that one or more of these is a key ingredient of treatment. A major premise behind many multimodal interventions, and of this thesis, is that the multimodal *cueing* is a potent component. For this reason, M-MAT includes a rich combination of modalities in its cueing hierarchy, including gesture, drawing, reading and writing, and yet no advantage was found compared to the lean cueing hierarchy in CIAT Plus, which uses only reading and spoken repetition. Conversely, a key rationale for most constraint therapies is that participants should avoid other modalities for communicating and/or self-cueing, yet the systematic implementation of other modalities in M-MAT did not negatively affect outcomes.

The equivalent effects of multimodal cueing and constraint found in our review (Chapter 2) and experimental trial (Chapter 5) could suggest that, (a) both cueing designs are equally effective, or (b) cueing is not the crucial component of these interventions. The treatments share other features that could be more important in improving confrontational naming. For example, both treatments include presentation of the written target word as a cueing step. A review by Sze, Hameau, Warren and Best (2020) conducted a random forest analysis of treatment components within single word production studies, incorporating data from 222 participants in 32 papers. They found that presentation of the written target as a cue, either in whole or part, was most consistently predictive of positive outcomes, regardless of whether the participant copied the word out or not. While their analysis excluded constraint and multimodal

treatments, the effect of presenting the written word might extend to these approaches nonetheless. Verbal repetition of the target was also common to both CIAT Plus and M-MAT. More simply still, targets were presented repeatedly over 30 hours. Repeated exposure to words without training has shown large effect sizes in past analyses, only marginally smaller than treated words (Wisenburn & Mahoney, 2009).

Results in this thesis highlight the need to systematically describe and compare elements of interventions in aphasia to determine the most potent aspect(s). These elements may not be those predicted based on clinical reasoning.

Interventions should be reported in detail

In order to examine the contribution of individual elements of treatment as described above, comprehensive descriptions of treatment are required. Reporting of aphasia interventions has lacked sufficient detail to date (RELEASE Collaboration, 2020) and this limits the ability to conduct secondary analyses.

The findings of this thesis established a lack of detail in treatment reports for constraint and multimodal approaches. The review in Chapter 2 found that the application of constraint was poorly described in the majority of studies despite being the namesake element of the treatment. Further, for those that did describe constraint, some restricted use of hands while some allowed gesture. Differing interpretations of constraint is less problematic than not reporting in sufficient detail but does highlight that the label of 'constraint-induced' does not reliably describe methods.

Descriptions of multimodal interventions were also problematic. The ambiguity of the term 'multimodal' was noted during the search process for the Chapter 2 review and led to a scoping review of the term that was not originally anticipated for this thesis (Chapter 3). While most studies did report key procedural aspects of treatment, a

plethora of different treatments with little in common all used the label 'multimodal.'
We proposed a taxonomy that differentiates treatments by primary aim and rationale,
but even if this is taken up by the research field, the labels remain broad, and
comprehensive descriptions of intervention will still be needed.

A number of frameworks for describing interventions have been developed or
are in development and these will advance the ability to compare and analyse such
elements. The TIDieR checklist (Hoffmann et al., 2014) is one such reporting method,
though recently its utility for allowing replication of aphasia interventions has been
questioned, given the complexity of aphasia therapy (Dipper et al., 2021). The
Rehabilitation Treatment Taxonomy Specification System (RTTSS) has been developed
with the explicit aim of discerning the active elements of a treatment within
rehabilitation by reporting the intervention based on the underlying rationale (Stan et
al., 2018; Turkstra et al., 2016). Third, the Intervention Taxonomy describes both
intervention and delivery aspects of treatments (Schulz et al., 2010) and has been used
to compare interventions for communication impairments (O'Rourke et al., 2018).

While we are not the first to call for more thorough reporting of intervention
protocols, our work cautions against relying on therapy labels or categories even in
established treatments.

Topic B – Intensity

The topic of intensity, or how often treatment should be provided, is a high
priority for patients, clinicians and service providers (Bayley et al., 2007). Stroke
rehabilitation guidelines typically recommend high intensity treatment for aphasia
where possible, yet "high intensity" is not defined and no specifics on times or

frequency are provided (Dignam et al., 2016). The assumption that higher intensity

treatment will result in greater neurological change is based on principles of

neuroplasticity which are derived from animal models and studies of motor tasks (Kleim

& Jones, 2008). However, it is not clear that language recovery is analogous to motor

recovery within the brain. The contrasting paradigm comes from cognitive psychology

where distributed learning has been shown to be superior for retention in a range of

motor, cognitive and language tasks (Dignam et al., 2016). Its application to aphasia

relies on the premise that rehabilitation is a form of learning (Dignam et al., 2016).

International surveys consistently report that clinical practice comprises very

low intensity therapy in the chronic phase: around 1-2 hours per week (e.g., Palmer et

al., 2018; Verna et al., 2009). This is probably due to practical and financial reasons

rather than clinicians favouring cognitive psychology over neuroplasticity[3]. As outlined

in the introduction for the systematic review of weekly intensity (Chapter 4), the topic

of intensity has widespread ramifications for clinicians, patients, service delivery models

and funding. Strong evidence of substantial benefit to people with aphasia, families and

society would likely be needed in order to change current service delivery, given the

constraints on healthcare resources (Code & Petheram, 2011).

Our systematic review did not find strong evidence for higher or lower intensity

of treatment in chronic aphasia. Overall, there was a lack of studies that carefully

examined this aspect of scheduling and there remains a need for further work on the

topic. Our trial was not a direct comparison of lower to higher intensity but adds data to

---

[3] My personal observation is that clinicians feel a sense of regret or even guilt that they cannot
provide high intensity services; there is a widespread assumption that higher intensity has
already been demonstrated to be the gold standard.

suggest that low-moderate intensity therapy can still be efficacious for confrontational naming. Results were grossly consistent with confrontational naming improvements in trials of high intensity treatment, in terms of change in correctly named items. Future comparison of this data with the main COMPARE trial, when results are released, will allow calculation of the effects of lower weekly intensity with other factors controlled.

It is possible that weekly intensity in aphasia may be more complex than previously assumed and this would not be the first time that counter-intuitive or surprising results have occurred in stroke rehabilitation. An RCT of acute stroke treatments found no advantage of a higher dose of aphasia therapy in the first twelve weeks compared to lower dose standard treatment (Godecke et al., 2020), challenging a straightforward interpretation of the neuroplasticity principle of intensity. Another RCT found less favourable outcomes following early, intensive mobilisation of stroke patients compared to usual physiotherapy care (AVERT Trial Collaboration group et al., 2015), supporting the principle that *time matters*, but not in a straightforward manner where *earlier is better* (Kleim & Jones, 2008).

Overall, this work encourages further questioning of the clinical assumption that higher intensity is better.

Treatment Generalisation should not be Assumed

Generalisation to Untreated Items

The untreated items of the naming battery did not show statistically significant improvement, consistent with most other naming interventions in aphasia. For

example, a review of 44 naming studies found the highest mean effect size for trained words while unexposed words showed minimal effect sizes (Wisenburn & Mahoney, 2009). This result accords with the neuroplasticity principle of *specificity*, which is based on findings that brain changes after training are highly region-specific (Kleim & Jones, 2008). Based on this principle, addressing word retrieval and sentence production using specific targets, as in CIAT Plus and M-MAT, would not be expected to improve untreated words. However, close 'neighbours' may show some change. For example, generalisation to untrained items within the same semantic category has been demonstrated in aphasia therapy (Kendall et al., 2019; Raymer et al., 2008), as has improvement in syntactically similar sentences following treatment for sentence comprehension and production (Raymer et al., 2008). In our trial, semantically related untrained words were not investigated and this could explain the lack of generalisation. We also did not investigate sentence production specifically using the syntax used in the treatment. Some researchers have looked at the ability to request objects—an element of the Go Fish game in CIAT—following Intensive Language Action Therapy (a form of CIAT) and did find significant improvement (Stahl et al., 2018). Future research may need to look for improvement in appropriately related targets or sentence structures following intervention, rather than unrelated, untrained items.

Generalisation to Impaired Systems

Given that the untreated items of the naming battery did not show significant improvement, is it reasonable to predict improvement on one or more subtests of the WAB-R or picture description tasks, when these were not the targets of therapy? From the perspective of specificity, there would be little reason to expect a change in

impairment level outcomes unless the tasks within the test coincided with the items or sentence structure targeted in therapy.

CIAT Plus and M-MAT were hypothesised to cause a rising tide that would lift all boats (i.e., a change to the impaired language system itself), not only the lifting of particular boats (i.e., item-specific improvement). In other words, it was hoped that language processing ability itself would improve and this would be evident on an aphasia battery.

The hypothesis was not without precedent. Some past high-quality papers have reported significant changes after constraint-induced aphasia therapy on another omnibus aphasia battery, the Aachen Aphasia Test (Meinzer et al., 2007; Pulvermüller et al., 2001; Stahl et al., 2018; Wilssens et al., 2015), while others have reported varying levels of individual change on the WAB-R and Boston Naming Test following CIAT Plus or M-MAT in smaller studies (Attard et al., 2013; Rose et al., 2013). More recently, a trial of high intensity constraint-induced aphasia therapy on 17 participants showed significant improvement on the WAB-R AQ (p = .001) (Heikkinen et al., 2019). Nevertheless, in our experiment, the WAB-R AQ failed to change at the group level.

There are several factors that might explain the discrepant findings to date. First, most high-quality studies with positive findings have used the Aachen Aphasia Test. The equivalence of the WAB-R and the AAT has not been demonstrated. Although they appear to test similar language functions, the psychometric properties likely differ and one may be over- or under-sensitive to change. Second, the dual baseline design employed in our experiment highlighted the variability of individual test scores, not only for the WAB-R but for most outcomes. This will cause Type I and Type II errors at times, depending on the sample size and the random factor of participant performance.

Finally, as discussed in Chapter 5, it is possible that the lower intensity of our trial was insufficient to affect change in the language system itself. Heikkinen et al., (2019) did demonstrate statistically significant improvements on the WAB AQ following high intensity CIAT Plus, though the mean change was small (3.4) and below the threshold for clinically significant change. As Harnish et al. (2008, p. 469) described: "...perhaps the chronically injured brain requires a jolt to the system to alter established pathways for processing language, as opposed to mild stimulation provided over a longer period of time." Six hours per week may have been below the theorised threshold that would stimulate broad neural reorganisation (Dignam et al., 2016; Harnish et al., 2008; Kleim & Jones, 2008), although our systematic review (Chapter 4) did not suggest any simple cut-off exists.

Our work raises questions about whether system-level language improvements occur in chronic aphasia following constraint-induced and multimodal interventions, and whether there is a minimum intensity required to achieve this. The data from our experiment will also serve as a valuable comparison to the COMPARE main trial in terms of the contribution of intensity.

Generalisation to Function and Quality of Life

Generalisation from therapy to a person's daily communicative activity, function and then quality of life is the ultimate goal of treatment in aphasia (Carragher et al., 2013). Yet, as our review found in Chapter 2, most studies of constraint and multimodal therapies have not measured outcomes at activity/participation and quality of life levels. This accords with aphasia research more broadly where there is a lack of investigation on these outcomes (Doedens & Meteyard, 2019).

Our experiment measured a number of outcomes beyond impairment-level, including proxy rated conversation function (CETI), communication function in simulated situations (Scenario Test), and self-reported quality of life (SAQOL-39). The CETI showed possible improvement while none of the remaining outcomes improved significantly at a group level.

Continuing with the hypothesis that the language system itself was not changed but only some of its inventory, it would be a steep ascent from the 80 nouns and verbs and the specific sentence structures trained to affecting real conversational ability and subsequently self-reported quality of life. The lack of transfer from 'brain training' apps to real life cognition may be an apt comparison (Simons et al., 2016).

However, there are several features of CIAT Plus and M-MAT that are designed to encourage generalisation to function. Word retrieval and production occurs within sentences rather than in isolation, the utterances form part of a game involving turn taking and responsivity in which they are communicative rather than drill-based, and participants are assigned home transfer tasks. It is also puzzling that the non-speech modalities trained in M-MAT (gesture, drawing, writing) did not result in improved scores on the Scenario Test compared to CIAT Plus.

Our review in Chapter 3 found mixed results when it came to outcomes of activity/participation. For both constraint and multimodal interventions, mean changes for the CETI were often less than the clinically significant difference of 12 despite statistical significance, although some individual changes ($n = 9$, B0 to Post) did exceed 12 points. Mean changes for the Scenario Test were also small, again with a wide range of individual changes. Other activity/participation measures used did show statistically significant improvement at a group level, including the Communicative Activity Log

(Pulvermüller et al., 2001) and the Amsterdam–Nijmegen Everyday Language Test

(Wilssens et al., 2015). Quality of Life data was not able to be extracted for individual

treatments in any high quality studies we reviewed but following administration of both

CIAT Plus and M-MAT in a crossover fashion, Rose et al. (2013) reported improvements

on SAQOL-39 subscales for some participants.

Two large RCTs of intervention for chronic aphasia provide conflicting results on

generalisation to function. Breitenstein et al. (2017) used the Amsterdam–Nijmegen

Everyday Language Test as the primary outcome measure and this improved

significantly following 22 to 49 hours of therapy. In contrast, Palmer et al. (2019) did not

see effects on functional communication (Therapy Outcome Measures) or self-ratings of

participation and quality of life following a mean of 28 hours of computer-based

therapy. One key difference is that Breitenstein et al. provided their treatment over

three weeks whereas the treatment duration of Palmer et al. was six months – an

average of 2.3 hours per week. Intensity may therefore play a role in carryover to

activity/participation and quality of life. However, the treatments in these studies were

also different and used different approaches to encouraging generalisation. In

Breitenstein et al., the intervention was partially focused on progressing through a

hierarchy of communicative functions (e.g., communicating personal information), with

targets drawn from baseline performance on those tasks. In Palmer et al., therapy

involved single word naming of 100 personally relevant words using software and the

protocol specified that volunteers would visit participants monthly to practise use of

target items in conversation. However, on average only 45 minutes of total

conversational practice was provided by volunteers over the six months. These transfer

tasks could be an important component for functional generalisation and require further research.

Both CIAT Plus and M-MAT might be modified to further encourage generalisation to function. The naming battery was developed to cover three levels of difficulty whilst balancing psycholinguistic variables, but the general frequency of words in a corpus is not necessarily representative of their relevance for an individual (Renvall et al., 2013). Selecting items that are personally relevant to the individual should ensure that practice of word retrieval will arise naturally and is particularly important given that the untreated items did not significantly change. The syntax of the target sentence might also be adjusted. While SVO structures, as used in the trial, are the most common structure in English (Roland et al., 2007), it could be that present tense, which was trained in CIAT Plus and M-MAT, is less common than past tense. Future research should investigate ways to personalise these treatments and whether such personalisation results in stronger generalisation.

Finally, given that generalisation in our study was limited – to untreated naming items, language systems and activity/participation and quality of life – a natural question is how valuable the treatments in our study were to people with aphasia. Longitudinal follow-up research of participants at, for example, one year after intervention would help answer this question.

Assessment performance and treatment response vary in aphasia

Intra-individual variability

The inclusion of two baseline points in our RCT in Chapter 5 is not a common design element but it provided a unique insight into the stability of the outcome measures prior to treatment. Although group means show little or no change on measures between baselines, when individual changes are charted (Figure 5-4), some individual participants show anomalies in scores. For example, in the treated items, one participant (uppermost orange) names eight more items from Baseline 1 compared to Baseline 0, only to drop down again at Post Intervention. Another (second orange) declines stepwise from 62 to 52 to 44. Thirdly, a participant (second green) scores an anomalously low score at Baseline 1 (-20) with the other three data points in line with the typical participant trend. In fact, all outcomes in Figure 5-4 have at least a small number of participants who show 'bounces' between baselines, a time period when no change was expected. These anomalies appear to be random, as they occur as both increases and decreases and cancel out when the group mean is calculated. They are therefore unlikely to represent true changes in aphasia severity (WAB-R AQ) and naming ability.

However, temporary drops or boosts in performance on an assessment is a feasible explanation, as aphasia is known to be highly variable within the individual (Nespoulous, 2000). In my clinical experience, I have had many, many patients with aphasia complain about the unpredictability of their language function. Variation across a one-hour session, across a day or longer periods is a source of frustration and could explain the individual variability found across baselines. In other words, a portion of the individual changes are due not to measurement error nor change in the function being

measured, but natural variability of participant performance. Possible causes of such variability might include the level of fatigue or stimulation, psychosocial stressors not disclosed to the assessors, performance anxiety, motivation, etc. Language performance in people with aphasia may be more vulnerable to such stressors, given the system is already impaired.

Inter-individual Variability

By examining individual changes between Baseline 0 and Post Intervention for the treated items of the naming battery, it is clear that the overall finding of a mean improvement is supported, even after accounting for intra-individual variability. However, it is also clear that a small number of participants do not show improvement after intervention. Such outliers do not nullify the results of the trial, as group designs are designed to examine mean group effects, but understanding them is important in applying treatments to the individual, whether in clinical practice or future aphasia research. Group level research has limitations when it comes to individuals, as the mean response across a group of participants is not sufficient to determine the clinician's choice of intervention for an individual in front of them (Menahemi-Falkov et al., 2020). However, data on what predicts an individual's response can help with these decisions.

The baseline characteristics measured in our trial did not correlate with changes on the WAB-R AQ or the naming battery but this is not conclusive – instead, it may be an issue of statistical power, considering that predictive factors likely interact. We used only demographics and cognitive and linguistic characteristics in examining possible relationships, but other proposed predictors include the psychosocial environment, age at time of stroke, lesion volume, lesion location and impact on connectivity and

perfusion characteristics (Kiran & Thompson, 2019). Discovering which treatment works

best for whom, and even which treatment *at which intensity* works best for whom

(Brady et al., 2016), requires so many variables that it is difficult to see a way to conduct

such research using traditional experimental designs. Large databases are being

established to predict the treatment most likely to be effective for a given patient based

on retrospective analysis (Boyd et al., 2017; Price et al., 2010). By measuring a range of

characteristics at baseline, we now have data that can contribute to such analyses and

to personalised treatments in future.

This work has clearly highlighted the intra- and inter-individual variability—or

variability in assessment performance and treatment response—present within chronic

aphasia. It demonstrates the threat of intra-individual variability to studies with small

sample sizes, which may draw inappropriate conclusions based on random error. In

contrast, the variable treatment response *between* individuals may result in Type II

errors: A treatment may be effective for a subset of participants in a trial who share

certain predictive characteristics, while the mean response is non-significant.

Distinguishing variable performance from treatment response is a difficult question but

some promising methods of calculating change thresholds for individuals and groups

have been recently applied to aphasia (Menahemi-Falkov et al., 2020). Interestingly,

some authors have suggested that intra-individual variability is not simply noise but may

itself be an important predictor of treatment response (Duncan et al., 2016).

Aphasia intervention research should avoid small group samples and instead use

large, powered sample sizes to account for variability. Alternatively, use of high quality

Single Case Experimental Designs allows the researcher to elucidate baseline variance (Goldstein, 2014).

Gains from Intervention in Chronic Aphasia Show Decline at Follow-Up

The data from our RCT show a clear reduction of naming scores at the 12 week Follow-Up and this was statistically significant compared to Post Intervention, though also significantly higher than baseline. Approximately half of the gains were lost by this timepoint in terms of mean number of items named. Additional follow-up assessments may show that gains return to baseline levels. Our results are consistent with other reports of follow-up assessments (Menahemi-Falkov et al., 2020). There are many activities which might support the maintenance of gains from treatment, including self-practice, additional therapy and aphasia groups (Menahemi-Falkov et al., 2020). Additional therapy has been shown to be beneficial even if it is a different treatment or provided at a different schedule (Menahemi-Falkov et al., 2020).

When implementing treatments such as CIAT Plus or M-MAT, clinicians should inform people with aphasia about the likely attenuation of gains in the weeks and months following intervention and the need for additional therapy and/or maintenance activities in order to avoid it. There may be a 'willingness to pay' threshold in terms of patient time exchanged for gains, or this may vary between people. If more longitudinal follow-up datapoints were available, the 'forgetting curve' could be calculated as a power function (Wixted & Ebbesen, 1997) which would allow visualisation of the drop-off and a reasonably accurate prediction of the eventual return to baseline. This data might help people with aphasia decide on treatment. Other factors would also probably influence patient decision making on the acceptability of treatments; for example,

interventions that clearly impact functional communication and/or quality of life may have people with aphasia more willing to commit to intervention.

Conclusion

In summary, this thesis has expanded the evidence base for constraint and multimodal approaches to aphasia intervention, highlighting gaps in conceptualisation, terminology, reporting, outcome measurement and methodological rigour, and importantly, it produced data that demonstrates that neither approach is superior for confrontational naming. It has explored the effects of intensity in constraint and multimodal approaches in chronic aphasia and revealed that a less intensive dose of 30 hours over 5 weeks is effective for naming outcomes. This work has also outlined further work to be done in measuring outcomes of intervention. Future research needs to account for intra-individual variability across assessments and inter-individual variability in response to treatment. Interventions should also be designed with the generalisation, maintenance and impact of outcomes as a primary goal.

These findings contribute to the incremental progress in aphasia treatment knowledge, with the hope that each advancement brings the field closer to improving the lives of people with aphasia.

References

Attard, M. C., Rose, M. L., & Lanyon, L. E. (2013). The comparative effects of Multi-Modality Aphasia Therapy and Constraint-Induced Aphasia Therapy-Plus for severe chronic Broca's aphasia: An in-depth pilot study. *Aphasiology*, *27*(1), 80–111. https://doi.org/10.1080/02687038.2012.725242

AVERT Trial Collaboration group, Bernhardt, J., Langhorne, P., Lindley, R. I., Thrift, A. G., Ellery, F., Collier, J., Churilov, L., Moodie, M., Dewey, H., & Donnan, G. (2015). Efficacy and safety of very early mobilisation within 24 h of stroke onset (AVERT): A randomised controlled trial. *The Lancet*, *386*(9988), 46–55. https://doi.org/10.1016/s0140-6736(15)60690-0

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders, 36*(3), 189–208. https://doi.org/10.1016/s0021-9924(03)00019-4

Bayley, M. T., Hurdowar, A., Teasell, R., Wood-Dauphinee, S., Korner-Bitensky, N., Richards, C. L., Harrison, M., & Jutai, J. W. (2007). Priorities for stroke rehabilitation and research: Results of a 2003 Canadian Stroke Network consensus conference. *Archives of Physical Medicine and Rehabilitation*, *88*(4), 526–528. https://doi.org/10.1016/j.apmr.2007.01.005

Boyd, L. A., Hayward, K. S., Ward, N. S., Stinear, C. M., Rosso, C., Fisher, R. J., Carter, A. R., Leff, A. P., Copland, D. A., Carey, L. M., Cohen, L. G., Basso, D. M., Maguire, J. M., & Cramer, S. C. (2017). Biomarkers of stroke recovery: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. *International Journal of Stroke*, *12*(5), 480–493. https://doi.org/10.1177/1747493017714176

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *The Cochrane Database of Systematic Reviews*, *6*(6), CD000425–CD000425. https://doi.org/10.1002/14651858.cd000425.pub4

Breitenstein, C., Grewe, T., Flöel, A., Ziegler, W., Springer, L., Martus, P., Huber, W., Willmes, K., Ringelstein, E. B., Haeusler, K. G., Abel, S., Glindemann, R., Domahs, F., Regenbrecht, F., Schlenck, K.-J., Thomas, M., Obrig, H., de Langen, E., Rocker, R., … Bamborschke, S. (2017). Intensive speech and language therapy in patients with chronic aphasia after stroke: a randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. *The Lancet*, *389*(10078), 1528–1538. https://doi.org/10.1016/s0140-6736(17)30067-3

Carragher, M., Sage, K., & Conroy, P. (2013). The effects of verb retrieval therapy for people with non-fluent aphasia: Evidence from assessment tasks and conversation. *Neuropsychological Rehabilitation*, *23*(6), 846–887. https://doi.org/10.1080/09602011.2013.832335

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

Dignam, J. K., Rodriguez, A. D., & Copland, D. (2016). Evidence for Intensive Aphasia Therapy: Consideration of Theories from Neuroscience and Cognitive Psychology. *American Academy of Physical Medicine and Rehabilitation*, *8*(3), 254–267. https://doi.org/10.1016/j.pmrj.2015.06.010

Dipper, L. T., Franklin, S., Aguiar, V. de, Baumgaertner, A., Brady, M., Best, W., Bruehl, S., Denes, G., Godecke, E., Gil, M., Kirmess, C., Markey, M., Meinzer, C., Orellana, M. M., Norvik, M., Nouwens, F., Rose, M. L., van de Sandt, M., Whitworth, A., & Visch-Brink, E. G. (2021). An Umbrella Review of Aphasia Intervention descriPtion In Research: the AsPIRE project. *Aphasiology*, 1–26. https://doi.org/10.1080/02687038.2020.1852001

Doedens, W. J., & Meteyard, L. (2019). Measures of functional, real-world communication for aphasia: a critical review. *Aphasiology*, *34*(4), 1–23. https://doi.org/10.1080/02687038.2019.1702848

Duncan, E. S., Schmah, T., & Small, S. L. (2016). Performance Variability as a Predictor of Response to Aphasia Treatment. *Neurorehabilitation and Neural Repair*, *30*(9), 876–882. https://doi.org/10.1177/1545968316642522

Godecke, E., Armstrong, E., Rai, T., Ciccone, N., Rose, M. L., Middleton, S., Whitworth, A., Holland, A., Ellery, F., Hankey, G. J., Cadilhac, D. A., Bernhardt, J., & Group, V. C. (2020). A randomized control trial of intensive aphasia therapy after acute stroke: The Very Early Rehabilitation for SpEech (VERSE) study. *International Journal of Stroke*, 174749302096192. https://doi.org/10.1177/1747493020961926

Goldstein, H. (2014). Throwing the baby out with the bathwater: Pitfalls of misrepresenting single-case experimental designs. *Aphasiology, 29*(5), 583–587. https://doi.org/10.1080/02687038.2014.987045

Harnish, S. M., Neils-Strunjas, J., Lamy, M., & Eliassen, J. C. (2008). Use of fMRI in the study of chronic aphasia recovery after therapy: a case study. *Topics in Stroke Rehabilitation*, *15*(5), 468–483. https://doi.org/10.1310/tsr1505-468

Heikkinen, P. H., Pulvermüller, F., Mäkelä, J. P., Ilmoniemi, R. J., Lioumis, P., Kujala, T., Manninen, R.-L., Ahvenainen, A., & Klippi, A. (2019). Combining rTMS With Intensive Language-Action Therapy in Chronic Aphasia: A Randomized Controlled Trial. *Frontiers in Neuroscience, 12*, 1036. https://doi.org/10.3389/fnins.2018.01036

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D. G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S. E., Dixon-Woods, M., McCulloch, P., Wyatt, J. C., Chan, A.-W., & Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, *348*, g1687–g1687. https://doi.org/10.1136/bmj.g1687

Kendall, D. L., Moldestad, M. O., Allen, W., Torrence, J., & Nadeau, S. E. (2019). Phonomotor Versus Semantic Feature Analysis Treatment for Anomia in 58 Persons With Aphasia: A Randomized Controlled Trial. *Journal of Speech, Language, and Hearing Research*, *62*(12), 4464–4482. https://doi.org/10.1044/2019_jslhr-l-18-0257

Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of Language Networks in Aphasia: Advances, Updates, and Future Challenges. *Frontiers in Neurology*, *10*, S225. https://doi.org/10.3389/fneur.2019.00295

Kleim, J. A., & Jones, T. A. (2008). Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research*, *51*(1), S225-39. https://doi.org/10.1044/1092-4388(2008/018)

Maher, L. M., Kendall, D. L., Swearengin, J. A., Rodriguez, A., Leon, S. A., Pingel, K., Holland, A. L., & J, G.-R., Leslie. (2006). A pilot study of use-dependent learning in the context of Constraint Induced Language Therapy. *Journal of the International Neuropsychological Society*, *12*(6), 843–852. https://doi.org/10.1017/s1355617706061029

Meinzer, M., Streiftau, S., & Rockstroh, B. (2007). Intensive language training in the rehabilitation of chronic aphasia: Efficient training by laypersons. *Journal of the International Neuropsychological Society*, *13*(5), 846–853. https://doi.org/10.1017/s1355617707071111

Menahemi-Falkov, M., Breitenstein, C., Pierce, J. E., Hill, A. J., O'Halloran, R., & Rose, M. L. (2020). *A systematic review of maintenance following intensive therapy programs in chronic post-stroke aphasia: Importance of individual response analysis [Manuscript submitted for publication]*.

Nespoulous, J.-L. (2000). Invariance vs Variability in Aphasic Performance. An Example: Agrammatism. *Brain and Language*, *71*(1), 167–171. https://doi.org/10.1006/brln.1999.2242

O'Rourke, A., Power, E., O'Halloran, R., & Rietdijk, R. (2018). Common and distinct components of communication partner training programmes in stroke, traumatic brain injury and dementia. *International Journal of Language & Communication Disorders*, *53*(6), 1150–1168. https://doi.org/10.1111/1460-6984.12428

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M., Bowen, A., Latimer, N., Julious, S., Cross, E., Alshreef, A., Harrison, M., Bradley, E., Witts, H., & Chater, T. (2019). Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): a multicentre, single-blinded, randomised controlled trial. *The Lancet. Neurology*, *18*(9), 821–833. https://doi.org/10.1016/s1474-4422(19)30192-9

Palmer, R., Witts, H., & Chater, T. (2018). What speech and language therapy do community dwelling stroke survivors with aphasia receive in the UK? *PLOS ONE*, *13*(7), e0200096. https://doi.org/10.1371/journal.pone.0200096

Pierce, J. E., Menahemi-Falkov, M., O'Halloran, R., Togher, L., & Rose, M. L. (2017). Constraint and multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation*, *29*(7), 1005–1041. https://doi.org/10.1080/09602011.2017.1365730

Pierce, J. E., O'Halloran, R., Togher, L., & Rose, M. L. (2019). What Is Meant by "Multimodal Therapy" for Aphasia? *American Journal of Speech-Language Pathology*, *28*(2), 706–716. https://doi.org/10.23641/asha.7646717

Price, C. J., Seghier, M. L., & Leff, A. P. (2010). Predicting language outcome and recovery after stroke: the PLORAS system. *Nature Reviews Neurology*, *6*(4), 202–210. https://doi.org/10.1038/nrneurol.2010.15

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, *32*(7), 1621–1626. https://doi.org/10.1161/01.str.32.7.1621

Raymer, A. M., Beeson, P., Holland, A. L., Kendall, D. L., Maher, L. M., Martin, N., Murray, L., Rose, M. L., Thompson, C. K., & Turkstra, L. (2008). Translational research in aphasia: From neuroscience to neurorehabilitation. *Journal of Speech, Language, and Hearing Research*, *51*, S259–S275. https://doi.org/10.1044/1092-4388(2008/020)

RELEASE Collaboration. (2020). Communicating simply, but not too simply: Reporting of participants and speech and language interventions for aphasia after stroke. *International Journal of Speech-Language Pathology*, *22*(3), 302–312. https://doi.org/10.1080/17549507.2020.1762000

Renvall, K., Nickels, L. A., & Davidson, B. (2013). Functionally relevant items in the treatment of aphasia (part I): Challenges for current practice. *Aphasiology*, *27*(6), 636–650. https://doi.org/10.1080/02687038.2013.786804

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*(3), 348–379. https://doi.org/10.1016/j.jml.2007.03.002

Rose, M. L., Attard, M. C., Mok, Z., Lanyon, L. E., & Foster, A. M. (2013). Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: A phase 1 study. *Aphasiology*, *27*(8), 938–971. https://doi.org/10.1080/02687038.2013.810329

Schulz, R., Czaja, S. J., McKay, J. R., Ory, M. G., & Belle, S. H. (2010). Intervention Taxonomy (ITAX): Describing Essential Features of Interventions. *American Journal of Health Behavior*, *34*(6). https://doi.org/10.5993/ajhb.34.6.15

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do "Brain-Training" Programs Work? *Psychological Science in the Public Interest*, *17*(3), 103–186. https://doi.org/10.1177/1529100616661983

Stahl, B., Mohr, B., Büscher, V., Dreyer, F. R., Lucchese, G., & Pulvermüller, F. (2018). Efficacy of intensive aphasia therapy in patients with chronic stroke: a randomised controlled trial. *Journal of Neurology, Neurosurgery & Psychiatry*, *89*(6), 586–592. https://doi.org/10.1136/jnnp-2017-315962

Stan, J. H. V., Dijkers, M. P., Whyte, J., Hart, T., Turkstra, L. S., Zanca, J., & Chen, C. (2018). The Rehabilitation Treatment Specification System: Implications for improvements in research design, reporting, replication, and synthesis. *Archives*

of Physical Medicine and Rehabilitation, *100*(1), 146–155. https://doi.org/10.1016/j.apmr.2018.09.112

Sze, W. P., Hameau, S., Warren, J., & Best, W. (2020). Identifying the components of a successful spoken naming therapy: a meta-analysis of word-finding interventions for adults with aphasia. *Aphasiology*, *38*(1), 1–40. https://doi.org/10.1080/02687038.2020.1781419

Turkstra, L. S., Norman, R., Whyte, J., Dijkers, M. P., & Hart, T. (2016). Knowing What We're Doing: Why Specification of Treatment Methods Is Critical for Evidence-Based Practice in Speech-Language Pathology. *American Journal of Speech-Language Pathology*, *25*(2), 164–171. https://doi.org/10.1044/2015_ajslp-15-0060

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. https://doi.org/10.1080/17549500902726059

Wilssens, I., Vandenborre, D., van Dun, K., Verhoeven, J., Visch-Brink, E., & Marien, P. (2015). Constraint-induced aphasia therapy versus intensive semantic treatment in fluent aphasia. *American Journal of Speech-Language Pathology*, *24*(2), 281–294. https://doi.org/10.1044/2015_ajslp-14-0018

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology*, *23*(11), 1338–1352. https://doi.org/10.1080/02687030902732745

Wixted, J. T., & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition*, *25*(5), 731–739. https://doi.org/10.3758/bf03211316

Appendices

Appendix 1 – Summary of all eligible papers for Chapter 2

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
Constraint	Level 2 - Randomised trial	(Pulvermüller et al., 2001)	Constraint-Induced Therapy of Chronic Aphasia After Stroke	CIAT *vs.* Standard therapy, less intensive	17 (10 constraint)	Impairment Activity/Participation
Constraint	Level 2 - Randomised trial	(Meinzer et al., 2007)	Intensive language training in the rehabilitation of chronic aphasia: efficient training by laypersons.	CIAT by therapists *vs.* CIAT by laypersons	20	Impairment
Constraint	Level 2 - Randomised trial	(Wilssens et al., 2015)	Constraint-Induced Aphasia Therapy Versus Intensive Semantic Treatment in Fluent Aphasia	CIAT *vs.* Standard Therapy	9 (5 constraint)	Impairment Activity/Participation
Constraint	Level 3 - Non-randomised, controlled cohort/followup study	(L. M. Maher et al., 2006)	A pilot study of use-dependent learning in the context of Constraint Induced Language Therapy.	CIAT *vs.* PACE	9 (4 constraint)	Impairment
Constraint	Level 3 - Non-randomised, controlled cohort/followup study	(Meinzer et al., 2005)	Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy	CIAT *vs.* CIAT Plus	27	Impairment Activity/Participation
Constraint	Level 4 - Case series	(Johnson et al., 2014)	An enhanced protocol for constraint-induced aphasia therapy II: a case series.	CIAT II	4	Impairment Activity/Participation
Constraint	Level 4 - Case series	(Faroqi-Shah & Virion, 2009)	Constraint-induced language therapy for agrammatism: Role of grammaticality constraints	CIAT *vs.* CIAT-G	4	Impairment
Constraint	Level 4 - Case series	(Kempler & Goral, 2011)	A comparison of drill- and communication-based treatment for aphasia	Generative CIAT *vs.* Drill CIAT	2	Impairment
Constraint	Level 4 - Case series	(Goral & Kempler, 2009)	Training verb production in communicative context: evidence from a person with chronic non-fluent aphasia	Modified CIAT (reduced intensity)	1	Impairment

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
Constraint	Level 4 - Case series	(Maul et al., 2014)	Using informative verbal exchanges to promote verb retrieval in nonfluent aphasia	Modified CIAT (reduced intensity, modelling of target sentences)	4	Impairment
Constraint	Level 4 - Case series	(Breier et al., 2009)	Behavioral and neurophysiologic response to therapy for chronic aphasia	CIAT	23	Impairment
Constraint	Level 4 - Case series	(Breier, Juranek, & Papanicolaou, 2011)	Changes in maps of language function and the integrity of the arcuate fasciculus after therapy for chronic aphasia	CIAT	1	Impairment
Constraint	Level 4 - Case series	(Breier et al., 2006)	Functional imaging before and after constraint-induced language therapy for aphasia using magnetoencephalography.	CIAT	6	Impairment
Constraint	Level 4 - Case series	(Breier, Maher, Schmadeke, Hasan, & Papanicolaou, 2007)	Changes in language-specific brain activation after therapy for aphasia using magnetoencephalography: A case study	CIAT	1	Impairment
Constraint	Level 4 - Case series	(MacGregor et al., 2015)	Ultra-rapid access to words in chronic aphasia: The effects of intensive language action therapy (ILAT)	CIAT	12	Impairment
Constraint	Level 4 - Case series	(Mohr et al., 2014)	Changes of right-hemispheric activation after constraint-induced, intensive language action therapy in chronic aphasia: fMRI evidence from auditory semantic processing.	CIAT	12	Impairment
Constraint	Level 4 - Case series	(Pulvermüller, Hauk, Zohsel,	Therapy-related reorganization of language in both hemispheres of patients with chronic aphasia	CIAT	10	Impairment

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
		Neininger, & Mohr, 2005)				
Constraint	Level 4 - Case series	(Richter, Miltner, & Straube, 2008)	Association between therapy outcome and right-hemispheric activation in chronic aphasia	CIAT	24 (16 constraint)	Impairment
Constraint	Level 4 - Case series	(Szaflarski et al., 2008)	Constraint-induced aphasia therapy stimulates language recovery in patients with chronic aphasia after ischemic stroke	CIAT (individualised goals and stimuli)	3	Impairment Activity/Participation
Constraint	Level 4 - Case series	(Kurland, Baldwin, & Tauer, 2010)	Treatment-induced neuroplasticity following intensive naming therapy in a case of chronic wernicke's aphasia	CIAT *vs.* PACE	1	Impairment
Constraint	Level 4 - Single case experimental design	(Kavian, Khatoonabadi, Ansari, Saadati, & Shaygannejad, 2014)	A Single-subject Study to Examine the Effects of Constrained-induced Aphasia Therapy on Naming Deficit.	CIAT	2	Impairment
Constraint	Level 4 - Single case experimental design	(Kurland et al., 2012)	Constrained Versus Unconstrained Intensive Language Therapy in Two Individuals With Chronic, Moderate-to-Severe Aphasia and Apraxia of Speech: Behavioral and fMRI Outcomes.	CIAT *vs.* PACE	2	Impairment
Combined Multimodal & Constraint	Level 4 - Single case experimental design	(Attard et al., 2013)	The comparative effects of Multi-Modality Aphasia Therapy and Constraint-Induced Aphasia Therapy-Plus for severe chronic Broca's aphasia: An in-depth pilot study.	CIAT Plus *vs.* MMAT	2	Impairment, Activity/Participation Quality of Life

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
Combined Multimodal & Constraint	Level 4 - Single case experimental design	(Rose et al., 2013)	Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: A phase 1 study	CIAT Plus *vs.* MMAT	11	Impairment, Activity/Participation Quality of Life
Gesture	Level 4 - Case series	(Carragher, Sage, & Conroy, 2013)	The effects of verb retrieval therapy for people with non-fluent aphasia: Evidence from assessment tasks and conversation	Semantic Feature Analysis + Gesture + phonemic cueing	9	Impairment
Gesture	Level 4 - Case series	(Marangolo et al., 2010)	Improving language without words: first evidence from aphasia	Action observation *vs.* Action observation and execution *vs.* Action observation and meaningless movement	6 (5 with stroke)	Impairment
Gesture	Level 4 - Single case experimental design	(Boo & Rose, 2011)	The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in Broca's aphasia	Repetition *vs.* semantic *vs.* gesture *vs.* semantic + gesture	2	Impairment Activity/Participation
Gesture	Level 4 - Single case experimental design	(Ferguson et al., 2012)	A comparison of intention and pantomime gesture treatment for noun retrieval in people with aphasia	Intention gesture *vs.* pantomime gesture	4	Impairment
Gesture	Level 4 - Single case experimental design	(Raymer et al., 2012)	Contrasting effects of errorless naming treatment and gestural facilitation for word retrieval in aphasia	Errorless naming *vs.* gesture	8	Impairment Activity/Participation
Gesture	Level 4 - Single case experimental design	(Raymer et al., 2006)	Effects of gesture+verbal treatment for noun and verb retrieval in aphasia	Gesture + verbal	9	Impairment

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
Gesture	Level 4 - Single case experimental design	(Rodriguez et al., 2006)	Effects of gesture+verbal and semantic-phonologic treatments for verb retrieval in aphasia	Gesture + verbal *vs.* Semantic + phonologic	4	Impairment
Gesture	Level 4 - Single case experimental design	(Rose & Douglas, 2008)	Treating a semantic word production deficit in aphasia with verbal and gesture methods	Verbal *vs.* Gesture *vs.* Verbal + gesture	1	Impairment
Gesture	Level 4 - Single case experimental design	(Rose, Douglas, & Matyas, 2002)	The comparative effectiveness of gesture and verbal treatments for a specific phonologic naming impairment	Verbal *vs.* Gesture *vs.* Verbal + gesture	1	Impairment
Gesture	Level 4 - Single case experimental design	(Rose & Sussmilch, 2008)	The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia	Semantic *vs.* Gesture *vs.* Semantic + Gesture *vs.* Repetition	3	Impairment Activity/Participation
Music	Level 3 - Non-randomised, controlled cohort/followup study	(Wan, Zheng, Marchina, Norton, & Schlaug, 2014)	Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia	MIT *vs.* no therapy	20 (11 MIT)	Impairment
Music	Level 3 - Non-randomised, controlled cohort/followup study	(Lim et al., 2013)	The therapeutic effect of neurologic music therapy and speech language therapy in post-stroke aphasic patients	MIT *vs.* Standard Therapy	21 (11 chronic, 6 chronic and music)	Impairment
Music	Level 4 - Case series	(Bonakdarpour, Eftekharzadeh, & Ashayeri, 2003)	Melodic intonation therapy in Persian aphasic patients	MIT (adapted to Farsi)	7	Impairment

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
Music	Level 4 - Case series	(Breier, Randle, Maher, & Papanicolaou, 2010)	Changes in maps of language activity activation following melodic intonation therapy using magnetoencephalography: two case studies.	MIT	2	Impairment
Music	Level 4 - Case series	(Goldfarb & Bader, 1979)	Espousing melodic intonation therapy in aphasia rehabilitation: a case study	MIT	1	Impairment
Music	Level 4 - Case series	(Morrow-Odom & Swann, 2013)	Effectiveness of melodic intonation therapy in a case of aphasia following right hemisphere stroke	MIT	1	Impairment Activity/Participation Quality of Life
Music	Level 4 - Case series	(Schlaug, Marchina, & Norton, 2008)	From singing to speaking: Why singing may lead to recovery of expressive language function in patients with Broca's aphasia	MIT *vs.* Speech Repetition Therapy	2	Impairment
Music	Level 4 - Case series	(Sparks et al., 1974)	Aphasia rehabilitation resulting from melodic intonation therapy	MIT	9	Impairment
Music	Level 4 - Case series	(van der Meulen, van de Sandt-Koenderman, & Ribbers, 2012)	Melodic Intonation Therapy: Present Controversies and Future Opportunities	MIT	2	Impairment
Music	Level 4 - Case series	(Wilson, Parsons, & Reutens, 2006)	Preserved Singing in Aphasia: A Case Study of the Efficacy of Melodic Intonation Therapy	MIT *vs.* Rhythmic therapy	1	Impairment
Music	Level 4 - Single case experimental design	(Hough, 2010)	Melodic intonation therapy and aphasia: Another variation on a theme	Modified MIT (no tapping)	1	Impairment Activity/Participation

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
						Quality of Life
Music	Level 3 - Non-randomised, controlled cohort/followup study	(Stahl et al., 2013)	How to engage the right brain hemisphere in aphasics without even singing: Evidence for two paths of speech recovery	Therapy *vs.* Repetition Therapy *vs.* Rhythmic Therapy	15 (5 music, 5 rhythm)	Impairment Activity/Participation
Music	Level 4 - Case series	(Zumbansen, Peretz, & Hébert, 2014b)	The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca's aphasia	Melodic therapy *vs.* Rhythmic therapy *vs.* Standard Therapy	3	Impairment Quality of Life
Pharma/stimulation	Level 2 - Randomised trial & Level 4 - Case series	(Barbancho et al., 2015)	Bilateral brain reorganization with memantine and constraint-induced aphasia therapy in chronic post-stroke aphasia: An ERP study.	CIAT + Memantine	28	Impairment
Pharma/stimulation	Level 2 - Randomised trial & Level 4 - Case series	(Berthier et al., 2009)	Memantine and Constraint-Induced Aphasia Therapy in Chronic Poststroke Aphasia	CIAT + Memantine	27	Impairment Activity/Participation
Pharma/stimulation	Level 4 - Case series	(Abo et al., 2012)	Effectiveness of Low-Frequency rTMS and Intensive Speech Therapy in Poststroke Patients with Aphasia: A Pilot Study Based on Evaluation by fMRI in Relation to Type of Aphasia	Constraint therapy + rTMS	24	Impairment
Pharma/stimulation	Level 4 - Case series	(Martin et al., 2014)	Language improvements after TMS plus modified CILT: Pilot, open-protocol study with two, chronic nonfluent aphasia cases.	Modified CILT + TMS	2	Impairment
Pharma/stimulation	Level 4 - Case series	(Vines et al., 2011)	Non-invasive brain stimulation enhances the effects of melodic intonation therapy	MIT + tDCS	6	Impairment

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
Pharma/stimulation	Level 4 - Single case experimental design	(Al-Janabi et al., 2014)	Augmenting melodic intonation therapy with non-invasive brain stimulation to treat impaired left-hemisphere function: two case studies	MIT + rTMS	2	Impairment
Writing	Level 4 - Case series	(Sugishita, Seki, Kabe, & Yunoki, 1993)	A material-control single-case study of the efficacy of treatment for written and oral naming difficulties	Naming with written cueing hierarchy	22 (3 chronic)	Impairment
Writing	Level 4 - Case series	(Weill-Chounlamountry, Capelle, Tessier, & Pradat-Diehl, 2013)	Multimodal therapy of word retrieval disorder due to phonological encoding dysfunction	Therapy software "Au fil des mots" (anagrams, copying, writing)	1	Impairment Activity/Participation
Writing	Level 4 - Single case experimental design	(Ball, de Riesthal, Breeding, & Mendoza, 2011)	Modified ACT and CART in severe aphasia	ACT + CART	3	Impairment
Writing	Level 4 - Single case experimental design	(Beeson & Egnor, 2006)	Combining treatment for written and spoken naming	CART + repetition	2	Impairment
Writing	Level 4 - Single case experimental design	(DeDe et al., 2003)	Teaching self-cues: A treatment approach for verbal naming	Written naming + tactile cueing + verbal naming	1	Impairment
Writing	Level 4 - Single case experimental design	(Hillis, 1989)	Efficacy and generalization of treatment for aphasic naming errors	Written naming + verbal naming	2 (1 within chronic criteria)	Impairment

Treatment	Oxford level	Publication	Title	Treatment(s)	n	Outcome type
Writing	Level 4 - Single case experimental design	(Wright et al., 2008)	Using a written cueing hierarchy to improve verbal naming in aphasia	Written cueing hierarchy based on CART	2	Impairment

Appendix 2A – Secondary sources for Chapter 3

Paper	Terms used
ASHA. (2016). Summary of the Clinical Practice Guideline - Australian Aphasia Rehabilitation Pathway. Retrieved October 28, 2017, from https://www.asha.org/articlesummary.aspx?id=8589971309	Multimodal treatment, used as a keyword for this CPG summary
ASHA. (n.d.). Aphasia - ASHA practice portal. Retrieved October 28, 2017, from https://www.asha.org/PRPSpecificTopic.aspx?folderid=8589934663&section=Treatment	Multimodal treatment, used as a keyword for this CPG summary
ASHA. (n.d.). Summary of the Clinical Practice Guideline - National Stroke Foundation Clinical Guidelines for Stroke Management. Retrieved October 29, 2017, from https://www.asha.org/articlesummary.aspx?id=8589960979	Multimodal treatment, used as a keyword for this CPG summary
ASHA. (n.d.). Summary of the Clinical Practice Guideline - New Zealand Clinical Guidelines for Stroke Management. Retrieved October 29, 2017, from https://www.asha.org/articlesummary.aspx?id=8589960475	Multimodal treatment, used as a keyword for this CPG summary
ASHA. (n.d.). Summary of the Clinical Practice Guideline - RCSLT Clinical Guidelines. Retrieved October 29, 2017, from https://www.asha.org/articlesummary.aspx?id=8589960345	Multimodal treatment, used as a keyword for this CPG summary
Cahana-Amitay, D., & Albert, M. L. (2015a). Neuroscience of aphasia recovery: the concept of neural multifunctionality. *Current Neurology and Neuroscience Reports*, *15*, 41. https://doi.org/10.1007/s11910-015-0568-7	multimodal cueing
Cahana-Amitay, D., & Albert, M. L. (2015b). Redefining Recovery from Aphasia. Oxford University Press. https://doi.org/10.1093/med/9780199811939.003.0009	multimodality aphasia treatment, multimodal cueing
Centeno, J., & Ansaldo, A. (2013). Aphasia in Multilingual Populations. In I. Papathanasiou, P. Coppens, & C. Potagas (Eds.), *Aphasia and Related Neurogenic Communication Disorders* (pp. 275–293). Burlington, MA: Jones & Bartlett Publishers.	multimodality stimulation

Cherney, L. R., & Robey, R. R. (2001). Aphasia Treatment: Recovery, Prognosis, and Clinical Effectiveness. In R. Chapey (Ed.), *Language interventions strategies in aphasia and related neurogenic communication disorders* (4 ed., pp. 186–202). Baltimore, MD: Lippincott Williams & Wilkins.	Schuell-Wepman-Darley Multimodal-stimulation (SWDM)
Duffy, J. R., & Coelho, C. A. (2001). Schuell's stimulation approach to rehabilitation. In R. Chapey (Ed.), *Language interventions strategies in aphasia and related neurogenic communication disorders* (4 ed., pp. 341–379). Baltimore: Lipincott Williams & Wilkins.	Multimodality stimulation
Fucetola, R., Holloran, S. M., Pratzel, J., Connor, L. T., & Tucker, F. (2011). Innovation in Evidence-Based Language Treatment: 10-Year Report [presentation slides]. *American Speech and Hearing Association Conference*. San Diego, CA. Retrieved from https://www.asha.org/Events/convention/handouts/2011/Fucetola-Holloran-Pratzel-Connor-Tucker/	Multimodal treatment
Hartley, M. L., Turry, A., & Raghavan, P. (2010). The role of music and music therapy in aphasia rehabilitation. *Music and Medicine*, *2*(4), 235–242.	Multimodal approach
Howard, D., Patterson, K., Franklin, S., Orchard-lisle, V., & Morton, J. (1985). The facilitation of picture naming in aphasia. *Cognitive Neuropsychology*, *2*(1), 49–80. https://doi.org/10.1080/02643298508252861	multimodal approach multimodal therapy
Huber, W., Springer, L., & Willmes, K. (1993). Approaches to Aphasia Therapy in Aachen. In A. Holland & M. Forbes (Eds.) (pp. 55–86). San Diego, CA: Singular.	Multimodal stimulation
Jakab, I. (1985). Nonverbal Expression and Congenital Aphasia. In P. Pichot, P. Berner, R. Wolf, & K. Thau (Eds.), *Clinical Psychopathology Nomenclature and Classification* (pp. 1029–1037). New York.	Multimodality treatment
Katz, R. C. (2001). Computer Applications in Aphasia Treatment. In R. Chapey (Ed.), *Language interventions strategies in aphasia and related neurogenic communication disorders* (4 ed., pp. 852–877). Baltimore, MD: Lippincott Williams & Wilkins.	Multi-modal treatment
Kearns, K. P., & Simmons, N. (1985). Group Therapy for Aphasia: A Survey of VA medical centres (pp. 176–183). Presented at the Clinical Aphasiology Conference,	Multimodality stimulation Multimodal stimulation

Ashland, OR. Retrieved from http://eprints-prod-05.library.pitt.edu/851/1/15-22.pdf	
Koul, R., & Corwin, M. (2011). The Process of Evidence-Based Practice: Informing AAC Clinical Decisions for Persons with Aphasia. In R. Koul (Ed.), *Augmentative and Alternative Communication for Adults with Aphasia: Science and Clinical Practice* (pp. 155–164). Bingley, UK: Emerald.	multimodal treatment package
Madden, E. B., Robinson, R. M., & Kendall, D. L. (2017). Phonological Treatment Approaches for Spoken Word Production in Aphasia. *Seminars in Speech and Language*, *38*(1), 62–74. https://doi.org/10.1055/s-0036-1597258	multimodal training
Morganstein, S., & Certner-Smith, M. (2001). Thematic Language-Stimulation Therapy. In R. Chapey (Ed.), *Language interventions strategies in aphasia and related neurogenic communication disorders* (4 ed., pp. 450–468). Baltimore, MD: Lipincott Williams & Wilkins.	Multimodality stimulation
Murray, L., & Coppens, P. (2013). Formal and Informal Assessment of Aphasia. In I. Papathanasiou, P. Coppens, & C. Potagas (Eds.), *Aphasia and Related Neurogenic Communication Disorders* (pp. 67–91). Burlington, MA: Jones & Bartlett Publishers.	Multimodal cue
Pierce, J. E., Menahemi-Falkov, M., O'Halloran, R., Togher, L., & Rose, M. L. (2017). Constraint and multimodal approaches to therapy for chronic aphasia: A systematic review and meta-analysis. *Neuropsychological Rehabilitation*, 1005–1041. https://doi.org/10.1080/09602011.2017.1365730	Multimodal approaches Multimodal therapy Multimodal treatment Multimodal cueing Multimodal training
Purdy, M., & Dietz, A. (2010). Acquired Communication Disorders and Cognitive Deficits: AAC Intervention Challenges. *Perspectives on Augmentative and Alternative Communication*, *19*(3), 62–10. https://doi.org/10.1044/aac19.3.62	Multimodality Communication Training Program
Robey, R. R. (1994). The efficacy of treatment for aphasic persons: a meta-analysis. *Brain and Language*, *47*(4), 582–608. https://doi.org/10.1006/brln.1994.1060	Multimodal stimulus/response

Robey, R. R. (1998). A Meta-Analysis of Clinical Outcomes in the Treatment of Aphasia. *Journal of Speech, Language, and Hearing Research*, *41*(1), 172–16. https://doi.org/10.1044/jslhr.4101.172	Multimodality stimulation hierarchies Schuell-Wepman-Darley Multimodal-stimulation (SWDM)
Rose, M. L. (2013). Releasing the constraints on aphasia therapy: the positive impact of gesture and multimodality treatments. *American Journal of Speech-Language Pathology*, *22*(2), S227–39. https://doi.org/10.1044/1058-0360(2012/12-0091)	Multimodality treatments
Rose, M. L., Raymer, A. M., Lanyon, L. E., & Attard, M. C. (2013b). A systematic review of gesture treatments for post-stroke aphasia. *Aphasiology*, *27*(9), 1090–1127. https://doi.org/10.1080/02687038.2013.805726	Multimodal treatments Multimodality treatment approaches Multimodality Aphasia Therapy Multimodality aphasia training Multimodality training
Stark, J. A. (2010, November 14). Reinventing the wheel? On the history of aphasia therapy [Poster]. *American Speech and Hearing Association Conference*.	Multimodality approach Multimodal approach
Wallace, S. (2013, September 1). More Than Words. *The ASHA Leader*, pp. 40–45. https://doi.org/10.1044/leader.FTR1.18092013.40	Multimodal strategies Multimodal Communication Treatment Multimodal treatment
Wertz, R. T. (2003). Efficacy of Aphasia Therapy, Escher, and Sisyphus. In I. Papathanasiou & R. de Bleser (Eds.), *The Sciences of Aphasia: From therapy to theory* (pp. 259–272). Oxford, United Kingdom: Pergamon.	Multimodality stimulation
Williamson, D. S., Richman, M., & Redmond, S. C. (2010). Group Treatment for Aphasia Based on a Hierarchical Framework. Presented at the American Speech and Hearing Association Conference. Retrieved from www.asha.org/Events/convention/handouts/2010/1685-Williamson-Darlene/	Multi-modal approach

Appendix 2B – Original research data extraction table for Chapter 3

Paper	Main terms used	Target of intervention	Input (stimulus)	Clinician input (cueing/prompting/modelling)	Participant output (response req'd)	On verbal errors or part of training protocol?	Timing of modalities[4]
Churney, K. (2014). *Drawing and multimodality communication training as an effective treatment option for individuals with nonfluent aphasia*. California State University, Long Beach.	Multimodality communication training	Total Communication (gesture, drawing, spoken expression, written)	Pictures	Not stated	Verbal, gesture, writing, drawing	*n/a*	Consecutive
Crossley, A. (2007). *Effects of multi-modality communication for people with aphasia (PWAs) and their communication partners (CPs)*. Dalhousie University, Canada.	multi-modality communication treatment	Total Communication (gesture, drawing, written, visual - symbols)	Pictures	Modelling gesture, drawing, writing, picture board (pointing)	Verbal (though not explicitly stated), Gesture, drawing, writing, picture board (pointing)	*n/a*	Participant choice
Carlomagno, S., Zulian, N., Razzano, C., De Mercurio, I., & Marini, A. (2013). Coverbal gestures in the recovery from severe fluent aphasia: A pilot study. *Journal of Communication Disorders*, *46*(1), 84–99. https://doi.org/10.1016/j.jcomdis.2012.08.007	Multimodal communication therapy	Total Communication (gesture, spoken expression)	Pictures	Modelling and feedback of gesture	Gesture + verbal	*n/a*	Simultaneous
Carr, S. A. (2013). *Effects of semantic + multimodal communication program for switching behavior in severe aphasia (Doctoral dissertation)*. Duquesne University, Ann Arbor.	Multimodal Communication Program	Total Communication (gesture, visual - symbols, visual - drawing, spoken expression)	Pictures	Modelling and feedback of verbal, communication board, gesture, drawing	Verbal, communication board (pointing), gesture, drawing	*n/a*	Consecutive (As per Thiel et al. 2015)

[4] Within total communication papers, modalities were trained simultaneously (e.g., producing gesture and speech within a sentence), separately (e.g., treating drawing in one session and writing in another). consecutively (e.g., spoken naming, then written naming, repetition and symbol pointing for the same target word in one session). Alternatively, participants were given free choice as to modalities and timing.

Reference	Program Name	Modalities	Stimuli	Treatment	Response Modes		Order
Carr, S. A., & Wallace, S. E. (2013). Effects of Semantic + Multimodal Communication Program for Switching Behavior in Moderate-Severe Aphasia. Presented at the 43rd Clinical Aphasiology Conference, Tucson, AZ. Retrieved from http://aphasiology.pitt.edu/2465/	Multimodal Communication Program	Total Communication (gesture, visual - symbols, visual - drawing, spoken expression)	Pictures	Modelling and feedback of verbal, communication board, gesture, drawing	Verbal, communication board (pointing), gesture, drawing	n/a	Consecutive
Schwartz, L., Nemeroff, S., & Reiss, M. (1974). An Investigation of Writing Therapy for the Adult Aphasic: The World Level. *Cortex*, *10*(3), 278–283. https://doi.org/10.1016/S0010-9452(74)80020-1	Multi-modality language therapy	Total Communication (spoken expression, gesture, reading comprehension, auditory comprehension, writing, drawing)	Pictures, written words, verbal	Not stated	Verbal, pictures (pointing), reading aloud, writing	n/a	Consecutive (As per Thiel et al. 2015)
Wallace, S. E., & Kayode, S. (2017). Effects of a semantic plus multimodal communication treatment for modality switching in severe aphasia. *Aphasiology*, *31*(10), 1127–1142. https://doi.org/10.1080/02687038.2016.1245403	Multimodal Communication Treatment	Total Communication (spoken expression, gesture, visual - picture pointing, drawing)	Pictures	Modelling and feedback of verbal, picture board (pointing), gesture, drawing	Verbal, picture board (pointing), gesture, drawing	n/a	Consecutive
Wallace, S. E., Purdy, M., & Skidmore, E. (2014). A multimodal communication program for aphasia during inpatient rehabilitation: A case study. *NeuroRehabilitation*, *35*(3), 615–625. https://doi.org/10.3233/NRE-141136	multimodal communication program	Total Communication (spoken expression, gesture, visual - picture pointing, drawing)	Pictures	Modelling and feedback of verbal, picture board (pointing), gesture, drawing	Verbal, picture board (pointing), gesture, drawing	n/a	Consecutive
Purdy, M., Duffy, R., & Coelho, C. A. (1994). An investigation of the communicative use of trained symbols following multimodality training, *22*, 345–256.	Multimodality training	Total Communication (spoken expression, gesture, visual - picture pointing)	Picture (for gesture and verbal response) Verbal (for picture response)	Modeling picture pointing, gesture and verbal, providing these on errors. Shaping of gestures. Phonemic, semantic and motor cues for verbal.	verbal gesture picture board (pointing)	n/a	Separate

Reference	Approach name						
Lasker, J., LaPointe, L., & Kodras, J. (2005). Helping a professor with aphasia resume teaching through multimodal approaches. *Aphasiology*, *19*(3-5), 399–410.	Multimodal approaches	Total Communication (spoken expression, visual - pictures, writing, text-to-speech)	n/a	n/a	Practised verbal output, written slides with pictures, developed text-to-speech utterances	*n/a*	Separate then simultaneous
Purdy, M., & Van Dyke, J. A. (2011). Multimodal Communication Training in Aphasia: A Pilot Study. *Journal of Medical Speech-Language Pathology*, *19*(3), 45–53.	Multimodal Communication Training	Total Communication (spoken expression, writing, gesture, drawing, visual - picture pointing)	Pictures	Modeling and shaping verbal, gesture, writing, picture board (pointing)	Verbal, gesture, writing, picture board (pointing)	*n/a*	Consecutive
Purdy, M., & Wallace, S. (2013). The Feasibility of a Multimodal Communication Treatment for Aphasia during Inpatient Rehabilitation. Presented at the Clinical Aphasiology Conference, Tucson, AZ. Retrieved from http://aphasiology.pitt.edu/2505/	Multimodal Communication Training	Total Communication (spoken expression, writing, gesture, drawing, visual - picture pointing)	Pictures	Modeling verbal, gesture, drawing writing, picture board (pointing)	Verbal, gesture, drawing writing, picture board (pointing)	*n/a*	Consecutive
Purdy, M., & Wallace, S. E. (2015). Intensive multimodal communication treatment for people with chronic aphasia. *Aphasiology*, *30*(10), 1071–1093. https://doi.org/10.1080/02687038.2015.1102855	Multimodal Communication Treatment	Total Communication (spoken expression, writing, gesture, drawing, visual - picture pointing)	Pictures	Part 1. Modeling verbal, gesture, drawing, writing, picture board (pointing) Part 2. Prompting for each modality without model unless needed	Verbal, gesture, drawing writing, picture board (pointing)	*n/a*	Consecutive
Macoir, J., Sauvageau, V. M., Boissy, P., Tousignant, M., & Tousignant, M. (2017). In-Home Synchronous Telespeech Therapy to Improve Functional Communication in Chronic Poststroke Aphasia: Results from a Quasi-Experimental Study. *Telemedicine*	multimodal language therapy	Total Communication (spoken expression, writing, gesture, drawing)	Pictures	Not stated	Verbal, gesture, writing (typing), drawing	*n/a*	Participant choice

Journal & E-Health, *23*(8), 630–639. https://doi.org/10.1089/tmj.2016.0235

Reference	Approach	Focus					
Beeson, P. M., & Ramage, A. E. (2000). Drawing from experience: the development of alternative communication strategies. *Topics in Stroke Rehabilitation*, *7*(2), 10–20.	Multimodal approach	Total Communication (visual - symbols, visual - drawing, written)	Unclear	Modeling drawing, encouraging picture board use (software), providing copy and recall and anagram treatments	Drawing, verbal, writing, picture board (software)	*n/a*	Separate
Brookshire, C. E., Conway, T., Pompon, R. H., Oelke, M., & Kendall, D. L. (2014). Effects of intensive phonomotor treatment on reading in eight individuals with aphasia and phonological alexia. *American Journal of Speech-Language Pathology*, *23*(2), S300–11. https://doi.org/10.1044/2014_AJSLP-13-0083	Multimodal treatment	Facilitation: Reading comprehension Phonological processing	Verbal (phonemes and syllables), pictures, writing (letters)	Verbal (motor placement descriptions), verbal (phoneme discrimination), verbal (repetition), verbal (phoneme/letter association)	Verbal (phonemes and syllables), auditory (phoneme discrimination)	Part of training	*n/a*
Rose, M. L., Mok, Z., Carragher, M., Katthagen, S., & Attard, M. C. (2015). Comparing multi-modality and constraint-induced treatment for aphasia: a preliminary investigation of generalisation to discourse. *Aphasiology*, 1–21. https://doi.org/10.1080/02687038.2015.1100706	Multi-Modality Aphasia Therapy	Facilitation: Spoken expression (discourse)	Pictures	Modeling verbal, gesture, drawing, written	Verbal (repetition), verbal (oral reading), drawing, written, gesture	On errors	*n/a*
Thomson, J. (2012). *Assessing the benefits of multimodal rehabilitation therapy for aphasia [Masters Dissertation]* (pp. 1–143). University of Manchester.	multimodal rehabilitation therapy multimodal item focused therapy	Facilitation: Spoken expression (noun retrieval, verb retrieval)	1. Verbal 2. Verbal 3. Verbal (questions), picture, written 4. Picture, written 5. Written	Verbal feedback or modeling	1. Pointing to picture 2. Pointing to correct written word 3. Verbal (yes/no semantic questions) 4. Verbal	Part of training	*n/a*

			(letters) 6. Written 7. Verbal		5. Written (unscramble) 6. Verbal 7. Verbal (repetition)		
Denman, A. (2017, September). Multi-modal errorless learning functional naming therapy - a single case study [poster]. *RCSLT Conference*. Glasgow, Scotland.	Multi-modal errorless learning functional naming therapy	Facilitation: Spoken expression (noun retrieval)	objects, photos, written words and spoken	Verbal	Verbal (repetition)	*n/a – only on input*	*n/a*
Dunn, I. (2010). *The effects of multimodality cueing on lexical retrieval in aphasic speakers (Doctoral Thesis)*. The William Paterson University of New Jersey.	Multimodality cueing	Facilitation: Spoken expression (noun retrieval)	Pictures	Modelling of gesture as needed along with semantic or phonological cues	Verbal (no evidence that subject produced gesture)	On errors	*n/a*
Hoodin, R. B., & Thompson, C. K. (1983). Facilitation of verbal labeling in adult aphasia by gestural, verbal, or verbal plus gestural training (pp. 62–64). Presented at the Clinical Aphasiology Conference, Phoenix, AZ.	multimodality training	Facilitation: Spoken expression (noun retrieval)	Not stated	Not stated	Verbal, gesture	Part of training	*n/a*
Kendall, D. L., Oelke, M., Brookshire, C. E., & Nadeau, S. E. (2015). The Influence of Phonomotor Treatment on Word Retrieval Abilities in 26 Individuals With Chronic Aphasia: An Open Trial. *Journal of Speech, Language, and Hearing Research*, *58*(3), 798–15. https://doi.org/10.1044/2015_JSLHR-L-14-0131	Multimodal therapy	Facilitation: Spoken expression (noun retrieval)	Mouth pictures + verbal (phonemes), mouth pictures + written letters	Provides placement descriptions, discriminations choices, repetition and sound/letter association	Verbal, objects (arranging coloured blocks), pointing (written letters)	Part of training	*n/a*
Rose, M. L., Attard, M. C., Mok, Z., Lanyon, L. E., & Foster, A. M. (2013a). Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: A phase 1 study. *Aphasiology*, *27*(8), 938–971.	Multi-Modality Aphasia Therapy	Facilitation: Spoken expression (noun retrieval)	Pictures	Modeling verbal, gesture, drawing, written	Verbal (repetition), verbal (oral reading), drawing, written, gesture	On errors	*n/a*

https://doi.org/10.1080/02687038.2013.810329

Weill-Chounlamountry, A., Capelle, N., Tessier, C., & Pradat-Diehl, P. (2013). Multimodal therapy of word retrieval disorder due to phonological encoding dysfunction. *Brain Injury*, *27*(5), 620–631. https://doi.org/10.3109/02699052.2013.767936	Multimodal therapy	Facilitation: Spoken expression (noun retrieval)	1. Pictures, written (scrambled letters) 2. Written, verbal (phonemes, syllables, words) 3. Written 4. Picture 5. Picture	Not stated	1. Written (unscramble) 2. Verbal (repetition) 3. Written (copying), verbal 4. Written, verbal 5. Verbal	Part of training	*n/a*
Rebstock, A. M. (2014). *Effects of semantic feature analysis + multimodal communication program for word retrieval and switching behavior in primary progressive aphasia (Doctoral dissertation)*. Duquesne University.	Multimodal Communication Program	Facilitation: Spoken expression (noun retrieval) Total Communication (spoken expression, gesture, drawing)	Pictures	Modeling verbal, gesture and drawing	Verbal, gesture, drawing	Part of training	*n/a*
Attard, M. C., Rose, M. L., & Lanyon, L. E. (2013). The comparative effects of Multi-Modality Aphasia Therapy and Constraint-Induced Aphasia Therapy-Plus for severe chronic Broca's aphasia: An in-depth pilot study. *Aphasiology*, *27*(1), 80–111. https://doi.org/10.1080/02687038.2012.725242	Multi-Modality Aphasia Therapy	Facilitation: Spoken expression (noun retrieval) Total Communication as contingency	Pictures	Modeling verbal, gesture, drawing, written	Verbal (repetition), verbal (oral reading), drawing, written, gesture	On errors	*n/a*
McCarthy, S. E. (2004). *The effects of a multimodality approach on sentence production using response elaboration training with a reading component on aphasic patients (Doctoral dissertation)*. East Tennessee State University.	Multimodality approach Multimodality treatment	Facilitation: Spoken expression (sentences)	Pictures, written	Verbal models	Verbal (sentence description), verbal (oral reading), unscrambling written words	Part of training	*n/a*

Reference	Treatment	Target	Stimuli	Clinician	Patient	Feedback	
Thompson, C. K., & McReynolds, L. V. (1986). Wh interrogative production in agrammatic aphasia: an experimental analysis of auditory-visual stimulation and direct-production treatment. *Journal of Speech and Hearing Research*, *29*(2), 193–206.	multimodal stimulation	Facilitation: Spoken expression (sentences)	Pictures, written words, verbal	Providing verbal stimuli	Verbal (repetition), verbal (Wh-question)	*n/a – only on input*	*n/a*
Henning, D. M. (2016). *Music and multimodal stimulation (M-STIM): A dynamic approach to increasing expressive and receptive language in severe global aphasia (Master's thesis)*. Northern Illinois University.	Music And Multimodal Stimulation (M-STIM)	Facilitation: Spoken expression (sentences) Auditory comprehension	Sensory/tactile, pictures, video, objects,	Modelled melodic target	Melodic verbal output	*n/a – only on input*	*n/a*
Rose, M. L., & Sussmilch, G. (2008). The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology*, *22*(7-8), 691–706. https://doi.org/10.1080/02687030701800800	multi-modal semantic treatment	Facilitation: Spoken expression (verb retrieval in sentences)	Pictures	Modeling verbal and gesture	Verbal + gesture	On errors	*n/a*
Fink, R., Brecher, A., Sobel, P., & Schwartz, M. (2005). Computer-assisted treatment of word retrieval deficits in aphasia. *Aphasiology*, *19*(10-11), 943–954.	multi-modality cueing multi-modality matching	Facilitation: Spoken expression (verb retrieval)	Pictures (cueing)	Software provides verbal or written cueing	Verbal (naming) Select written and/or spoken word (matching)	On errors	*n/a*
Kearns, K. P., Simmons, N., & Sisterhen, C. (1982). Gestural sign (Amer-Ind) as a facilitator of verbalization in patients with aphasia (pp. 183–191). Presented at the Clinical Aphasiology Conference, Oshkosh, WI. Retrieved from http://eprints-prod-05.library.pitt.edu/725/1/12-23.pdf	multimodality training	Facilitation: Spoken expression (verb retrieval)	Pictures	Providing gesture and verbal for imitation	Verbal + gesture	Part of training, then on errors	*n/a*
Thiel, L., Sage, K., & Conroy, P. (2015). Comparing uni-modal and multi-modal therapies for improving writing in acquired dysgraphia after stroke. *Neuropsychological Rehabilitation*, 1–29. https://doi.org/10.1080/09602011.2015.1026357	Multi-modal therapy	Facilitation: Writing (words)	Verbal	Providing verbal stimuli, feedback	1. Pointing to written word, then verbal, writing (copying) 2. Pointing to symbol	Part of training	*n/a*

	representing word, then verbal, writing 3. Pointing to written word, then verbal, writing (copying)

Appendix 3 – Data extraction table for Chapter 4

Study	Design	n (Intensity allocations)	MPO	Aphasia type	Severity	Age	Sex	Treatment	Outcome measures (bold = primary outcome) I	A/P	QoL	Total dose per arm	Schedule - higher	Schedule – lower	High-low ratio‡
Marcotte et al., 2018	Case series	2 (1 higher, 1 lower)	12, 36	2 Broca's	WAB AQ 66 WAB AQ 62	59, 58	1M 1F	Phonological Component Analysis	**BOLD signal changes on fMRI during picture naming** Picture naming – treated items Picture naming – untreated items	-	-	30 hrs	12 hrs/wk x 2.5 wks	3hrs/wk x 10 wks	4:1
Harnish, Neils-Strunjas, Kamy & Eliassen, 2008	Crossover pre-post case series	1 (Higher then lower)	≈96	1 Conduction	WAB AQ 71.6	52	1F	Traditional – naming, picture description, cueing hierarchy, writing to dictation	**BOLD signal changes on fMRI during letter decision task** BNT	-	-	15 hrs	7.5 hrs/wk x 2 wks	2 hrs/wk x 7.5 wks	3.75:1
Sage, Snell & Lambon Ralph, 2011	Crossover design (phase 1 only ≈ RCT)	8 (3 higher, 5 lower)	Range not available. Mean 58.3 (SD 41.1)	5 Fluent 3 Non-fluent	6 Mild 1 Moderate 1 Severe*	Mean 61.2 (SD 8)	6M 2F	Picture naming with cueing hierarchy	**Picture naming – treated items** Untreated items	-	-	10 sessions (?hrs)	5/wk x 2 wks	2/wk x 5 wks	2.5:1

Stahl et al., 2018	RCT	30 (15 higher, 15 lower)	12-243	23 Broca's 5 Global 1 Anomic 1 Wernicke's	AAT mean 50.5 (SD 2.9)	33-84	17M 13F	Constraint Induced Language Therapy (no cueing provided)	**AAT (four subscales)** Action Communication Test	-	-	24 hrs†	12hrs/wk x 2 wks†	6hrs/wk x 4 wks†	2:1
Dignam et al., 2015	NonRCT	34 (16 higher, 18 lower)	4-225	Not reported	CAT Severity scores 51.6 (higher), 52.3 (lower)	35-77	28M 6F	Language Impairment and Functioning Therapy – impairment, computer, functional & group	**BNT**	CETI CCRSA	ALA	48 hrs	16hrs/wk x 3 wks	6hrs/wk x 8 wks	2.67:1
Mozeiko, Coelho & Myers, 2015	SCED – AB	8 (4 higher, 4 lower)	13-134	3 Broca's 1 Global 1 Anomic 1 Conduction 2 Unclassifiable	WAB AQ Mean 48.5, range 24-84	26-77	5M 3F	Constraint Induced Language Therapy	**CIU count** CIUs/word count CIUs/min *WAB AQ (Pre-post)*	*CADL (pre-post)*	-	30 hrs	15hrs/wk x 2 wks	3hrs/wk x 10 wks	5:1
Ramsberger & Marie, 2007	SCED – Multiple baseline across conditions	4 (Higher & lower phases for each participant)	6-72	1 Broca's 1 Wernicke's 1 Conduction 1 Anomic	WAB AQ 69 WAB AQ 53 ADP SS 25th %ile ADP SS 68th %ile	63-74	3F 1M	Cued Naming module of MossTalk Words software – cues customised per participant and across sessions	**Picture naming – treated items**	-	-	**P1/P2:** 15 sessions **P3/P4:** 20 sessions	**P1/P2:** 5/wk x 3 wks **P3/P4:** 5wk x 4 wks	**P1/P2:** 2/wk x 7.5 wks **P3/P4:** 2/wk x 10 wks	2.5:1

| Raymer, Kohen & Saffell, 2006 | SCED – multiple baseline across conditions | 5 *(Higher & lower phases for each participant)* | 4-42 | 2 Broca's 2 Conduction 1 Mixed transcortical | WAB AQ Mean 53.3, range 33-76 | 51-82 | 3F 2M | Three matching modules of MossTalk Words software | **Picture naming – treated items** Semantic word/picture decisions – treated items (C01, C02) *WAB (Pre-post)* *BNT (Pre-post)* | - | - | 12 hrs | 3-4hrs/wk x 3-4wks | 1-2hrs/wk x 6-12wks | 2:1 - 4:1 |

*Our interpretation of reported language scores for multiple assessments

† Reported schedule for our pre-endpoint sub-analysis; complete study investigated more therapy hours

‡ Ratio of weekly treatment in higher intensity to low intensity arms

Note. RCT – Randomised Controlled Trial; SCED – Single Case Experimental Design, MPO – Months Post Onset from stroke; WAB AQ – Western Aphasia Battery Aphasia Quotient; BNT – Boston Naming Test; ADP SS – Aphasia Diagnostic Profiles standard score; I – Impairment level; A/P – Activity/Participation level; QoL – Quality of life level; CCRSA - Communication Confidence Rating Scale for Aphasia ; AAT – Aachen Aphasia Test; ALA – Assessment of Living with Aphasia; CIU – Correct Information Units

Appendix 4 – Supplementary Study Data for Chapter 5

Main trial to Substudy Gap

Histogram of number of weeks between main trial and substudy randomisation

Baseline characteristics

Beeswarm plots of participant characteristics at baseline. One dot = one participant.

Semantic Impairment

Attention

Cognitive Flexibility

Working Memory

Short Term Memory

Raven's Progressive Matrices Score
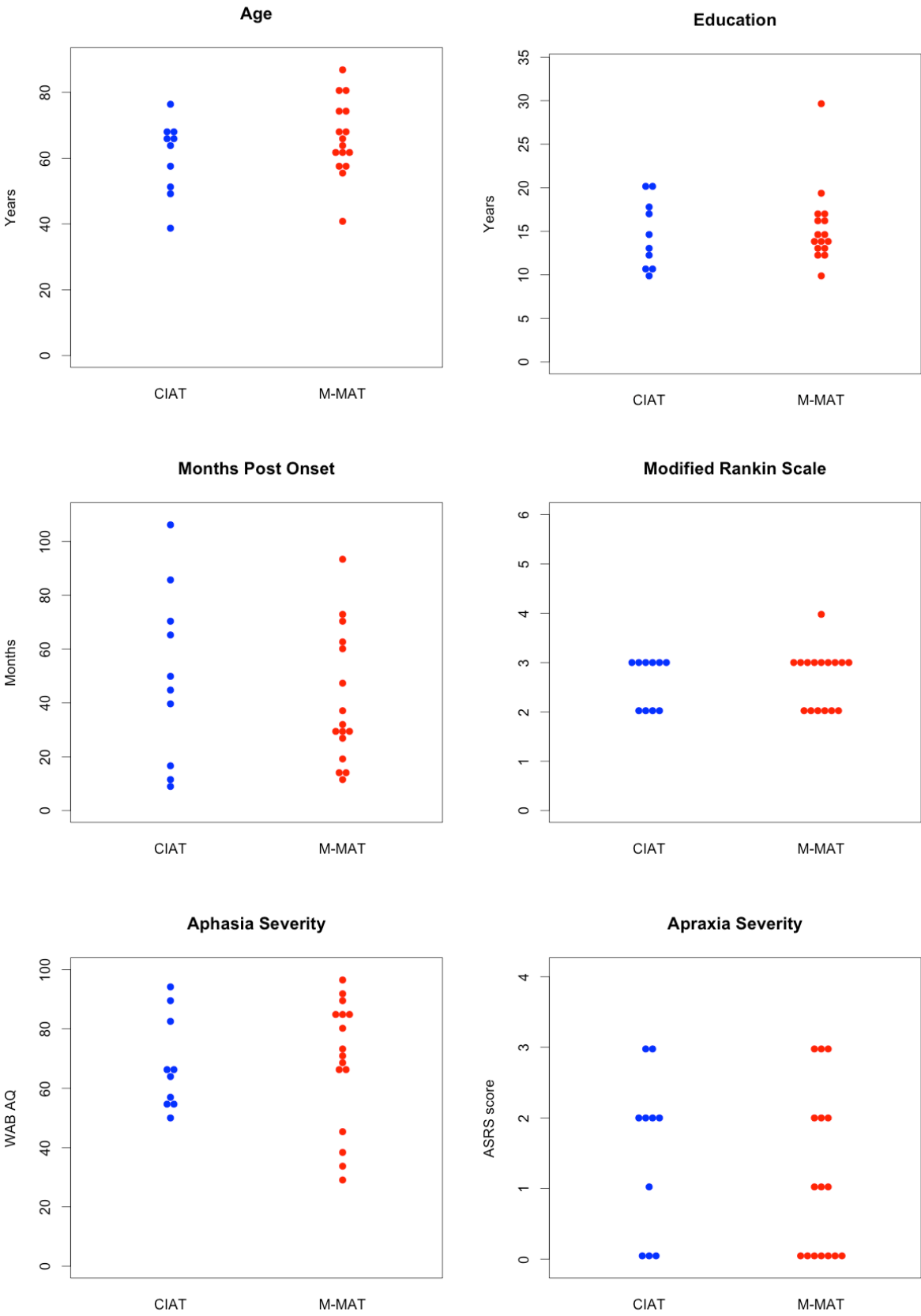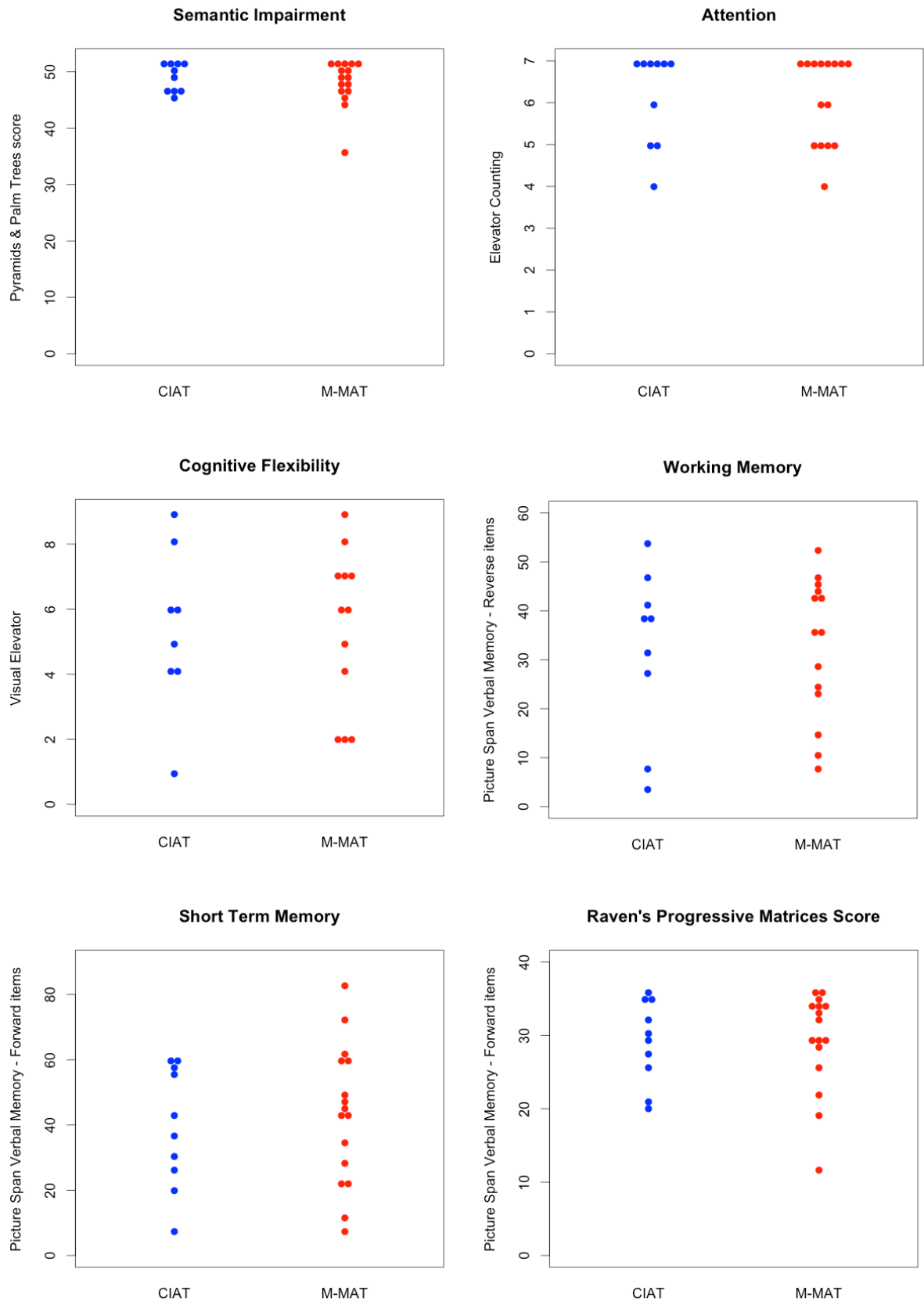
Participants who withdrew

	Completed (n=26)	Withdrew (n=6)
Age	64.0 (11.5)	66.5 (13.4)
Female:Male (% female)	6:20 (23%)	3:3 (50%)
Years of education	15.2 (4.2)	15.5 (5.4)
Months Post Onset Median (IQR)	38.5 (42.5)	39.7 (18.8)
WAB AQ Mean (SD)	68.6 (19.0)	82.5 (6.2)
Mild n (%)	16 (62 %)	6 (100%)
Moderate n (%)	9 (35%)	0
Severe n (%)	1 (3%)	0
Apraxia Severity Rating Scale		
No Apraxia	10 (38%)	3 (50%)
Mild	4 (15%)	1 (17%)
Mild-Moderate	7 (27%)	0
Moderate	5 (19%)	2 (33%)

Mixed Repeated Measure ANOVAs

QQ plots generated in R using the following script:

```
library("ggpubr")

QQ <- function(outcome, b1, b2, b3, b4){

  outcome <- na.omit(outcome)

  outcome <- outcome %>%

    gather(key = "time", value = "score", b1, b2, b3, b4)

  outcome <- outcome %>% convert_as_factor(Group, time)

  ggqqplot(outcome, "score") + facet_grid(Group ~ time)

}

e.g.

WABAQ <- read.csv("C. WAB AQ/WAB AQ.csv")

QQ(WABAQ, "WAB_AQ_B0", "WAB_AQ_B1", "WAB_AQ_Post",

"WAB_AQ_FU")
```
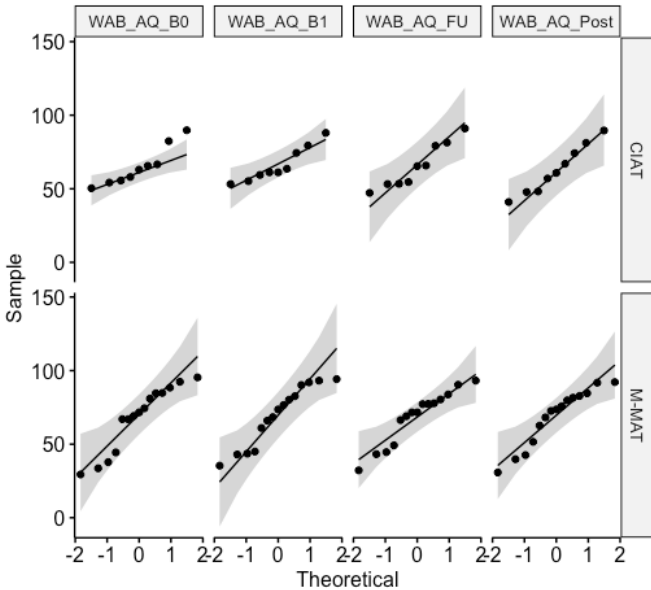
WAB-R AQ

Dropped cases	1 M-MAT and 1 CIAT dropped from analysis due to missing datapoint
Normality and outliers	WAB-R AQ was approximately normally distributed for each group and timepoint as assessed by the Shapiro Wilks test (p>0.05) and examination of QQ plots. There were no outliers, as assessed by examination of studentized residuals for values greater than ±3.
Equality of Variance	WAB-R AQ met the assumption of equality of variances under Levene's Test (p > .05).
Sphericity	Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, $\chi^2(5) = 9.66$, $p = .086$.
Main effect of time	The main effect of time did **not** show a statistically significant difference on WAB-R AQ at the different time points, $F(3, 66) = 0.917$, $p = 0.44$, partial $\eta^2 = .040$
Main effect of group	There was also **no** statistically significant main effect of intervention $F(1, 22) = .255$, $p = .618$, partial $\eta^2 = .011$
Interaction	There was **no** statistically significant interaction between the intervention and time on WAB-R AQ, $F(3, 66) = 0.523$, $p = .67$, partial $\eta^2 = .023$.

Naming battery – treated items

Dropped cases	1 M-MAT and 2 CIAT dropped from analysis due to missing datapoints
Normality and outliers	Normality was not met for CIAT at Baseline 0 according to the Shapiro Wilks test ($p = .014$) and this was demonstrated by one outlier on QQ plots. However, there were no outliers as assessed by
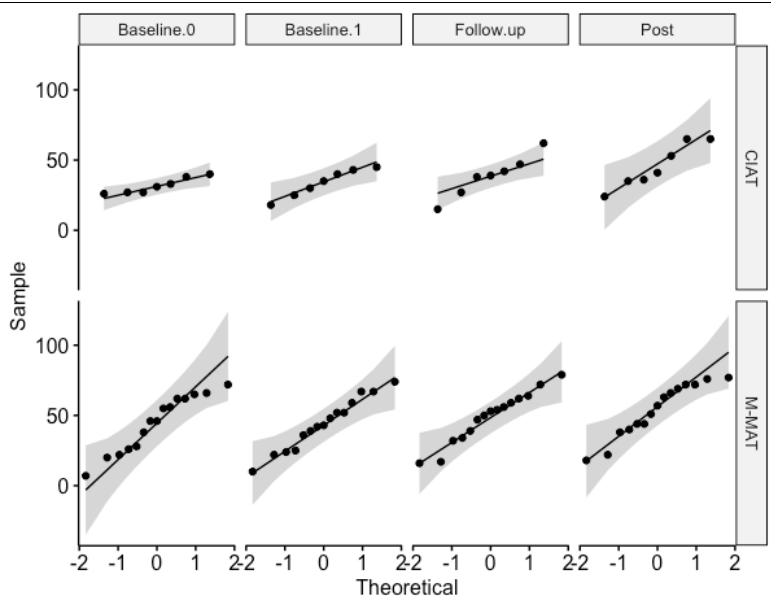
studentized residuals for values greater than ±3. Analysis was completed *with* and *without* the identified outlier.
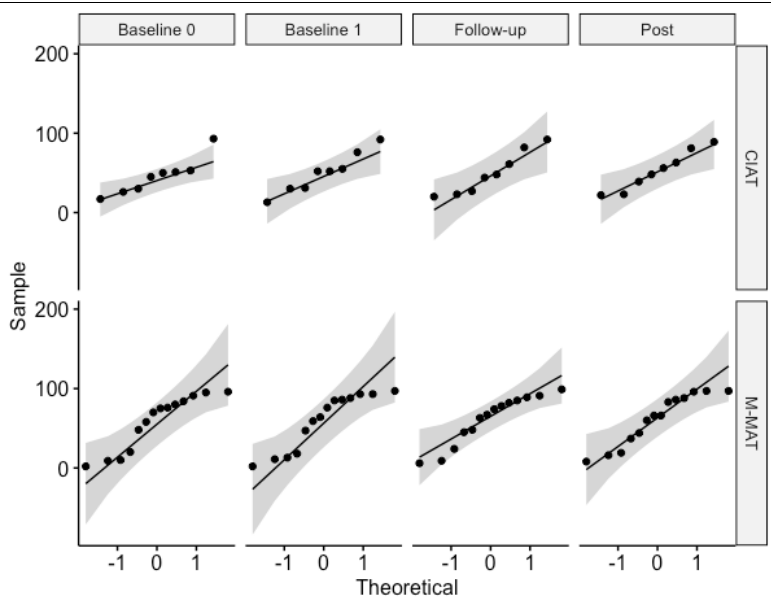
With outlier included	
Equality of Variance	There was equality of variances under Levene's Test (p > .05)
Sphericity	The assumption of sphericity was not met using Mauchly's test of sphericity $\chi^2(5) = 12.639$, $p = .027$. Greenhouse-Geisser correction was therefore applied to interpret the two-way interaction.
Main effect of time	There **was** a significant main effect of time $F(2.3, 48.6) = 14.392$, $p < .001$, partial $\eta2 = .407$. • Post Intervention **was** significantly different to B0 and B1 (p<.001), and Follow-Up ($p = .040$) • Follow-Up was significantly different to B1 ($p = .021$)
Main effect of group	There was **no** main effect of group $F(1, 21) = .732$, $p = .402$, partial $\eta2 = .034$
Interaction	There was **no** significant interaction $F(2.3, 48.6) = .438$, $p = .667$, partial $\eta2 = .020$
Without outlier	
Normality	Normality was met according to Shapiro-Wilks (p > .05) and QQ plots. There were no outliers according to Studentised Residuals.

Equality of Variance	Data did not meet the assumption of equality of variances under Levene's Test at Baseline 0 (p = .004)
Sphericity	The assumption of sphericity was not met using Mauchly's test of sphericity $\chi^2(5)$ = 12.653, p = .027. Greenhouse-Geisser correction was therefore applied to interpret the two-way interaction.
Main effect of time	There **was** a significant main effect of time $F(2.3, 45.9)$ = 12.648, p < .001, partial η2 = .387 • Post intervention **was** significantly different to Baseline 0 (p < .001), Baseline 1 (p < .001) and Follow-Up (p = .040) • Follow-Up was not significantly different to Baselines 0 and 1 (p = .275 and .054 respectively)
Main effect of group	There was **no** main effect of group $F(1, 20)$ = 1.974, p = .175, partial η2 = .090
Interaction	There was **no** significant interaction $F(2.3, 45.9)$ = .405, p = .697, partial η2 = .020
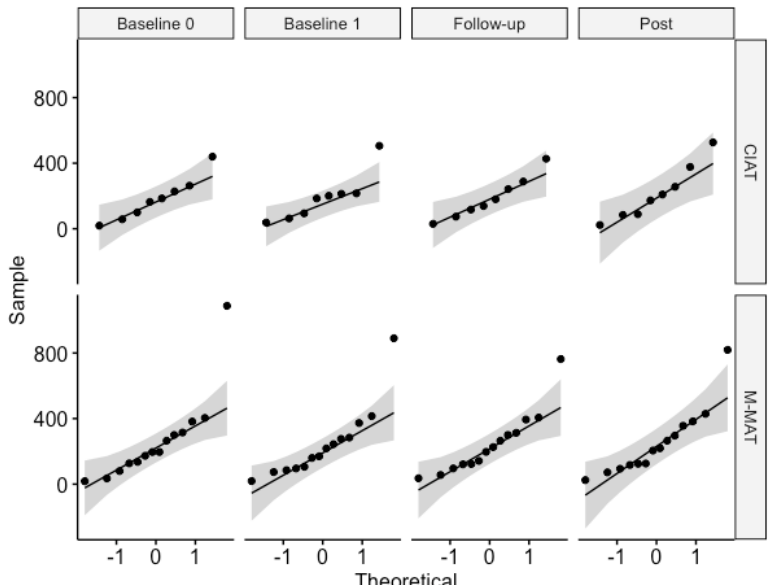
Naming battery – untreated items

Dropped cases	2 CIAT and 2 M-MAT missing due to missing data points
Normality and outliers	Data was **not** normally distributed at MMAT Baseline 0 (p = .037) and B1 (p = .028) as assessed by the Shapiro-Wilks test but was approximately normal according to QQ plots. There were no outliers, as assessed by examination of Studentised residuals for values greater than ± 3.

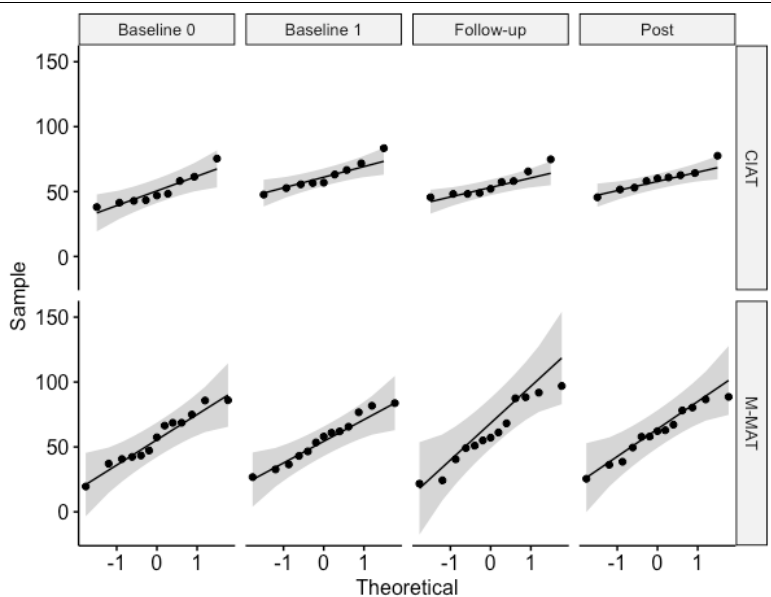| Equality of Variance	Data at each group timepoint met the assumption of equality of variances under Levene's test (p > .05)
Sphericity	Mauchly's test of sphericity indicated that the assumption of sphericity was not met for the two-way interaction, $X^2(5) = 11.849$, $p = .037$. Greenhouse-Geisser correction was therefore applied to interpret the two-way interaction.
Main effect of time	The main effect of time was **not** statistically significant $F(2.086, 41.722) = 2.560$, $p = .063$, partial $\eta^2 = .114$
Main effect of group	There was **no** main effect of group $F(1, 20) = 0.645$, $p = .431$, partial $\eta^2 = .031$
Interaction	There was **no** statistically significant interaction between the intervention and time, $F(2.086, 41.722) = .411$, $p = .674$, partial $\eta^2 = .20$

CIU count
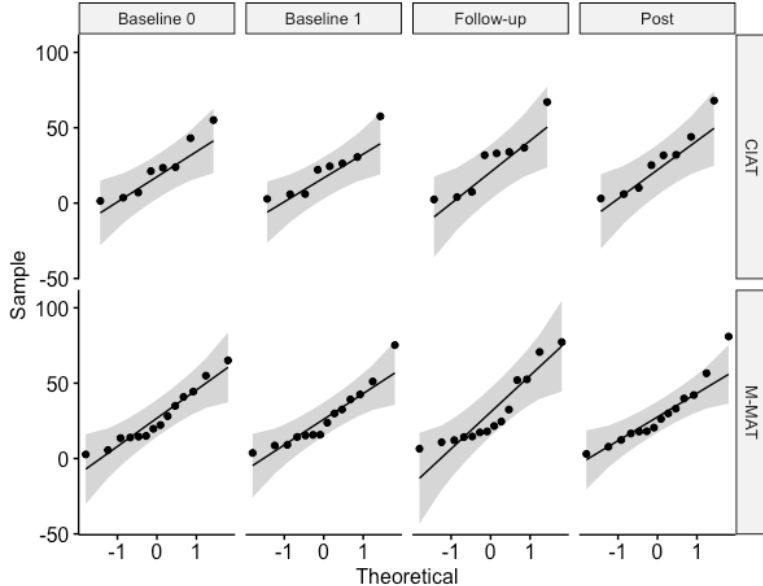
Dropped cases	2 CIAT and 2 M-MAT missing due to missing data points
Normality and outliers	Data was not normally distributed for M-MAT at any timepoint as assessed by the Shapiro-Wilks test ($p < .05$). One outlier was identified which had a studentised residual value >3 for each datapoint. The repeated measures ANOVA was run with and without

	the participant

With outlier included	
Equality of Variance	Data at each group timepoint met the assumption of equality of variances under Levene's test ($p > .05$)
Sphericity	Mauchly's test of sphericity indicated that the assumption of sphericity was not met for the two-way interaction, $X^2(5) = 19.2$, $p = .002$. Correction was therefore applied using Greenhouse-Geisser.
Main effect of time	The main effect of time was **not** statistically significant $F(1.791, 35.814) = 0.728$, $p = .476$, partial $\eta^2 = .035$
Main effect of group	There was **no** main effect of group $F(1, 20) = .449$, $p = .511$, partial $\eta^2 = .022$
Interaction	There was **no** statistically significant interaction between the intervention and time, $F(1.791, 35.814) = 0.967$, $p = .382$, partial $\eta^2 = .046$
Without outlier	
Normality	Data met the Shapiro-Wilks test of normality for each group and timepoint after removal of the above outlier. There were no other outliers, as assessed by examination of Studentised residuals for values greater than ± 3.

Equality of Variance	Data at each group timepoint met the assumption of equality of variances under Levene's test ($p > .05$)
Sphericity	Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, $X^2(5) = 5.482$, $p = .361$
Main effect of time	The main effect of time was **not** statistically significant $F(3, 57) = 1.678$, $p = .182$, partial $\eta^2 = .081$
Main effect of group	There was **no** main effect of group $F(1, 19) = 0.022$, $p = .884$, partial $\eta^2 = .001$
Interaction	There was **no** statistically significant interaction between the intervention and time, $F(3, 57) = 0.769$, $p = .516$, partial $\eta^2 = .039$

CIUs per minute

Dropped cases	2 CIAT and 2 M-MAT missing due to missing data points
Normality and outliers	Data was **not** normally distributed for M-MAT at Follow-Up as assessed by the Shapiro-Wilks test ($p = .015$), however the QQ plot demonstrated approximate normality. Data was not normally distributed for M-MAT at Follow-Up as assessed by the Shapiro-Wilks test ($p = .015$), however the QQ plot demonstrated approximate normality.

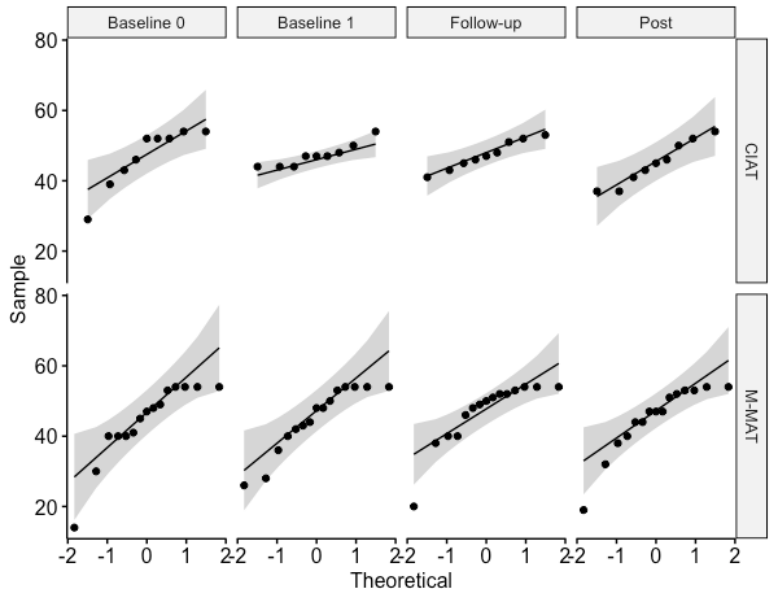| Equality of Variance	Data at each group timepoint met the assumption of equality of variances under Levene's test ($p > .05$)
Sphericity	Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, $X^2(5) = 9.318$, $p = .097$
Main effect of time	The main effect of time **was** statistically significant $F(3, 60) = 4.492$, $p = .007$, partial $\eta^2 = .183$; however there were no statistically significant difference between timepoints on post hoc analysis using one-way ANOVAs.
Main effect of group	There was no main effect of group $F(1, 20) = .156$, $p = .697$, partial $\eta^2 = .008$
Interaction	There was **no** statistically significant interaction between the intervention and time, $F(3, 60) = .529$, $p = .664$, partial $\eta^2 = .026$
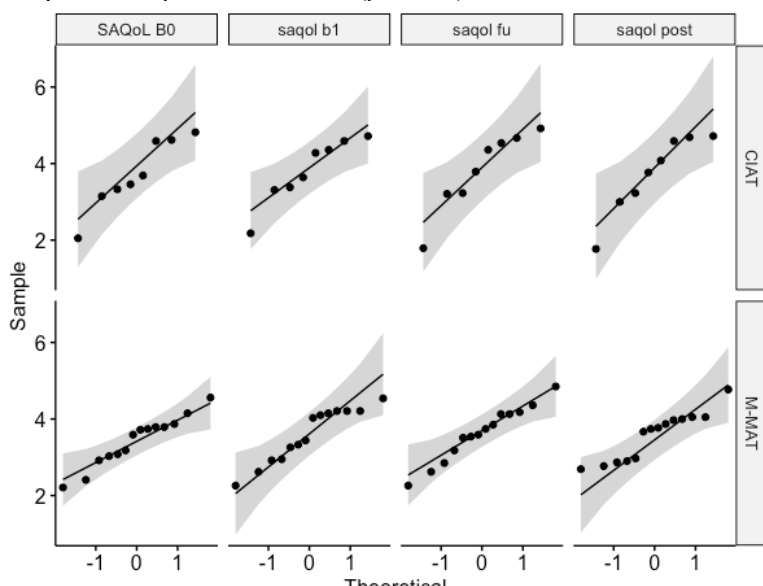
CETI

Dropped cases	1 CIAT and 3 M-MAT missing due to missing data points
Normality and outliers	Data was normally distributed at each group and timepoint as assessed by the Shapiro-Wilks test ($p > .05$) and QQ plots.

	There were no outliers, as assessed by examination of Studentised residuals for values greater than ± 3
Equality of Variance	Data did not meet the assumption of equality of variances under Levene's test at Post (p = .048) or Follow-Up (p = .022). This was likely to not have had an effect on results as there was no main effect of group.
Sphericity	Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, X^2(5) = 9.126, p = .105
Main effect of time	The main effect of time **was** statistically significant F(3, 60) = 3.173, p = .031, partial η² = .137 CETI scores were **significantly different** between Baseline 0 and Post intervention (p = .036)
Main effect of group	There was no main effect of group F(1, 20) = .074, p = .788, partial η² = .004
Interaction	There **was** a statistically significant interaction between the intervention and time, F(3, 60) = 3.006, p = .037, partial η² = .131 - There were **no** significant between-group differences at any timepoint - CIAT within-subjects: The simple main effect of time **was** significant, F(3, 24) = 8.692, p < .001, partial η² = .521. Pairwise comparison revealed a statistically significant difference between Baseline 0 and Baseline 1 in the CIAT group (p = .004). - MMAT within-subjects: Simple main effect of time was **not** significant, F(3, 36) = 1.305, p < .288, partial η² = .098

SAQOL

Dropped cases	2 CIAT and 2 M-MAT missing due to missing data points

Normality and outliers	Data was normally distributed at each group and timepoint as assessed by the Shapiro-Wilks test ($p > .05$)
	There were no outliers, as assessed by examination of Studentised residuals for values greater than ± 3
Equality of Variance	Data at each group timepoint met the assumption of equality of variances under Levene's test ($p > .05$)
Sphericity	Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, $X^2(5) = 2.454$, $p = .784$
Main effect of time	The main effect of time was **not** statistically significant $F(3, 60) = 1.091$, $p = .360$, partial $\eta^2 = .052$
Main effect of group	There was **no** main effect of group $F(1, 20) = .398$, $p = .536$, partial $\eta^2 = .019$
Interaction	There was **no** statistically significant interaction between the intervention and time, $F(3, 60) = .194$, $p = .900$, partial $\eta^2 = .010$

Scenario Test

Dropped cases	1 CIAT and 1 M-MAT missing due to missing data points
Normality and outliers	Data was not normally distributed at CIAT B0, or for MMAT at any timepoint, as assessed by the Shapiro-Wilks test ($p > .05$) and MMAT also displayed deviation from normality for the Post and Follow-Up QQ plots.

	There was one outlier for B0, B1 and Follow-Up, as assessed by examination of Studentised residuals for values greater than ± 3
Equality of Variance	Data at B1 did not meet the assumption of equality of variances under Levene's test (p = .011)
Sphericity	Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction, $X^2(5)$ = 9.771, p = .082
Main effect of time	The main effect of time was **not** statistically significant $F(3, 66)$ = 1.180, p = .324, partial η^2 = .051
Main effect of group	There was **no** main effect of group $F(1, 22)$ = .168, p = .686, partial η^2 = .008
Interaction	There was **no** statistically significant interaction between the intervention and time, $F(3, 66)$ = 0.614, p = .609, partial η^2 = .027
The above findings are likely not valid given the violation of assumptions. The Friedman test therefore run to determine change across time. Scenario Test scores were not significantly different at the four timepoints, $\chi^2(3)$ = 4.327, p = .228.	

Correlation analysis

Normality was checked using Shapiro Wilks test in RStudio:

```
Correlation <- read.csv(file = "I.

Correlation/Correlation.csv", header = TRUE, sep = ",")

apply(Correlation[,2:12], 2, shapiro.test)
```

Shapiro Wilks Test

| Variable	W statistic	p
Age	0.97732	.7975
MPO	0.93975	.1201
Severity	0.90157	.0145*
Semantic	0.82598	.0004*
Attention	0.73579	<.0001*
CogFlex	0.94645	.3164
STMtotal	0.9695	.6103
Wmtotal	0.92563	.0880
Nonverbal	0.82598	.0004*
WAB-R AQ change	0.92695	.0582
Treated Items Change	0.98056	.8747

Spearman's Rank Correlation

Conducted in Rstudio. Example script:

```
cor.test(Data$Age, Data$WABchange, alternative =
"two.sided", method = "spearman", use = "complete.obs")
```
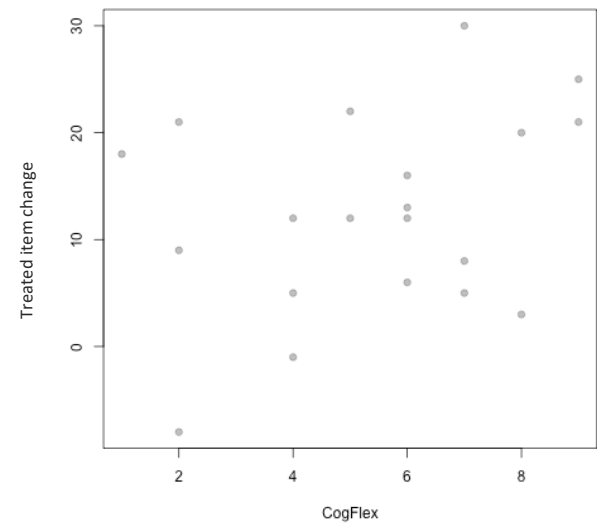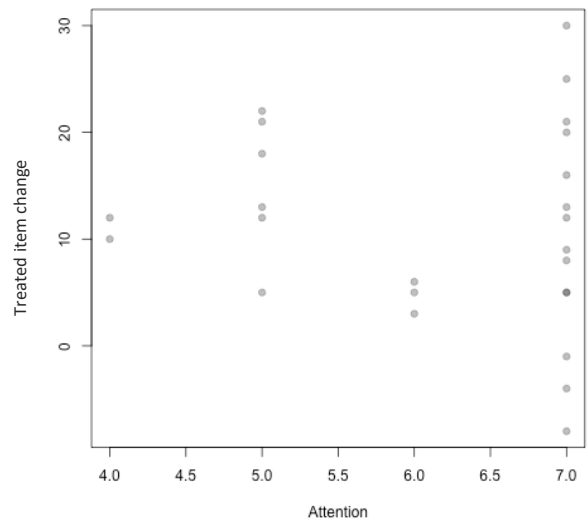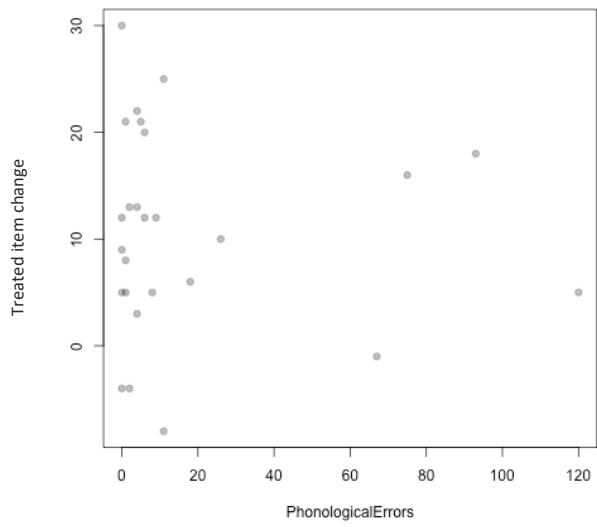
Scatter plots

WAB-R AQ change vs characteristics
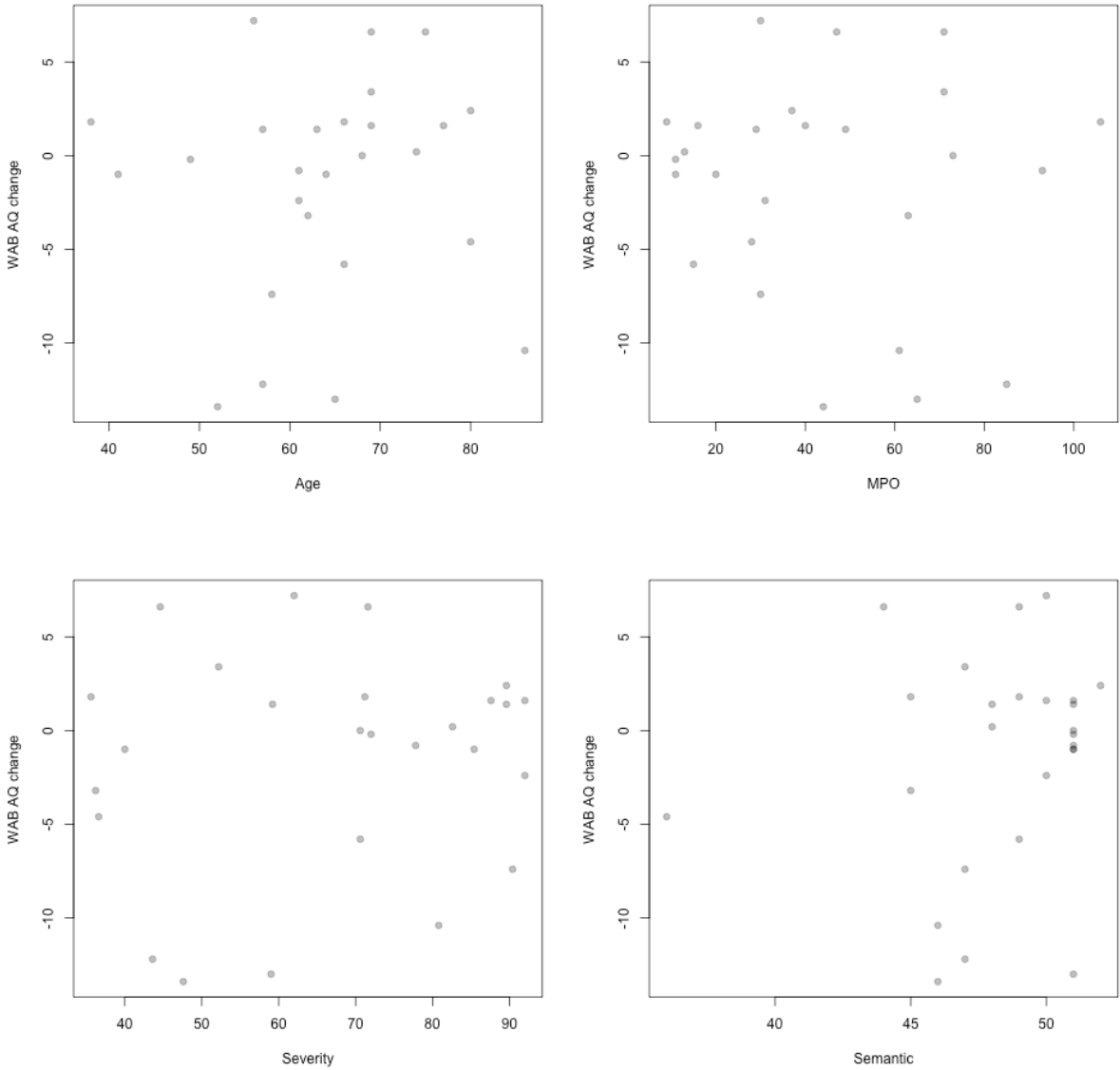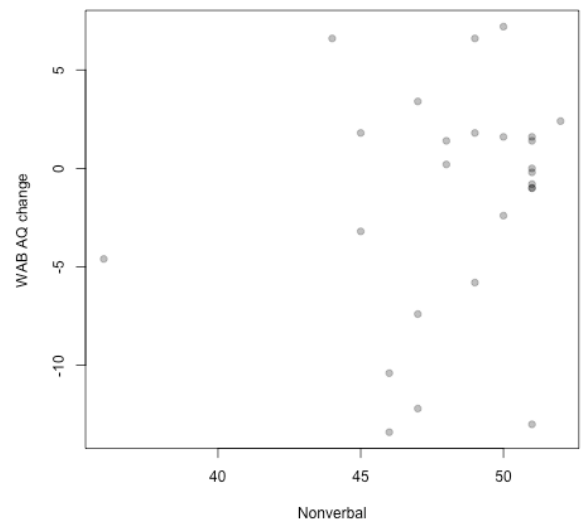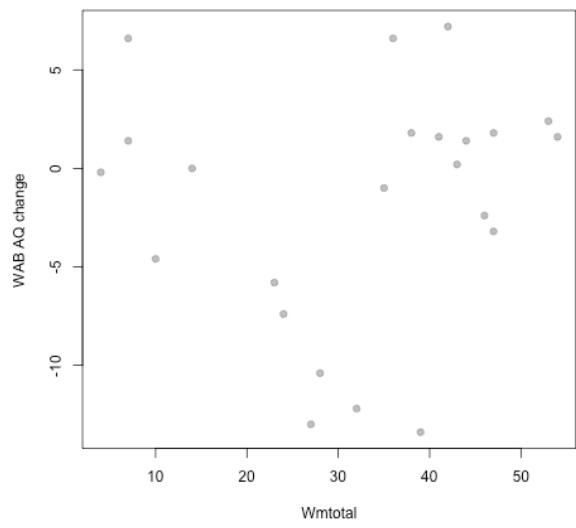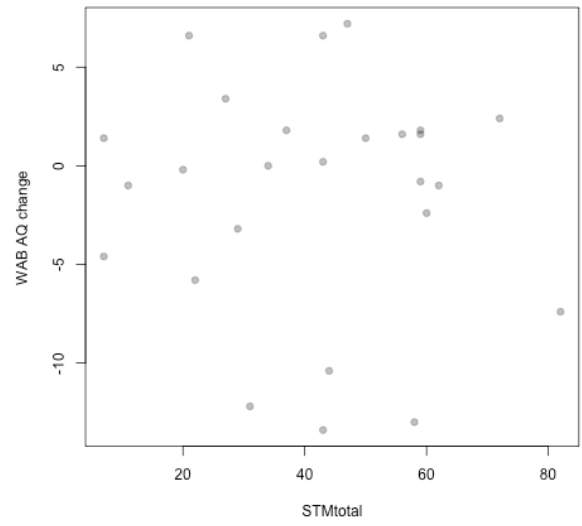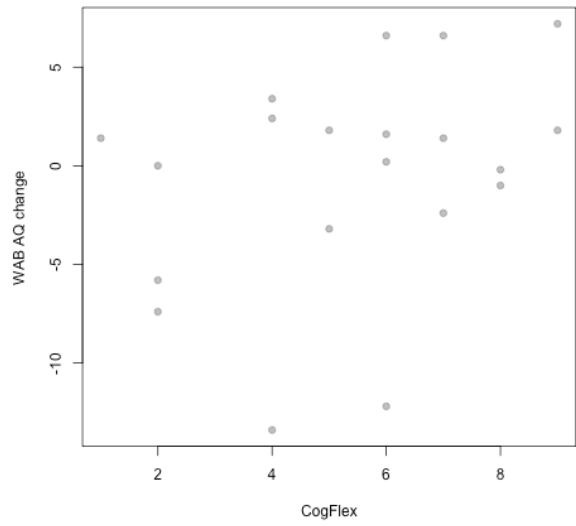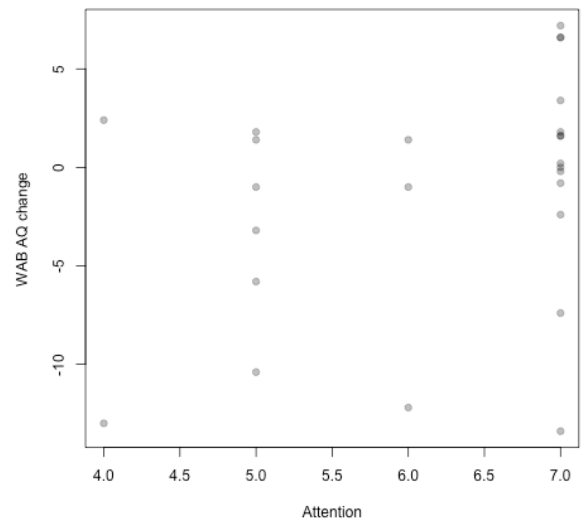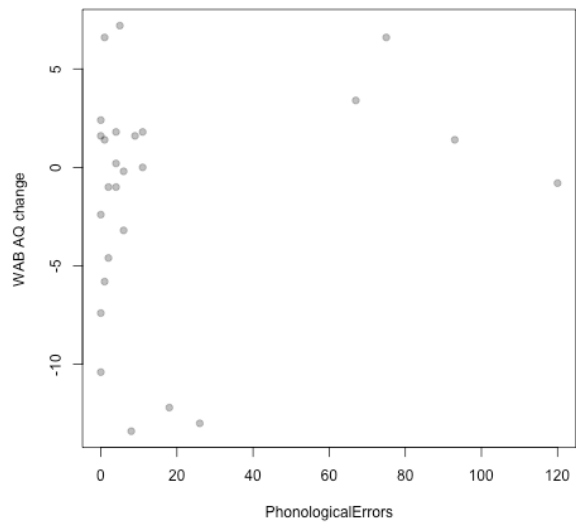
Treated items naming change vs characteristics
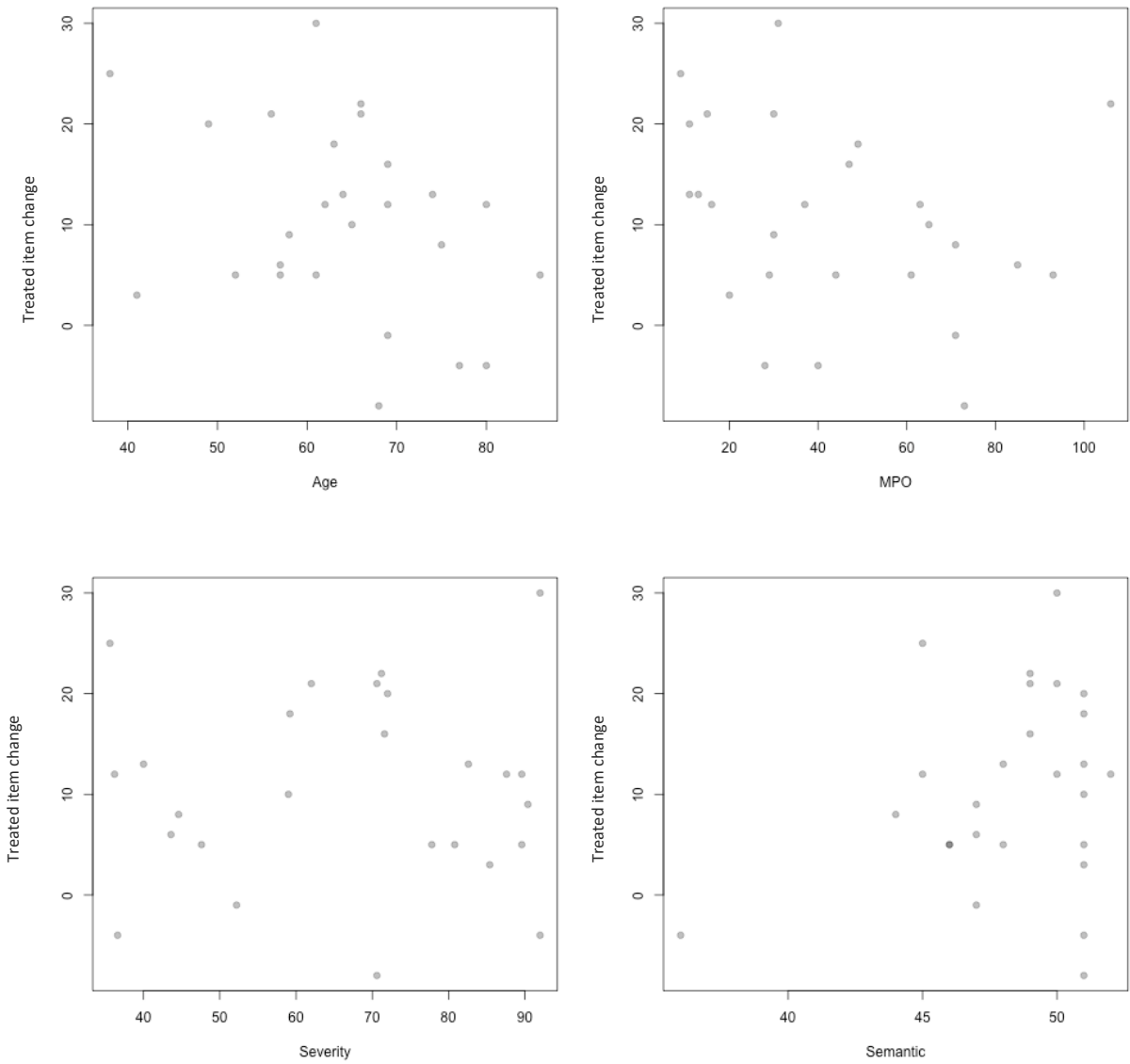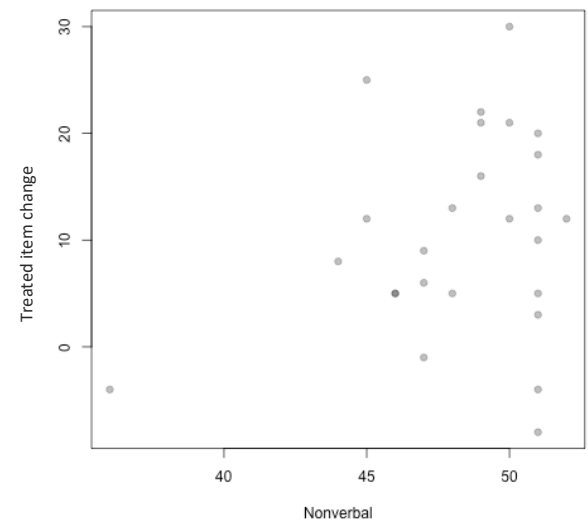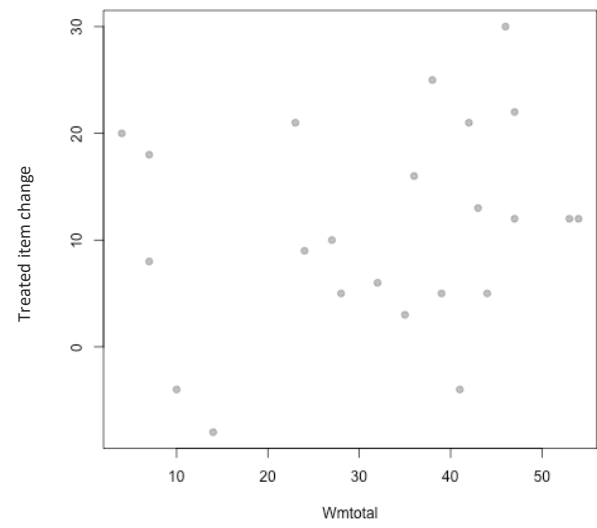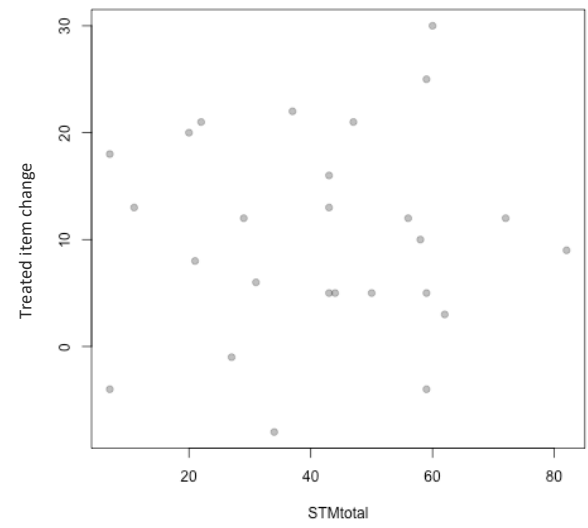
WAB-R AQ change vs treated items naming change