2023

Treatment integrity and differentiation in the very early rehabilitation in SpEech (VERSE) trial

Emily Brogan
Edith Cowan University

Natalie Ciccone
Edith Cowan University

Erin Godecke
Edith Cowan University

10.1080/02687038.2021.1988891
Brogan, E., Ciccone, N., & Godecke, E. (2023). Treatment integrity and differentiation in the very early rehabilitation in SpEech (VERSE) trial. *Aphasiology*, 37(1), 99-121.
https://doi.org/10.1080/02687038.2021.1988891

Aphasiology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/paph20

Treatment Integrity and Differentiation in the Very Early Rehabilitation in SpEech (VERSE) trial

Emily Brogan, Natalie Ciccone & Erin Godecke

To cite this article: Emily Brogan, Natalie Ciccone & Erin Godecke (2021): Treatment Integrity and Differentiation in the Very Early Rehabilitation in SpEech (VERSE) trial, Aphasiology, DOI: 10.1080/02687038.2021.1988891

To link to this article: https://doi.org/10.1080/02687038.2021.1988891

Published online: 29 Oct 2021.

APHASIOLOGY
https://doi.org/10.1080/02687038.2021.1988891

Routledge
Taylor & Francis Group

Treatment Integrity and Differentiation in the Very Early Rehabilitation in SpEech (VERSE) trial

Emily Brogan [a,b], Natalie Ciccone [a] and Erin Godecke [a,b]

[a]School of Medical and Health Sciences, Edith Cowan University, Perth, Western Australia; [b]Speech Pathology, Sir Charles Gairdner Hospital, Perth, Western Australia

ABSTRACT

Background: Key elements of treatment fidelity include treatment integrity (adherence to the treatment protocol) and treatment differentiation (the difference in treatment ingredients in the control and intervention groups). The Very Early Rehabilitation in SpEech (VERSE) trial established treatment fidelity at the macro level for key components of therapy.

Aims: To complete a detailed analysis of treatment integrity and differentiation at the utterance level of a therapeutic interaction.

Methods: This was an observational study of therapy videos collected as part of the VERSE trial. Participants were people with aphasia in the very early phase of recovery post stroke (n = 44) and speech-language pathologists (n = 25). Therapist video recorded sessions in the intensive arms of the trial (VERSE-prescribed therapy and Usual Care Plus) and 53 therapy videos (12%) were randomly selected for analysis. Therapy sessions were transcribed, and key measures reflective of therapeutic inputs and client acts were coded to determine treatment integrity and differentiation. A descriptive analysis and a Welch's t-test for unequal variances were used to analyse the sessional data.

Results: Therapists in the VERSE (prescribed intervention) arm of the study, were highly adherent to the treatment protocol at the utterance level (M = 97%). Treatment differentiation between the intensive conditions in this sample was not achieved for cueing and error handling suggesting the treatment delivered between groups was similar.

Conclusions: Within this sample, treatment integrity to the prescribed condition was maintained. Despite significant differences on a broad level, there was not significant differentiation in the therapy provided in the two arms of the trial at the utterance level. This result supports the null finding in effectiveness between the two intensive arms of the treatment as potential key measures were not different in dosage.

ARTICLE HISTORY
Received 17 May 2021
Accepted 30 September 2021

KEYWORDS
Aphasia; treatment fidelity; randomised controlled trial; stroke; therapy

CONTACT Emily Brogan e.brogan@ecu.edu.au Edith Cowan University, 270 Joondalup Drive, Joondalup 6027, Perth, Western Australia, Australia

Introduction

Treatment fidelity concept

Assessing and monitoring treatment fidelity is an important component of intervention studies; however, a single, widely agreed upon definition has not been developed (O'Shea et al., 2016). Most definitions, at a fundamental level, describe treatment fidelity as the degree to which the administration of a treatment corresponds to the specified protocol for the implementation of that treatment (Kaderavek & Justice, 2010). However, the definition of treatment fidelity has widened over the last twenty years to incorporate the concepts of treatment integrity, treatment differentiation, and treatment receipt (Borrelli et al., 2005). Treatment integrity or adherence refers to the extent to which the treatment is given as intended, while treatment differentiation determines if the therapy provided in the intervention and control groups is sufficiently different in relation to the key or "active" ingredients of the therapy provided (Moncher & Prinz, 1991). Treatment receipt is the processes implemented to monitor and improve the participant's ability to comprehend and perform the behaviours targeted in the treatment (Bellg et al., 2004). This broader definition of treatment fidelity reflects the increased complexity of research and the demand for an evolving evidence base to guide clinical practice. To address this, there is a need to reliably report on elements of intervention that are regularly omitted in treatment studies (Walton et al., 2017). Strengthening processes for the planning, implementation, monitoring, and reporting on treatment fidelity in clinical trials help to ensure that interventions are delivered as per the study protocol, and the integrity of the treatment is preserved.

Addressing treatment fidelity can help to explain study findings, minimise errors in interpreting study outcomes, and inform future modifications to the intervention (Hoffmann et al., 2014). Monitoring and evaluating treatment fidelity within studies assists in the interpretation of study outcomes and may increase confidence in the relationship between intervention components and the outcomes of a trial (Rixon et al., 2016). While treatment fidelity is broached with caution by some, due to the real-world applicability of strict treatment protocol procedures (Karas & Plankis, 2016), implementation of thorough treatment fidelity procedures has the aim of building a robust body of evidence for interventions. If the scientific basis for clinical practice is built on studies that have not effectively investigated treatment fidelity, then systematic reviews, meta-analyses, and clinical practice guidelines may be skewed (Wheeler et al., 2006). Additionally, studies with high fidelity monitoring and reporting have increased external and internal validity and are more likely to be replicable (Borrelli, 2011). High levels of fidelity reporting also allow comparisons to be made between treatments (Hildebrand et al., 2012; Moncher & Prinz, 1991; Resnick et al., 2005; Schlosser, 2002). While the importance of planning for and assessing treatment fidelity has been highlighted, treatment fidelity processes are not always incorporated within intervention studies (Bellg et al., 2004). For example, Borrelli et al. (2005) evaluated 342 behaviour change articles between 1990 and 2000 and found that overall, 54% of the studies did not report intervention fidelity.

Treatment fidelity in complex interventions

Complex interventions are "health service interventions that are not drugs or surgical procedures but have many potential active ingredients" (Oakley et al., 2006, p. 413). Behavioural interventions, such as those used in speech-language pathology (SLP), are complex, and treatments implemented have been described as "black boxes" referring to the fact that they may contain many potential active ingredients that shape patient outcomes (DeJong et al., 2005; Walker et al., 2017). Additionally, due to the inherent complexity of behavioural interventions, there may be more variation when different sites and providers are involved in intervention studies (O'Shea et al., 2016) resulting in potential issues with training of providers and delivery of treatment such as contamination between conditions or therapist drift. Therapist drift refers to a decrease in desired therapy skills or adherence to protocol over time. Treatment fidelity is especially relevant to behavioural change interventions due to this complexity (O'Shea et al., 2016). Investigating therapy fidelity measures encourages researchers to deconstruct and make explicit what is inside the "black box" of rehabilitation intervention (Hand et al., 2018). Currently, the incorporation of treatment fidelity processes is more established in the field of psychotherapy research compared to medical rehabilitation research (Hildebrand et al., 2012). Treatment fidelity has been made of such value in the psychotherapy literature that only research that includes substantial attention to this is considered as reliable evidence for the efficacy of a treatment. In 2008, it was recommended that psychotherapy sessions should be videotaped and a random selection of 20% should be rated for treatment adherence and therapy competence (Öst, 2008). Although not as established as the psychotherapy literature, increasing attention is being given to treatment fidelity in rehabilitation interventions in stroke. Similar to psychotherapy, stroke interventions are often behavioural in nature, delivered via a therapist and are likely to have a range of potentially active ingredients.

Treatment fidelity processes should be incorporated when designing a study, when implementing the study and also when reporting the findings (Borrelli, 2011; Brogan et al., 2019). The importance of assessing treatment fidelity has been emphasized in recent guidelines and recommendations for evaluating complex interventions (Rixon et al., 2016); however, guidelines are needed to define treatment fidelity concepts and to provide standardisation regarding key aspects of treatment fidelity (Gearing et al., 2011). The Treatment Fidelity Workgroup of the National Institutes of Health Behaviour Change Consortium (Bellg et al., 2004) reviewed treatment fidelity practices used within the identified literature and developed recommendations to embed treatment fidelity practices within intervention studies. Recommendations for addressing treatment fidelity in behaviour change studies were outlined across five main areas: study design, training providers, delivery of treatment, receipt of treatment, and enactment of treatment skills (Bellg et al., 2004; Borrelli, 2011). Additionally, Gearing et al. (2011) provided a treatment fidelity guide, identifying four core components from the treatment fidelity literature with less of an emphasis on treatment enactment, the fifth category in the Bellg et al. (2004) framework. Gearing et al.'s (2011) guide provides a chronological outline incorporating study design, training, monitoring intervention delivery, and monitoring intervention receipt. The guide encourages researchers to address some broader areas of study design,

such as the overall framework. The elements included in both the Bellg et al. (2004) framework and Gearing et al.'s (2011) guide are very similar, with Bellg et al. (2004) being the most widely applied in research (O'Shea et al., 2016).

Documenting fidelity involves investigating therapy effectiveness, including key ingredients and the dosage delivered of these (Kaderavek & Justice, 2010). Specification of therapy ingredients is central to all areas of Bellg et al.'s (2004) framework, and dosage is important for determining whether the therapist adhered to the protocol and whether the intervention groups received different amounts or types of intervention. Warren et al. (2007) proposed a comprehensive model for defining dosage and calculating cumulative intervention intensity. The model included the term "dose form" which is defined as the task or activity within which the teaching episodes are delivered and it may contain the important active ingredients of the intervention. Potential active ingredients may involve therapeutic inputs (therapist behaviours) and client acts (client behaviours) (Baker, 2012). When designing a treatment, it is important that researchers give significant consideration to the theoretical underpinnings of the treatment so that the potential active ingredients and the required dosage are identified and then adequately monitored and evaluated (Borrelli, 2011). As such, the planning of treatments may be as complex and detailed as the treatment itself. If researchers fail to recognise potential factors that may facilitate change in the participants at the design stage, the appropriate data may not be collected, evaluated, and reported (Walker et al., 2017). Specification of ingredients in the study design allows therapists to be trained in delivering those ingredients and decisions on the monitoring of implementation of the ingredients in therapy can be made. How to monitor treatment delivery presents a challenge as there are few validated tools to use for the investigation of treatment fidelity in behavioural interventions (Borrelli et al., 2005).

Treatment integrity has been incorporated as a feature of behavioural intervention studies (Damschroder et al., 2016; Hildebrand et al., 2012; Seng & Lovejoy, 2013; Thomas et al., 2016) with stricter protocol adherence linked to improved study outcomes (O'Donell, 2008; Perepletchikova & Kazdin, 2005). Treatment integrity is commonly investigated by measuring the therapist's adherence to the treatment protocol through the use of structured observations. These observations are carried out in person or via recording and typically involve the completion of checklists (Borrelli, 2011). As a gold standard, the checklist should be developed using a priori coding categories, which reflect the potential factors that are identified during the development of the study design (Kaderavek & Justice, 2010). The checklist may also be used for the collection and analysis of data for treatment differentiation (Hildebrand et al., 2012). One criticism is that direct observation may be more prone to bias because the treatment variable is predefined and an observer might report what the therapist is supposed to do rather than what actually happened (Schlosser, 2002).

Treatment fidelity information should be included in detailed treatment descriptions and in the reporting of trial results (Moher, 2018) although this is infrequently done (Conlon et al., 2020). Current papers may limit their investigation and reporting of treatment fidelity to study design or protocol adherence only (O'Shea et al., 2016; Spell et al., 2020) and not addressing the five areas of treatment fidelity (Bellg et al., 2004). Several guidelines now exist to help authors include important intervention details including information on treatment fidelity (Moher, 2018). For example, checklists such as the Template for Intervention Description and Replication (TIDieR) (Hoffmann et al.,

2014) were established to encourage more complete reporting of treatments and to address the "remarkably poor" (p. 1) intervention description quality. While the TIDieR checklist includes general items related to the therapy such as task selection, therapy location, and dosage, it also includes treatment fidelity-specific items related to planned (item 11) and actual (item 12) treatment fidelity (Hoffmann et al., 2014). These items go beyond simple receipt of the intervention and refer to how well the intervention was received or delivered (Hoffmann et al., 2014). With a recent emphasis on treatment fidelity, publications addressing this area specifically within clinical trials have been published (Behn et al., 2018; Carragher et al., 2019; Conlon et al., 2020; McLennon et al., 2016; Resnick et al., 2011; Spell et al., 2020) providing exemplars for the reporting of treatment fidelity analyses. This has included the specific quantitative reporting of protocol adherence.

Treatment fidelity in aphasia

The adequacy of SLP randomised controlled trials (RCT) intervention descriptions, across all SLP practice areas, not just within aphasia trials, was assessed against the TIDieR checklist (Hoffmann et al., 2014) with 46% of studies included reporting on treatment fidelity (Ludemann et al., 2017). Hinckley and Douglas (2013) investigated treatment fidelity reporting within the aphasia literature and found 21/149 (14%) of studies reported treatment fidelity between 2002 and 2011. More recently Brogan et al. (2019) reported 9/42 (21%) aphasia RCTs published between 2012 and 2017 explicitly reported on treatment fidelity processes. The least addressed aspect of treatment fidelity was ensuring participants used the skills gained in treatment in appropriate life settings, with only two (2%) articles including this. Poor monitoring of treatment fidelity was also reported specifically in verb treatments (Hickin et al., 2020).

Increasingly published research protocols are using the current gold standard technique of video recording therapy sessions and assessing these according to a priori criteria. These studies include the Aphasia Action Success Knowledge (ASK) trial (Worrall et al., 2016), Predicting and Promoting Sub-acute Aphasia Recovery (PAPAR) (Copland, 2017), SUpporting well-being through PEeR-Befriending (SUPERB) trial (Behn et al., 2018) and the COMPARE trial (M. Rose et al., 2019). Other studies such as BIG CACTUS (Palmer et al., 2015) have advantages with computerised therapy in obtaining fidelity data via logs without relying on therapist reports or monitoring therapist participant interactions. An additional level of therapy fidelity is the monitoring of recorded sessions and providing feedback to the therapist to change aspects of their therapy delivery and increase protocol adherence, while treatment is still being delivered. ASK (Carragher et al., 2019), COMPARE (M. L. Rose et al., 2019) and VERSE (Godecke et al., 2020b) have incorporated this element.

Very Early Rehabilitation in SpEech (VERSE) trial

This study involved the analysis of a subset of data from the VERSE trial. The findings of the VERSE trial indicated that communication outcomes were equivalent for usual care therapy and early intensive therapy at 12 and 26 weeks post stroke (Godecke et al., 2020b). Usual care was documented to be therapy provided an average of two to three

times per week for approximately 40 min (9.5 hours in total in the first 50 days after stroke) and this resulted in improved communicative ability after stroke as measured by the outcome on the Western Aphasia Battery Revised – Aphasia Quotient (WABR-AQ) (Kertesz, 2006). Intensive intervention (22 hours in total in the first 42 days after stroke) provided within the Usual Care-Plus and the prescribed VERSE intervention arms of the trial did not cause harm but did not provide a statistically significant benefit to participants over and above the usual care regimen (Godecke et al., 2020b).

The interventions in the trial were as follows:

(i) Usual Care: Participants randomised to this group received care that is typical for aphasia management in the Australian and New Zealand healthcare settings and was at the discretion of the treating SP. It also included management of other speech pathology impairments such as dysphagia, dysarthria, and/or apraxia of speech. Only direct aphasia therapy time was included in the analysis for the primary VERSE result. Usual care therapy was recorded over a period of 20 working days.

(ii) Usual Care-Plus: Participants received treatment that is typical of direct aphasia therapy, at the discretion of the treating SP as per the Usual Care group, but with a defined intensity therapy regimen of daily sessions for 45–60 minutes duration over 20 sessions. Direct aphasia therapy included 1:1 impairment-based therapy, impairment-based computer training, social training, group impairment-based therapy, group social training, and Augmentative and Alternative Communication (AAC) training.

(iii) VERSE intervention: The intensity of this arm of therapy matched that of the Usual Care-Plus arm but the intervention was impairment-based, prescribed, and standardised according to a specific VERSE intervention protocol. As per treatment fidelity recommendations, when planning the prescribed intervention, the theoretical underpinnings of the treatment were highlighted, allowing these to be monitored and evaluated. This intervention was founded on principles to promote neurorecovery: i) massed practice, ii) error-free learning, iii) task complexity, iv) salience, and v) maximising communicative success.

The current study

Treatment fidelity processes were developed and implemented within the VERSE trial. Treatment integrity and differentiation were established in the main study, at a broad level, between treatment conditions. However, with a trial of this size, including 8915 completed therapy sessions, fine-grained analysis of all therapy sessions was not feasible. The researchers placed importance on investigating a proportion of the sample with greater detail. In the present treatment fidelity study, we investigated the dose form, as per Warren et al. (2007), given in the two arms of the trial to establish treatment integrity and differentiation. In the VERSE trial, the intensity was prescribed for the two intervention arms at twenty sessions; however, the dose form given and received within a session was not prescribed. This therefore provides an opportunity to investigate treatment integrity and differentiation. According to Baker (2012) dose form involves both the therapeutic inputs and client acts of a task. Therapeutic inputs refer to behaviours by the therapist

within the task that are thought to be therapeutic in nature. Similarly, client acts are behaviours that may contribute to the therapeutic effect. Integrity and differentiation results have been framed according to these elements within the current study. As the interventions provided within the VERSE RCT were behavioural and interactional in nature, treatment integrity and differentiation were investigated at the utterance level in this study, with the following aims:

(i) To determine treatment integrity (protocol adherence) for the task (dose form) and therapeutic inputs (cueing and error handling) to the VERSE intervention protocol.
(ii) To determine the level of treatment differentiation (dose form given within sessions) for the task, therapeutic inputs (cueing) and client acts (verbal output, errors), between the intensive conditions (Usual Care-Plus and VERSE) at the utterance level.

Method

Participants

This study used therapy videos collected as part of the VERSE RCT. Participants were recruited to the VERSE trial if they had aphasia secondary to an acute intracerebral haemorrhage or ischaemic stroke; were 18 years of age or over; had corrected hearing and vision; were medically stable at or before 14 days post stroke; and could participate in aphasia therapy in English without the need for an interpreter. Exclusion criteria included pre-existing aphasia prior to admission into the hospital; a history of progressive neurological disease, neurosurgery, major depression, subdural or subarachnoid haemorrhage; and an inability to maintain alertness for 30 consecutive minutes at 14 days post stroke. VERSE recruited 246 participants, 81 in Usual Care, 82 in Usual Care-Plus, and 83 in VERSE, across Australia and New Zealand. Participants were assessed at a baseline of 12 and 26 weeks post stroke on a range of impairment, psychosocial, and economic measures. The primary outcome measure for VERSE was the WABR-AQ (Kertesz, 2006) at 12 weeks post stroke.

The participants for the current study represent a sample subset from the broader VERSE RCT. As part of the treatment fidelity processes embedded within the VERSE RCT, VERSE and Usual Care-Plus therapists were required to video record one therapy session per week, resulting in four or five recordings per participant (please see below for more details on the VERSE RCT treatment fidelity processes). For practical reasons, it was suggested to therapists to video record session numbers five, 10, 15, and 20, however, therapy videos for any session were accepted.

An independent research assistant used a computer-generated block randomisation to select therapy videos for this study. This process was stratified for aphasia severity (mild, moderate, and severe aphasia as determined by the WABR-AQ at baseline). For the videos to be included within the study, participants were required to have outcome scores at 12 weeks (primary outcome) and 26 weeks on the WABR-AQ and have completed the full treatment protocol. Additionally, videos needed to be at least 40 min in length, not contain the present author as the therapist and be playable on Windows Media Player. The sample comprised 53 videos, which was 12% of the 434 therapy videos received

Table 1. Participant demographic and stroke characteristics.

Demographic	Whole group (n = 44)[a]	VERSE (n = 25)	Usual Care-Plus (n = 19)	VERSE High Intensity cohort (n = 164)[b]
Age, Mean (SD)	72.0 (14.8)	75.0 (14.2)	68.0 (14.7)	75 (18)
Female	19 (43%)	12 (48%)	7 (37%)	80 (49%)
Oxford Stroke Classification [c]				
Haemorrhagic	3 (7%)	2 (8%)	1 (5%)	13 (8%)
PACs	34 (77%)	18 (72%)	16 (84%)	110 (67%)
PoCs	1 (2%)	0 (0%)	1 (5%)	6 (4%)
TACs	6 (14%)	5 (20%)	1 (5%)	35 (21%)
Baseline WABR-AQ Severity				
Mild	15 (34%)	8 (32%)	7 (37%)	47 (29%)
Moderate	13 (30%)	7 (28%)	6 (32%)	49 (30%)
Severe	16 (36%)	10 (40%)	6 (32%)	68 (41%)
Modified Rankin Scale Score (mRS)				
Low disability (mRS 0–2)	6 (14%)	5 (20%)	1 (5%)	16 (10%)
High disability (mRS 3–5)	38 (86%)	20 (80%)	18 (95%)	148 (90%)
National Institutes of Health Scale Score, Mean (SD)	9.1 (7.1)	9.4 (7.5)	8.7 (6.5)	9 (6)

Note. [a] Within this substudy, 53 videos were selected through randomisation with 8 participants having more than one video in the sample. Therefore, 44 different participants are included in this demographics table. [b] This column represents the VERSE high intensity cohort as a whole from the primary study for comparison purposes. [c] PACs = Partial Anterior Circulation syndrome; PoCs = Posterior Circulation syndrome; TACs = Total Anterior Circulation syndrome.

through the trial. See Table 1 for participant demographics included in this study. Of the 53 randomly selected videos, seven participants had two videos in the sample (total 14 videos), one participant had three videos in the sample (total of 3 videos), and the remaining 36 participants had one video each (total 36 videos). This resulted in the inclusion of 44 different participants. Twenty-seven SLPs are also included in the videos. SLPs were hired specifically for the trial and were required to be eligible for membership to Speech Pathology Australia. All treating SLPs underwent training (3 hours) and received procedural and training manuals relevant to the arm of therapy they were providing. Clinical support was provided by trial clinical staff as required throughout the trial. VERSE therapists were then given additional training (2 hours) to administer the prescribed therapy (Godecke et al., 2020a).

Treatment fidelity within VERSE

All therapists within the trial completed standardised therapist training, including the provision of manual and logged session data. Additionally, therapists in the intensive arms of the trial (Usual Care-Plus and VERSE) were required to video record a set number of sessions across the intervention period. The Therapy Fidelity Monitor (TFM) checked that therapy for participants in the intensive arms of the study was commenced on or before day 15 post stroke and that it did not continue beyond 50 days post stroke. The duration and frequency of each session within the intensive arms of the trial were also monitored to ensure it was 45–60 minutes of direct aphasia therapy, for a maximum of five sessions per week for 20 sessions and completed within 4 weeks. For the prescribed VERSE intervention arm, key therapy ingredients were identified within the study design. This guided therapist training and the monitoring of therapy delivery. VERSE intervention therapists received a specific therapy manual and received one on one support in order to

implement the therapy as prescribed. The TFM monitored that prescribed targets as per the protocol were met, including delivering conversation-based therapy at the target goal level. They also monitored major VERSE protocol elements broadly, such as the timing and type of cueing used by the therapist and scaffolding of correct productions. This resulted in an overall rating of the session as adherent or non-adherent to VERSE protocol. The TFM and Therapy Fidelity Co-ordinator was responsible for feeding back to the therapists about any deviations from the VERSE protocol or any general questions that arose about the treatment procedures.

Procedure

Transcription

In this study, each video was transcribed verbatim and utterances segmented following the guidelines provided within the Systematic Analysis of Language Transcripts (SALT) (Miller, 2008) software and as per SALT guidelines (available at http://saltsoftware.com/resources/tranaids).

Coding

During transcription, codes developed specifically for this research were applied at the word and utterance level of the transcript for both therapeutic inputs and client acts. Variables of interest were selected as possible key points of difference between treatments based on the theoretical underpinnings of the VERSE therapy protocol. For example, the quantity and accuracy of the participant's production was a key consideration as the VERSE protocol incorporated the principle of maximising communicative success through the use of salient communication-based tasks and error minimisation strategies (error-free production). Rationales for the coding systems are explained in the relevant sections below.

Treatment integrity – protocol adherence.

Task (dose form). Integrity or adherence to the VERSE prescribed treatment required therapists to implement conversation-based tasks aimed at eliciting an accurate phrase structure at the appropriate level for the participant. This appropriate level is called a "goal level". The rationale for verbal conversation tasks in the VERSE protocol was i) that improvement in therapy is experience dependent; ii) a lack of experience of practicing a task may result in learned non-use; and iii) conversation is salient to the person with aphasia, a key neuroplasticity principle (Pulvermuller & Berthier, 2008).

A summary of the goal levels of the prescribed treatment and subsequent coding is presented in Table 2. For example, if goal two was chosen as the starting point for the participant based on baseline data, the therapist was trained to facilitate the production of verbal output at a minimum of single word level. If the participant independently achieved 80% success during the session, they progressed to the next goal. Utterances from the participant were coded as either at, above, or below the targeted goal level. The target goal level of each session was coded, to account for incremental improvement throughout the treatment period. These codes were then counted to allow a calculation of the percentage of utterances that were at or above the target goal level. The mean length of utterance (MLU) during the therapy session, as per the SALT analysis, was also

Table 2. Summary of coding for task adherence.

Goal level	Definition	Utterance level coding	Minimum MLU Expected in Session
1a	Receptive: Identification of verb pictures from spoken words	Any verbal output is appropriate. Cannot mark as being below goal level. Can mark as above goal level for any verbal output that is not an error.	0
1b	Receptive: Identification of noun pictures from spoken words		0
1c	Receptive: Identification of adjectives pictures from spoken words		0
2	Verbal production of single words	If not responding verbally then mark as below the appropriate goal level. Can mark as above goal level if above this level.	1
3	Verbal production of two element phrases or clauses	If responding at single-word level including yes/no responses then mark as below the appropriate goal level. Can mark as above goal level if above this level.	2
4	Verbal production of three element phrases or clauses	If responding below three element clause then mark as below the appropriate goal level. Can mark as above goal level if above this level.	3
5	Verbal production of complex clauses and/ or phrases	If responding with three element clauses or less then mark as below the appropriate goal level. Can mark as above goal level if above this level.	4
6	Verbal production of complex phrases (verb and noun) and clauses	If responding below a complex phrase mark as below the appropriate goal level. Can mark as above goal level if above this level.	5
7	Verbal conversation about familiar topics	Cannot mark as above goal level. At conversation level utterances must be above a complex phrase otherwise they will be marked as below the appropriate goal level.	>5
8	Verbal conversation about unfamiliar topics		>5

Note. MLU = Mean Length of Utterance

used to determine protocol adherence to see whether it matched the predetermined goal level of the session. This measure is average and broader than the utterance by utterance analysis. The MLU analysis provides an overall measure of the mean length of utterance for the entire session, whereas the utterance analysis was a point-to-point count.

Therapist inputs. The video transcriptions for the VERSE intervention group were coded according to therapeutic inputs that adhered to VERSE treatment elements and those that did not. All therapist utterances that did not meet the criteria for non-adherence were deemed adherent to the therapy protocol. After coding and counting, a percentage of therapeutic inputs that adhered to VERSE treatment elements for each session was established.

There was a specific focus on how the therapist responded to participant errors with cueing. Cues were used by the therapist to promote accuracy and participation by assisting in word retrieval and/or accurate speech production. In the VERSE protocol, therapists were required to establish the type of cue that is most effective for the participant. Therapists were encouraged not to repeat cues that were ineffective. An utterance would be marked as a protocol deviation if the therapist used the same cue type more than once for the same stimulus and it was unsuccessful.

VERSE therapy incorporated the principle of error-free practice. Therefore, therapists were encouraged to limit the participant to no more than three errors on a stimulus. The utterance was marked as a protocol deviation if the participant made more than three

errors for one stimulus and a model was not given after the third error. Every subsequent cue and utterance for the same stimulus by the therapist, which allowed the participant to continue making errors, was marked as a protocol deviation.

Treatment differentiation.

Task (dose form). Tasks were coded with a description of the activity (e.g., picture naming), the target language process (i.e., expressive verbal, expressive written, receptive auditory, or receptive written) and the target phrase level (i.e., single-word level, sentence, or conversation) to enable dose form differentiation. For conversation to be included as a task, it needed to be the target of therapy, including therapeutic inputs such as cueing and feedback. Filler conversation and rapport building at the beginning or end of the session was not included.

Therapist inputs. Measures related to therapeutic inputs were coded in order to assess whether the treatment provided in the two intervention groups was significantly different. The therapeutic inputs of interest were related to cueing. The specific cues focused on in this study were phonemic, semantic, orthographic, visual, forced alternative, sentence completion, articulatory placement, and direct model cues. A cue was coded if a relevant cue was used by the therapist. If the cue was successful in eliciting an appropriate response from the participant, it was marked as a cue used with success. Turkstra et al. (2016) suggested investigating ingredients, such as handling of errors and their effect on target attainment, more generally across a therapy session rather than looking at the impact of these ingredients separately for each activity. This approach was used within this study and so the treatment sessions were examined as a whole to analyse therapist tailoring of cues and whether the use of the cues resulted in successful participant/client performance. Codes were then counted to provide totals and percentages for each session. In addition, calculations were performed by the SALT software (total therapist utterances, total therapist words, and utterances per minute) and used in the analysis.

Client acts. Measures related to client acts were coded to assess whether the participants' response to the treatment provided in the two intervention groups was significantly different. Aphasia treatment approaches are often described as "error full" or "errorless" referring to whether patients are encouraged to attempt target attainment irrespective of the error quantity or whether errors are minimized to avoid strengthening of associated negative neural networks. Reducing errors in therapy relates to Hebbian theory surrounding coactivation of neurons and that a mistake occurring while producing a word may wire together these two behaviours (Hebb, 1949; Varley, 2011). It aims to avoid therapy being the practice of errors (Varley, 2011). Errorless learning may reflect implicit learning techniques that may be suited to the motor learning required in speech production (Page, Wilson, Norris 2006). As such, the codes of focus reflected the quantity and accuracy of utterances. Each utterance was marked as error free if it contained no errors. An error made by the participant was marked at the word level. It was further coded for error types including filled and unfilled pauses, phonology, perseveration, circumlocution, semantic, grammatical, repetition fluency, receptive, and unintelligible errors. Errors that were self-corrected prior to therapist intervention were coded as self-corrections. The codes were then counted to provide totals and percentages for each session. In addition, calculations

were performed by the SALT software and used in the differentiation analysis. These were mean length of utterance (MLU) in both words and morphemes, total verbal utterances, total words, words per minute (WPM), utterances per minute (UPM) and mean turn length in words.

Additionally, calculations were performed manually according to the formulas in Table 3 and used in the analysis.

Statistical analysis

Descriptive statistical analyses for treatment integrity and differentiation measures were completed. As the session length of the videos was not equal, the time of the session was used to standardise the data [measure/session time (mins)] for the following measures: total error-free utterances, total errors, spontaneous verbal output, total self-corrections, total verbal utterances, total words, utterance-level errors, word-level errors, total cues used with success, total therapist utterances, and total therapist words. A Welch's t-test for unequal variances was performed using the above measures to determine whether Usual Care-Plus and VERSE conditions were significantly different. This test was used as it is robust to unequal sample sizes. A Bonferroni correction was applied for multiple comparisons and significance was set at $p < 0.001$.

Reliability

Six videos (one from each aphasia severity in Usual Care-Plus and VERSE groups) were re-coded for inter- and intra-rater reliability. This comprised 11% of the total sample in this study. Reliability was established using the intra-class correlation coefficient (ICC) in SPSS (IBM Corp, 2015) with a consistency 2-way mixed-effects model. The ICC was established on the four key measures of interest in the study – error-free utterances, total errors made, (iii) total self-corrections, and (iii) the total number of cues used with success by the therapist. Koo and Li (2016) guidelines for reporting the ICC have been adhered to in the tables below.

Table 3. Client act calculations for differentiating treatments.

Calculation	Formula
Spontaneous verbal utterances	The number of verbal utterances that were generated spontaneously, e.g., not a repetition of a model or a gesture (verbal utterances) – (repetitions + gestures)
Utterance Level Errors	Total of the number of errors that occur at the utterance level based on codes in the transcript.
Word Level Errors	Total of the number of errors that occurred at the word level based on codes in the transcript.
% errors self-corrected	Self-corrections/Total errors x 100
Average no. of errors per utterance	Verbal errors/Verbal utterances
Average no. of utterances per error	Verbal utterances/Verbal errors
% total words with errors	Errors/Total words x 100
Verbal utterances	Total utterances – Gestures

Inter-rater reliability

As per Supplement 1, three of the measures had ICCs in excess of 0.90, and, therefore, inter-rater reliability can be said to be excellent. The self-corrections measure was lower; however, ICCs in the range of 0.75–0.90 remain in the good reliability range (Koo & Li, 2016).

Intra-rater reliability

Intra-reliability is rated as excellent, given the ICCs are all in excess of 0.90 as per Supplement 2.

Results

Treatment integrity

Protocol adherence-task (dose form)

The majority of participant utterances within tasks were at the target goal level. The goal level refers to the target verbal output level during the session and ranges from a target MLU of 0 to >5. An average of 67% of utterances was at the target goal level within a session. On average, 8% of utterances by the participant in a session were above the target goal level of the session. Table 4 provides descriptive statistics for the therapists' adherence to the target goal level for the VERSE condition measured at the utterance level.

A higher level of dosage adherence was seen using the MLU analysis compared to the utterance by utterance dosage adherence measures above. The MLU analysis provides an overall measure of the mean length of utterance for the entire session, whereas the dosage adherence measures above were point-to-point count. Of the 28 VERSE sessions analysed, 24 (86%) had an MLU at or above the target goal level for the session. This represents a high level of protocol adherence.

Protocol adherence – therapeutic inputs

Therapists produced, on average, 671 utterances (SD = 278.1) per session in the VERSE condition. There were on average 15 protocol deviant utterances per session (inclusive of the same cue and error deviations) with 97.6% of utterances by the therapist being adherent to the VERSE intervention protocol. Where deviations were observed, allowing the participant to make more than three errors was observed most frequently, occurring on average ten times per session (1.5%). Table 5 summaries the protocol adherence data with descriptive statistics.

Treatment differentiation

Task (dose form)

Descriptive statistics for the tasks completed in this study are presented in Table 6. The prescribed protocol was designed to focus on verbal language output and encourage conversation. The VERSE therapy group had a greater percentage of verbal and conversation-level tasks.

Table 4. Measures of central tendency for protocol adherence to task.

Measure	VERSE (n = 28 videos)
Error free and at appropriate goal level	
Mean (SD)	197.2 (124.6)
Median	172.5
Interquartile range (IQR)	161
% total utterances	38.6%
Error free not at appropriate goal level[a]	
Mean (SD)	128.4 (87.6)
Median	122.0
IQR	114.0
% total utterances	25.1%
Contains an error and at the appropriate goal level	
Mean (SD)	185.7 (123.9)
Median	149.0
IQR	129
% total utterances	36.4%
Contains an error and not at appropriate goal level [a]	
Mean (SD)	64.0 (56.3)
Median	48.0
IQR	55.0
% total utterances	12.5%
Above goal level	
Mean (SD)	36.2 (31.7)
Median	31.0
IQR	63.0
% total utterances	7.0%
% utterances above goal level	
Mean (SD)	8.2 (8.5)
Median	6.1
IQR	12.1
Not at target goal level	
Mean (SD)	192.4 (124.3)
Median	159.0
IQR	175
% utterances at correct goal level	
Mean (SD)	67.2 (24.4)
Median	66.3
IQR	38.2

Note. [a] Not at appropriate goal level includes utterances that were below the target goal level only.

Table 5. Means and standard deviations for session protocol adherence – therapeutic inputs.

Measure	VERSE (n = 28 videos)
Total deviant utterances Mean (SD)	14.9 (17.3)
Protocol deviant utterances (same cue)	4.3 (7.5)
Protocol deviant utterances – x 3 errors	10.0 (12.5)
Protocol adherent behaviours	656.3 (273.6)
% Utterances that adhere to protocol	97.6 (3.2)

Therapist inputs

Descriptive statistics for therapeutic inputs are presented in Table 7. To maintain context, the means and standard deviations are presented in raw format prior to time adjustment [measure/session time (mins)]. After statistical analysis, the groups were not found to be significantly different with $p > 0.001$.

Table 6. Tasks for overall treatment differentiation.

Measure	Whole group (n = 53)	VERSE (n = 28)	Usual Care-Plus (n = 25)
Total number of tasks	152	56	96
Total number of verbal tasks (%)	114 (75%)	48 (86%)	66 (69%)
Total verbal tasks at single word level (%)	53 (46%)	13 (27%)	40 (61%)
Total verbal tasks at conversation level (%)	16 (14%)	12 (25%)	4 (6%)

Table 7. Descriptive and t-test statistic for treatment differentiation – therapeutic inputs.

Measure	VERSE (n = 28 videos)	Usual Care-Plus (n = 25 videos)	*p* value
Therapist total utterances Mean (SD)	671.1 (278.1)	679.6 (226.0)	.577
Cues used with success by therapist	29.6 (27.4)	38.5 (48.5)	.424
Total cues used by therapist	77.3 (72.3)	86.6 (74.8)	.631
% Cues that were successful	41.5 (20.0)	38.2 (26.9)	.625
Therapist total words	2687 (948.5)	2848.6 (857.9)	.518
Therapist utterances per minute	12.5 (4.9)	14.5 (7.6)	.283

Table 8. Descriptive and t-test statistic for treatment differentiation – client acts.

Measure	VERSE (n = 28 videos)	Usual Care-Plus (n = 25 videos)	*p* value
Total Verbal Utterances Mean (SD)*	474.2 (213.7)	483.4 (225.1)	.651
Error free utterances*	311.8 (135.5)	302.5 (139.6)	.979
Errors*	242.9 (134.6)	211.9 (135.7)	.521
Spontaneous verbal*	250.0 (161.0)	257.3 (144.5)	.661
Self-Corrections*	49.2 (64.0)	22.0 (22.0)	.040
% errors self-corrected	17.6 (18.1)	12.6 (14.4)	.276
Total words*	1847.4 (1460.7)	1511.6 (1128.2)	.402
MLU (words) [a]	3.5 (2.1)	2.9 (1.4)	.230
MLU (morphemes) [a]	3.9 (2.3)	3.2 (1.5)	.212
Average errors per utterance (errors/utt.)	0.5 (0.3)	0.4 (0.2)	.115
Average utterances per error (utt/errors)	2.4 (1.3)	2.8 (1.3)	.349
Utterance Level Errors	24.3 (27.2)	23.4 (34.0)	.854
Word Level Errors*	232.4 (162.0)	227.4 (162.3)	.997
% total words with errors	33.5 (52.1)	22.5 (25.5)	.326
Utterances per minute	9.7 (3.1)	9.8 (3.6)	.892
Words per minute	35.7 (27.7)	29.5 (19.8)	.355

Note. [a] Mean Length of Utterance, * Variables that were adjusted for time as the session length of the videos was not equal (measure/session time (mins).

Client acts

Descriptive statistics for client acts are presented in Table 8. To maintain context, the means and standard deviations are presented in raw format prior to time adjustment (measure/session time (mins)). After statistical analysis, the groups, Usual Care-Plus, and VERSE were not found to be significantly different with $p > 0.001$.

Discussion

This study sought to establish treatment integrity and treatment differentiation for the VERSE trial in a detailed transcript analysis of the therapy sessions. VERSE was the largest RCT in early aphasia recovery completed to date. The analysis of the wealth of treatment fidelity data collected in the trial is imperative to further analyse and interpret the trial outcomes.

Treatment integrity

Results in this sub-study suggest therapists, providing therapy in the intensive VERSE arm of the trial, were adherent to the prescribed treatment protocol. This was established by measuring the therapists' adherence to the target task level. On average, the target level of verbal output was achieved in 67% of participant utterances in this sample of sessions from the VERSE arm, as per protocol. When looking at a broader session view, the target phrase structure achieved 86% of the time when measured by the mean length of utterance. Protocol adherence for the therapeutic inputs was high at 98% with minimal deviations. Therefore, the key VERSE elements of minimising errors and encouraging verbal output were delivered by the therapist, and it is possible to establish treatment integrity to the dose form within this sample. There is no accepted minimum level of integrity that is required in a complex, behavioural RCT; however, the TF literature suggests protocol adherence of 80% or greater to be considered as a high fidelity level (Borrelli, 2011; Borrelli et al., 2005; Perepletchikova & Kazdin, 2005). Regardless, it is important to report what was intended compared to what was received (Brogan et al., 2019). Additionally, adherence to the intensity of the protocol is more commonly reported than adherence to key therapeutic elements within sessions (Bakheit et al., 2007; Resnick et al., 2011). However, this study has shown that it is possible for aphasia studies to monitor adherence to key therapeutic elements within therapy sessions and report quantitative values for this. Our findings are consistent with other levels of reported protocol adherence in the aphasia literature, in excess of 90% therapist adherence (Behn et al., 2018; Conlon et al., 2020). A current therapy fidelity protocol for the Action Success Knowledge (ASK) trial has set integrity at 100% adherence to essential therapeutic criteria in a session and 80% overall adherence to protocol to deem therapists as meeting criteria (Carragher et al., 2019). It is important to note that instances where sessions deviated from the target goal level may have been a reflection of clinical intuition and responsiveness to the participant and inadvertently caused infidelity to the treatment protocol.

Treatment differentiation

At the broader task level, the VERSE treatment arm contained more verbal output, less single-word tasks, and more conversation-based tasks than the Usual Care-Plus treatment arm as per protocol. However, when the therapeutic inputs and client acts were examined at the utterance level, the treatments did not differ significantly. As expected, the tasks completed in the prescribed VERSE therapy were different from the treatment provided in Usual Care-Plus, however the therapist input, related to the amount of cueing provided to support participant success as well as the amount of language produced by the therapist, was similar for therapists in both arms of the trial. It may be that even if different therapy tasks were used, the manner in which clinicians supported the participants within the treatment sessions was similar. This may reflect clinicians' perspectives around best practice for supporting patients in treatment sessions or their understanding of facilitating neural recovery through limiting errors and providing individual opportunities to produce verbal output. A related explanation concerns the underlying theoretical

mechanism underpinning language therapy. As hypothesized previously, the underlying therapeutic mechanism at work in language therapy may be irrespective of task or therapeutic inputs (Godecke et al., 2016).

In exploring treatment fidelity, it is important to consider the way in which the therapy is delivered as well as the role of the person who receives the therapy, the person with aphasia in this case (Bellg et al., 2004). People with aphasia are a heterogeneous population (Brady et al., 2016), with considerable variation in relation to how aphasia may present. Even within the same aphasia classification, such as Broca's or Wernicke's aphasia, the number of errors made or the types of cues that are helpful may be very different from one person to the next. However, homogeneity can be observed within PWA during therapy. The tasks used between therapy groups and between therapists were different; however, the interaction frequently elicited the same behaviour from people with aphasia. It is possible that people with aphasia were reasonably predictable once their individual error pattern and response to cues was established. Therefore, regardless of the assigned therapy group, the presence or absence of a protocol, the therapist's attempts to manipulate a session, minimise or maximise errors, PWA responded in a particular pattern. And so, another possible explanation for the lack of treatment differentiation at the utterance level in client acts is that regardless of the therapy provided, the response of PWA was similar across the intervention groups. This may relate to the reinforcement of error patterns by PWA regardless of the approach taken by the therapist (Conroy et al., 2009).

Interpretation of main trial results

Therapy fidelity data is vital to support the interpretation of the main trial outcomes, especially in behavioural research (Walton et al., 2017). The main VERSE trial results found that there was not a main effect for treatment intensity (Godecke et al., 2020b). Clinicians within in the VERSE and Usual Care-Plus arms of the trial were required to record 20% of their treatment sessions, in line with recommendations for long-term monitoring (Borrelli, 2011). This study provided a detailed examination of 12% of the videos received through the trial. As this is only a small percentage of the total number of therapy sessions completed, we cannot say with certainty whether this sample is representative of all therapy given, however, we feel the analyses completed and results found provide important insight into the therapy provided within this trial. If the results from this study are extrapolated to the broader VERSE RCT the findings suggest that therapists were adherent to the therapy protocol for the prescribed VERSE arm of the trial and that while the therapies, provided in the VERSE and Usual Care-Plus arms of the trial, contained different tasks and target language levels, how the therapists interacted with participants was not different. Specifically, the number of cues used by therapists to support participants, the success of these cues, the amount of verbal output from the therapist, and the PWA, were not significantly different. We are cautious with extrapolating this result to the entire VERSE RCT sample, however it could provide one potential explanation for why the prescribed VERSE therapy arm and Usual Care-Plus did not perform differently in statistical analyses for the primary outcome measures. It is likely that both conditions share the key elements of aphasia therapy, that is therapists within the Usual Care-Plus condition provided treatment elements at the micro level that share the key elements of the VERSE intervention. This may

relate to the VERSE intervention being a good reflection of existing practice, therefore promoting adherence. Additionally, therapists delivering Usual Care-Plus may already have been employing the key neuroplasticity principles that the VERSE prescribed intervention contained. It is unlikely that this is due to contamination between conditions as the VERSE intervention was shown to be statistically different at the task level of intervention prescription.

Recommendations for future trials

We would encourage researchers to undertake fine-grained analyses of treatments provided within complex, behavioural interventions in the early (pilot) phases of intervention studies. These fine-grained analyses may reveal important underlying factors in the way the intervention is provided or received by participants that require further investigation and understanding in order to interpret trial results. Examining differentiation in the treatments provided may not be sufficient if participant behaviour or acts have the potential to impact clinician behaviour as the therapy session progresses. It will be important to consider this relationship further in complex behavioural trials.

This study established that reviewing aphasia therapy videos from an RCT and coding these for key measures is feasible, however data analyses were lengthy due to the in-depth nature of the analyses completed. Further research is needed to develop a greater understanding of the behaviours that should be investigated in establishing treatment integrity and differentiation and the most efficient ways of evaluating these components. As this study was the first undertaken of this nature, a scoping process was used. If a pilot study is used and the researcher is confident that key intervention ingredients have been identified resulting in a more targeted approach, the practicability may be increased.

Limitations

This study is the first of its kind to provide a detailed analysis of therapy sessions within an aphasia-based RCT. However, it is acknowledged that a small number (12%) of the available trial videos were analysed and so there is a limit to the extrapolation of results to across the trial as a whole. It is also possible that the study was underpowered to find differences between the groups. The primary rater for the videos was not blinded to the treatment condition; however, unblinding was necessary for monitoring protocol adherence, and treatments were also easily distinguishable based on task selection. We do acknowledge the potential for bias.

Conclusion

Aphasia RCTs are encouraged to routinely report therapy fidelity data to assist research consumers with their interpretation of trial results. We report high treatment integrity within the VERSE trial. Determining the active ingredients and dosage level to ensure treatment differentiation is an ongoing goal for aphasia trials. Micro-level analysis of therapy sessions may reveal important underlying factors for further investigation and

only when the treatments are determined to be sufficiently different and the target dosage is achieved, can the efficacy question be answered.

Acknowledgments

The VERSE study was funded by the National Health and Medical Research Council (NHMRC) (#1044973), The Tavistock Trust for Aphasia (UK), and Edith Cowan University, Australia. We would like to thank all participants and their families for taking part in this trial and the therapists who allowed us to video record the therapy sessions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Declarations

The first author discloses receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the inspiring minds post graduate scholarship and Australian postgraduate award from Edith Cowan University. There are no further conflicts of interest.

Author contributions

EB contributed to the concept and design of the work, data collection and analysis, and drafting and editing of the article and approval of the version to be published.

NC and EG contributed to the concept and design of the work, data analysis, and drafting and editing of the article and approval of the version to be published.

ORCID

Emily Brogan http://orcid.org/0000-0001-9604-4558
Natalie Ciccone http://orcid.org/0000-0002-1822-7217
Erin Godecke http://orcid.org/0000-0002-7210-1295

References

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

Bakheit, M., Shaw, S., Barrett, L., Wood, J., Carrington, S., Griffiths, S., Searle, K., & Koutsi, F. (2007). A prospective, randomized, parallel group, controlled study of the effect of intensity of speech and language therapy on early recovery from poststroke aphasia. *Clinical Rehabilitation*, *21*(10), 885–894. https://doi.org/10.1177/0269215507078486

Behn, N., Hilari, K., Marshall, J., Simpson, A., Northcott, S., Thomas, S., Flood, C., Goldsmith, K., & McVicker, S. (2018). SUpporting well-being through PEeR-Befriending (SUPERB) trial: An exploration of fidelity in peer-befriending for people with aphasia. *Aphasiology*, *32*(sup1), 21–23. https://doi.org/10.1080/02687038.2018.1489120

Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., Ogedegbe, G., Orwig, D., Ernst, D., & Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *23*(5), 443–451. https://doi.org/10.1037/0278-6133.23.5.443

Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *Journal of Public Health Dentistry*, *71*(s1), S52–S63. https://doi.org/10.1111/j.1752-7325.2011.00233.x

Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A. J., Czajkowski, S., Breger, R., DeFrancesco, C., Levesque, C., Sharp, D. L., Ogedegbe, G., Resnick, B., & Orwig, D. (2005). A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *Journal of Consulting and Clinical Psychology*, *73*(5), 852–860. https://doi.org/10.1037/0022-006X.73.5.852

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Cochrane Database of Systematic Reviews. *Speech and Language Therapy for Aphasia following Stroke*, (6). Article CD000425. https://doi.org/10.1002/14651858.CD000425.pub4

Brogan, E., Ciccone, N., & Godecke, E. (2019). Treatment fidelity in aphasia randomised controlled trials. *Aphasiology*, *33*(7), 759–779. https://doi.org/10.1080/02687038.2019.1576442

Carragher, M., Ryan, B., Worrall, L., Thomas, S., Rose, M., Simmons-Mackie, N., Khan, A., Hoffmann, T. C., Power, E., Togher, L., & Kneebone, I. (2019). Fidelity protocol for the Action Success Knowledge (ASK) trial: A psychosocial intervention administered by speech and language therapists to prevent depression in people with post-stroke aphasia. *British Medical Journal Open*, *9*(5), e023560. https://doi.org/10.1136/bmjopen-2018-023560

Conlon, E. L., Braun, E. J., Babbitt, E. M., & Cherney, L. R. (2020). Treatment Fidelity Procedures for an Aphasia Intervention Within a Randomized Controlled Trial: Design, Feasibility, and Results. *American Journal Of Speech-Language Pathology*, *29*(1S), 412–424. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0227

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009). Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology*, *23*(11), 1311–1337. https://doi.org/10.1080/02687030902756439

Copland, D. (2017). Do measures of brain activity predict recovery of the ability to communicate after stroke and predict response to speech therapy? (PAPAR). Australian New Zealand Clinical Trials Registry. Retrieved 4th August from https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=372403

Damschroder, L. J., Goodrich, D. E., Kim, H. M., Holleman, R., Gillon, L., Kirsh, S., Richardson, C. R., & Lutes, L. D. (2016). Development and validation of the ASPIRE-VA coaching fidelity checklist (ACFC): A tool to help ensure delivery of high-quality weight management interventions. *Translational Behavioral Medicine*, *6*(3), 369–385. https://doi.org/10.1007/s13142-015-0336-x

DeJong, G., Horn, S. D., Conroy, B., Nichols, D., & Healton, E. B. (2005). Opening the black box of poststroke rehabilitation: Stroke rehabilitation patients, processes, and otcomes. *Archives of Physical Medicine and Rehabilitation*, *86*(12,Suppl.), 1–7. https://doi.org/10.1016/j.apmr.2005.09.003

Gearing, R. E., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. (2011). Major ingredients of fidelity: A review and scientific guide to improving quality of intervention research implementation. *Clinical Psychology Review*, *31*(1), 79–88. https://doi.org/10.1016/j.cpr.2010.09.007

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. (2020a). A randomized control trial of intensive aphasia therapy after acute stroke: The Very Early Rehabilitation for SpEech (VERSE) study. *International Journal of Stroke : Official Journal of the International Stroke Society*, *16*(5), 1747493020961926. https://doi.org/10.1177/1747493020961926

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. (2020b). A randomized control trial of intensive aphasia therapy after acute stroke: The Very Early Rehabilitation for SpEech (VERSE) study. *International Journal of Stroke : Official Journal of the International Stroke Society*, 1747493020961926. https://doi.org/10.1177/1747493020961926

Godecke, E., Armstrong, E. A., Rai, T., Middleton, S., Ciccone, N., Whitworth, A., Rose, M., Holland, A., Ellery, F., Hankey, G. J., Cadilhac, D. A., & Bernhardt, J. (2016). A randomized controlled trial of very early rehabilitation in speech after stroke. *International Journal of Stroke : Official Journal of the International Stroke Society*, *11*(5), 586–592. https://doi.org/10.1177/1747493016641116

Hand, B. N., Darragh, A. R., & Persch, A. C. (2018). Thoroughness and psychometrics of fidelity measures in occupational and physical therapy: A systematic review. *American Journal of Occupational Therapy*, *72*(5), 1–10. https://doi.org/10.5014/ajot.2018.025510

Hebb, D. O. (1949). *The organisation of behaviour. A neuropsychological theory*. John Wiley.

Hickin, J., Cruice, M., & Dipper, L. (2020). A Systematically Conducted Scoping Review of the Evidence and Fidelity of Treatments for Verb Deficits in Aphasia: Verb-in-Isolation Treatments. *American Journal Of Speech-language Pathology / American Speech-Language-Hearing Association*, *29*(1s), 530–559. https://doi.org/10.1044/2019_ajslp-cac48-18-0234

Hildebrand, M. W., Host, H. H., Binder, E. F., Carpenter, B., Freedland, K. E., Morrow-Howell, N., Baum, C. M., Doré, P., & Lenze, E. J. (2012). Measuring treatment fidelity in a rehabilitation intervention study. *American Journal Of Physical Medicine & Rehabilitation*, *91*(8), 715–724. https://doi.org/10.1097/PHM.0b013e31824ad462

Hinckley, J., & Douglas, N. F. (2013). Treatment fidelity: Its importance and reported frequency in aphasia treatment studies. *American Journal Of Speech-Language Pathology*, *22*(2), S279–S284. https://doi.org/10.1044/1058-0360(2012/12-0092)

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. *British Medical Journal*, *348*, g1687. https://doi.org/10.1136/bmj.g1687

IBM Corporation. (2015). In (Version 23.0) IBM Corp. IBM SPSS Statistics for Windows.

Kaderavek, J. N., & Justice, L. M. (2010). Fidelity: An essential component of evidence-based practice in speech-language pathology. *American Journal of Speech-Language Pathology*, *19*(4), 369–379. https://doi.org/10.1044/1058-0360(2010/09-0097)

Karas, S., & Plankis, L. (2016). Consideration of treatment fidelity to improve manual therapy research. *Journal of Manual & Manipulative Therapy*, *24*(4), 233–237. https://doi.org/10.1080/10669817.2015.1135555

Kertesz, A. (2006). *Western Aphasia Battery - Revised* (pp. Pro–Ed). Pearson Assessments.

Koo, T. K., & Li, M. Y. (2016). A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Ludemann, A., Power, E., & Hoffmann, T. C. (2017). Investigating the adequacy of intervention descriptions in recent speech-language pathology literature: Is evidence from randomized trials useable? [Article]. *American Journal Of Speech-Language Pathology*, *26*(2), 443–455. https://doi.org/10.1044/2016_AJSLP-16-0035

McLennon, S. M., Hancock, R. D., Redelman, K., Scarton, L. J., Riley, E., Sweeney, B., Habermann, B., Jessup, N. M., & Bakas, T. (2016). Comparing treatment fidelity between study arms of a randomized controlled clinical trial for stroke family caregivers. *Clinical Rehabilitation*, *30*(5), 495–507. https://doi.org/10.1177/0269215515585134

Miller, J. (2008). Systematic Analysis of Language Transcripts (SALT), English Demo Version 2008 [Computer Software]. In. Wisconsin: SALT Software LLC.

Moher, D. (2018). Reporting guidelines: Doing better for readers [journal article]. *BMC Medicine*, *16*(1), 233. https://doi.org/10.1186/s12916-018-1226-0

Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, *11*(3), 247–266. https://doi.org/10.1016/0272-7358(91)90103-2

O'Donell, C. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33–84. https://doi.org/10.3102/0034654307313793

O'Shea, O., McCormack, R., Bradley, J., & O'Neill, B. (2016). Fidelity review: A scoping review of the methods used to evaluate treatment fidelity in behavioural change interventions. *Physical Therapy Reviews*, *21*(3–6), 207–214. https://doi.org/10.1080/10833196.2016.1261237

Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *British Medical Journal*, *332*(7538), 413–416. https://doi.org/10.1136/bmj.332.7538.413

Öst, L.-G. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy*, *46*(3), 296–321. https://doi.org/10.1016/j.brat.2007.12.005

Page M, Wilson BA, Shiel A, Carter G, Norris D. (2006). What is the locus of the errorless-learning advantage? *Neuropsychologia,44*(1), 90–100. doi:10.1016/j.neuropsychologia.2005.04.004.

Palmer, R., Cooper, C., Enderby, P., Brady, M., Julious, S., Bowen, A., & Latimer, N. (2015). Clinical and cost effectiveness of computer treatment for aphasia post stroke (Big CACTUS): Study protocol for a randomised controlled trial. *Trials*, *16*(1), 18. https://doi.org/10.1186/s13063-014-0527-7

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*, *12*(4), 365–383. https://doi.org/10.1093/clipsy.bpi045

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

[Record #255 is using a reference type undefined in this output style.]

Resnick, B., Bellg, A. J., Borrelli, B., De Francesco, C., Breger, R., Hecht, J., Sharp, D. L., Levesque, C., Orwig, D., Ernst, D., Ogedegbe, G., & Czajkowski, S. (2005). Examples of implementation and evaluation of treatment fidelity in the BCC studies: Where we are and where we need to go [journal article]. *Annals of Behavioral Medicine*, *29*(2), 46. https://doi.org/10.1207/s15324796abm2902s_8

Resnick, B., Michael, K., Shaughnessy, M., Nahm, E. S., Sorkin, J. D., & Macko, R. (2011). Exercise intervention research in stroke: Optimizing outcomes through treatment fidelity. *Topics in Stroke Rehabilitation*, *18*(1), 611–619. https://doi.org/10.1310/tsr18s01-611

Rixon, L., Baron, J., McGale, N., Lorencatto, F., Francis, J., & Davies, A. (2016). Methods used to address fidelity of receipt in health intervention research: A citation analysis and systematic review. *BMC Health Services Research*, *16*, 663. https://doi.org/10.1186/s12913-016-1904-6

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 : Official Journal of the International Stroke Society*, *14*(9), 972–976. https://doi.org/10.1177/1747493019870401

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 : Official Journal of the International Stroke Society*, *14*(9), 972–976. https://doi.org/10.1177/1747493019870401

Schlosser, R. (2002). On the importance of being earnest about treatment integrity. *Augmentative and Alternative Communication*, *18*(1), 36–44. https://doi.org/10.1080/aac.18.1.36.44

Seng, E. K., & Lovejoy, T. I. (2013). Reliability and validity of a treatment fidelity assessment for motivational interviewing targeting sexual risk behaviors in people living with HIV/AIDS [journal article]. *Journal of Clinical Psychology in Medical Settings*, *20*(4), 440–448. https://doi.org/10.1007/s10880-012-9343-y

Spell, L. A., Richardson, J. D., Basilakos, A., Stark, B. C., Teklehaimanot, A., Hillis, A. E., & Fridriksson, J. (2020). Developing, Implementing, and Improving Assessment and Treatment Fidelity in Clinical Aphasia Research. *American Journal Of Speech-Language Pathology*, *29*(1), 286–298. https://doi.org/10.1044/2019_AJSLP-19-00126

Thomas, S. A., Coates, E., Das Nair, R., Lincoln, N. B., Cooper, C., Palmer, R., Walters, S. J., Latimer, N. R., England, T. J., Mandefield, L., Chater, T., Callaghan, P., & Drummond, A. E. R. (2016). Behavioural Activation Therapy for Depression after Stroke (BEADS): A study protocol for a feasibility randomised controlled pilot trial of a psychological intervention for post-stroke depression [journal article]. *Pilot and Feasibility Studies*, *2*(1), 45. https://doi.org/10.1186/s40814-016-0072-0

Turkstra, L., 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 [Article]. *American Journal Of Speech-Language Pathology*, *25*(2), 164–171. https://doi.org/10.1044/2015_AJSLP-15-0060

Varley, R. (2011). Rethinking aphasia therapy: A neuroscience perspective. *International Journal of Speech-language Pathology*, *13*(1), 11–20. https://doi.org/10.3109/17549507.2010.497561

Walker, M. F., Hoffmann, T. C., Brady, M. C., Dean, C. M., Eng, J. J., Farrin, A. J., Felix, C., Forster, A., Langhorne, P., Lynch, E. A., Radford, K. A., Sunnerhagen, K. S., & Watkins, C. L. (2017). Improving the development, monitoring and reporting of stroke rehabilitation research: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. *International Journal of Stroke : Official Journal of the International Stroke Society*, *12*(5), 472–479. https://doi.org/10.1177/1747493017711815

Walton, H., Spector, A., Tombor, I., & Michie, S. (2017). Measures of fidelity of delivery of, and engagement with, complex, face-to-face health behaviour change interventions: A systematic review of measure quality. *British Journal of Health Psychology*, *22*(4), 872–903. https://doi.org/10.1111/bjhp.12260

Warren, S., Fey, M., & Yoder, P. (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

Wheeler, J. J., Baggett, B. A., Fox, J., & Blevins, L. (2006). Treatment integrity: A review of intervention studies conducted with children with autism. *Focus on Autism and Other Developmental Disabilities*, *21*(1), 45–54. https://doi.org/10.1177/10883576060210010601

Worrall, L., Ryan, B., Hudson, K., Kneebone, I., Simmons-Mackie, N., Khan, A., Hoffmann, T., Power, E., Togher, L., & Rose, M. (2016). Reducing the psychosocial impact of aphasia on mood and quality of life in people with aphasia and the impact of caregiving in family members through the Aphasia Action Success Knowledge (Aphasia ASK) program: Study protocol for a randomized controlled trial. *Trials*, *17*, 153. https://doi.org/10.1186/s13063-016-1257-9