Speech and Language Sciences Section

School of Education, Communication and Language Sciences

Determining the Efficacy of a Usage-Based Language Intervention in the Early Years: A Non-Randomised Pilot Study

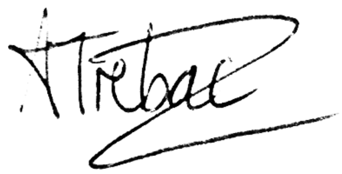
Anastasia Grace Sophia Trębacz

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

October 2020

Declaration of Originality

The material presented in this thesis is the original work of the candidate except as otherwise acknowledged. It has not been submitted previously in part or in whole, for any award at any university, at any other time.



Anastasia Grace Sophia Trębacz

Acknowledgements

Firstly, enormous thanks are due to my supervisory team Professor Cristina McKean, Dr Helen Stringer, Professor James Law and Dr Lauren Ackerman for conceiving the vision of this work and equipping me with the tools and confidence to make it a reality. I could not have asked for more supportive and empowering supervisors.

Thank you to Professor Nicola Botting and Professor Carolyn Letts for their detailed reading of this thesis and for all of their thoughtful and constructive feedback.

A number of other individuals warrant special thanks:

- The Research Excellence Academy at Newcastle University for the generous funding of this project
- Carrie Hughes and her team of Speech and Language Therapists, whose tireless support ensured schools and teachers were enthusiastic and willing to take part
- All of the schools, teachers, children, parents and carers who so generously took part in and supported this study
- Dr Sean Pert for his support throughout the project
- Professor Carolyn Letts and Professor Julie Morris for their yearly panel input which saw this thesis develop immeasurably
- Professor Ghada Khattab, Dr Nick Riches, Dr Carol Moxam, Dr Christos Salis, Dr Jalal-eddin Al-Tamimi, Dr Christine Jack, Helen Nazlie, Dr Faye Smith, Dr Laurence White, Dr Elaine Lopez. I feel so privileged to have worked with such an incredible and lovely department of people who made me feel welcome from the beginning
- To the Psychology Placement Students Phyllis Leung and Laura Barlow, and the Undergraduate SLT students Sarah Atkinson, Iona Kapstick, Louise Bulman, Janey Eastwood, and Bryony Mann, for supporting the blinded assessment delivery. These SLT students were also able to use data from this project in their own Undergraduate dissertations, an experience which I hope was valuable to them.
- My fabulous friends and colleagues in Speech and Language Sciences and beyond… There are too many to mention, but special thanks to Joe Willoughby, Jojo Tulip, Penny Levickis, Helen Wareham, Kerry Bray and Scott Kirkman who ensured I always had somebody to laugh with, to share ideas with, and that I was never doing this alone.
- To the amazing friends I have met along this journey, I really couldn't have done it without you.
- To my parents Miriam and Stefan, my Grandmother Anita, and Kate and Thomas, who were always only a phone call away.
- To Theodora for her love, acceptance and artwork.
- To Oma and Dziadziu, who are in everything I do.
- To my Uncle and friend, Tim Clark, who taught me that miracles do happen.

This thesis is dedicated to the memory of Alan Clark and Maddie Daisy Eatwell; both inspired me throughout their lives and continue to today.

Abstract

Introduction Language skills are crucial for positive social, emotional, academic and economic outcomes across the lifespan. Whilst most children acquire language relatively easily, a number of children experience difficulties. Early intervention can mitigate the risks associated with poor language skills, and the importance of evidence-based intervention is widely recognised by Speech and Language Therapists. Despite this, there are few robustly evaluated language interventions for young children.

Building Early Sentences Therapy (BEST) is a complex language intervention designed to support children to understand and produce two, three and four clause-element sentences (McKean et al., 2013).

This thesis presents a non-randomised quasi-experimental study of BEST with sign, BEST without sign and Treatment as Usual with four key goals: 1) to evaluate the efficacy of BEST, 2) to assess the use of sign as an 'active ingredient', 3) determine the effect sizes from goals 1 and 2, and 4) to inform further evaluation.

Methods Schools were assigned to three treatment arms matched with respect to classroom oral language environment and socioeconomic status. Participants were 3;5-4;5-year-old children with expressive, receptive or mixed language difficulties below the 16^{th} centile.

The intervention was 16 sessions of BEST with or BEST without sign. The TAU group received their usual classroom provision.

Children's outcomes on production, comprehension, content and morphology on targeted and non-targeted language structures were measured.

Results Visual trends showed improvement for children receiving BEST.

There was evidence for the efficacy of BEST compared to TAU for Production, Content and Morphology outcomes.

There was also evidence that BEST with Sign improved outcomes further for Production and Sentence Morphology outcomes.

Content Analysis of teacher feedback led to the generation of recommendations for future studies evaluating BEST.

Discussion Children receiving BEST make progress with production, content and morphology outcomes. The role of sign in further supporting progress is positively indicated. A further trial of BEST is warranted.

Dissemination

Presentations arising from this thesis

- McKean C, Stringer H, Letts C, Pert S, Masidlover M, Jack C, **Trebacz A**, Wareham H, Preston E, Baker J, Benson K. Staying faithful: Exploring the challenges of treatment fidelity. *In: North East SLT Research Network: Symposium*. 2020, North East SLT Research Network.
- Letts C, Pert S, Czaplewska E, Ashton E, Benson K, Preston E, Stringer H, **Trebacz A**, Wareham H, McKean C, Jack C. Working with bilingual children at risk of developmental language disorder: adapting assessment and intervention for the nursery-aged bilingual child. *In: NALDIC 28 Conference 2020: Multilingual Britain: Successes, challenges and future directions*. 2020, Virtual: National Association for Language Development in the Curriculum.
- **Trebacz A**, McKean C, Stringer H, Law J. A protocol for the refinement and evaluation of language intervention for children with language difficulties in the early years. *In: Child Language Symposium*. 2019, Sheffield, UK.
- **Trebacz A**, McKean C, Stringer H, Law J. **Evaluating Building Early Sentences Therapy (BEST); a Novel Speech & Language Intervention**. *In: Great North Children's Research Community Conference*. 2019.

Publications and Pre-registrations arising from this thesis

- McKean C, Benson K, Jack C, Letts CA, Pert S, Preston E, **Trebacz A**, Stringer H, Wareham H. Language intervention in the early years - comparing the effectiveness of language intervention approaches for pre-school children with language difficulties. BioMed Central Ltd, 2020. Available at: https://doi.org/10.1186/ISRCTN10974028.

Abbreviations

ANOVA: Analysis of Variance
ANCOVA: Analysis of Covariance
ASD: Autism Spectrum Disorders
BCRP: Better Communication Research Programme
BCT: Behaviour Change Theory
BCW: Behaviour Change Wheel
BEST: Building Early Sentences Therapy
BSL: British Sign Language
CGC: Computational Grammatical Complexity (account)
COM-B Model: Capability, Opportunity Motivation-Behaviour Model
CONSORT: Consolidated Standards of Reporting Trials
CPD: Continuing Professional Development
CSCOT: Communication Supporting Classrooms Observation Tool
DLD: Developmental Language Disorder
EAL: English as an Additional Language
EEF: Education Endowment Foundation
EYFS: Early Years Foundation Stage
EIF: Early Intervention Foundation
EOI: Extended Optional Infinitives (account)
EUCC: Extended unique checking constraint (account)
FOCUS: Focus on the Outcomes of Communication Under Six
GDPR: General Data Protection Regulations
G-SLI: Grammatical SLI
HT: Head Teacher
IDACI: Income Deprivation Affecting Children Index
IMD: Index of Multiple Deprivation
IRAS: Integrated Research Application System
LA: Local Authority
LAD: Language Acquisition Device
LI: Language Impairment
LD: Learning Disability
MLM: Multilevel Modelling
MRC: Medical Research Council
NCI: National Cancer Institute
NELI: Nuffield Early Language Intervention
NHS: National Health Service
NIHR: National Institute for Health Research
NRDLS: New Reynell Developmental Language Scales
OFSTED: Office for Standards in Education, Children's Services and Skills
OL: Oral Language (programme)
OLS: Ordinary Least Squares
PAS: Predicate Argument Structure
PCIT: Parent-Child Interaction Therapy
PDHL: Procedural Deficit Hypothesis
PGSS Paget Gorman Signed Speech
PLI: Primary Language Impairment
RA: Research Assistant
RCSLT: Royal College of Speech and Language Therapists
RCT: Randomised Controlled Trial
RDDR: Representational Deficit for Dependent Relations (account)
REC: Research Ethics Committee
RT: Response Times

RTI: Response To Intervention
SENCO: Special Educational Needs Coordinator
SEBD: Social, Emotional and Behavioural Difficulties
SES: Socio-Economic Status
SDQ: Strengths and Difficulties Questionnaire
SLCN: Speech Language and Communication Needs
SMD: Standardised Mean Difference
SLT: Speech and Language Therapist/Therapy
T1, T2, T3: Time point 1, Time point 2, Time point 3
TA: Teaching Assistant
TAU: Treatment as Usual
TD: Typically Developing
TDF: Theoretical Domains Framework
TFA: Theoretical Framework of Acceptability
WHO: World Health Organisation
WM: Working Memory

Chapter 1

Theoretical and Clinical Motivation for Evaluating Building Early Sentences Therapy (BEST); a Novel Usage-Based Language Intervention

1.1 Introduction

The majority of children acquire their first language with relative ease, however a conservative estimate suggests that 9.9% of children have a language difficulty severe enough to hinder their academic progress (Norbury et al., 2016). Language plays an essential role in positive social-emotional, mental health, academic and employment outcomes into adulthood (Conti-Ramsden et al., 2018; Conti-Ramsden et al., 2009; Durkin et al., 2009; Law et al., 2009; Snowling et al., 2006) and promoting positive language outcomes is therefore a priority.

Language difficulties are highly unstable in the early years (Reilly et al., 2014) and if they do not resolve by school entry, they are likely to persist until at least age 11 (McKean et al., 2017). It is therefore crucial to provide effective early intervention in the pre-school years to support children with language difficulties (Law et al., 2017). Despite the high level of need (Bercow, 2008, 2018), there are few robustly evaluated interventions for children in this age range (Law et al., 2017; Law et al., 2012). Fewer still have clearly articulated, explicit theoretical motivation (Roulstone et al., 2012).

1.1.1 Long-Term Consequences of Language Difficulties

Children who enter school with poor language are more likely to suffer adverse outcomes in their educational, social, emotional, and behavioural development (Conti-Ramsden & Botting, 2004, 2008; Snowling et al., 2001; Tomblin, 2008). The consequences of poor language skills during childhood include difficulties in the making and maintenance of peer relationships (Botting & Conti-Ramsden, 2008; Conti-Ramsden et al., 2013), social, emotional and educational difficulties (Conti-Ramsden & Botting, 2004, 2008; Conti-Ramsden et al., 2013; Snowling et al., 2001; Tomblin, 2008), and long-term negative effects on social inclusion, mental health, employability and understanding one's legal rights (Conti-Ramsden et al., 2013; Law et al., 2009; Rost & McGregor, 2012; Schoon et al., 2010).

This thesis presents a pilot study examining the application of Building Early Sentences Therapy (BEST), a novel, usage-based language intervention for pre-school children with severe language difficulties. BEST manipulates features of the language input and learning contexts to harness a number of cognitive mechanisms hypothesised to promote language learning. The research addresses four goals: 1) to evaluate whether BEST is effective in improving children's language development; 2) to examine whether the signing component of BEST is a key 'active ingredient' of the intervention; 3) to determine the effect sizes of findings from 1) and 2) above; and 4) to inform further stages in the intervention evaluation process through consideration of acceptability and feasibility of study processes.

The following sections of this chapter provide the background and rationale for this thesis.

The approach to terminology and labelling of language difficulties in the early years is first presented to contextualise the research within the target populations. The prevalence and issues surrounding identification of difficulties in pre-school children are then described. Next, the current evidence base of pre-school language interventions is reviewed, identifying the gap addressed by the BEST intervention. A review of current theories of the underpinning mechanisms of language difficulties and Developmental Language Disorder (DLD) is then presented to motivate the application of usage-based theory to intervention design.

The potential of BEST to address the issues raised in the literature review is considered. This is followed by a review of the use of sign in language interventions and a description of the hypothesised benefits of sign within the BEST intervention. Current frameworks for the evaluation of complex interventions are then discussed, to situate the present study within the appropriate phase of the evaluation process. Finally, the research questions addressed in this thesis are presented.

1.1.2 Terminology used throughout this Thesis

1.1.2.1 Diagnostic Terms

Developmental Language Disorder (DLD) is the recognised diagnostic label for a persistent language disorder that affects a child's functioning in the absence of an underlying biomedical condition. The term DLD was defined by the Delphi consensus method CATALISE project (Bishop et al., 2017). If the child does have an underlying biomedical conditions such as a neurological injury, developmental disorder or genetic condition, they would instead likely be diagnosed with 'Language Disorder Associated with [Condition X]' (Bishop, 2017; Bishop et al., 2016).

Prior to consensus adoption of the term DLD, the principal diagnosis for language difficulties with an absence of any frank hearing, neurological or emotional impairments was Specific Language Impairment (SLI) (Leonard, 2014), sometimes also referred to as Primary Language Impairment (PLI). Grammatical Specific Language Impairment (G-SLI) is a sub-type of SLI in which children present with grammatical deficits, or deficits affecting certain aspects of grammar (Friedmann & Novogrodsky, 2004; van der Lely, 1998, 2005). Some questioned the validity of the G-SLI subgroup (Leonard, 2014). SLI was omitted from the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013), after its validity was questioned by numerous researchers (Reilly et al., 2014; Tomblin & Records, 1996).

Throughout the literature review, the term DLD is used. This chapter draws on the older SLI literature, but children with SLI fulfil the criteria for a diagnosis of DLD and as such the focus of this review is on children who meet the criteria for DLD. Where the fact the participants meet strict SLI criteria has a specific bearing on the interpretation of the research, this is indicated.

1.1.2.2 *Language Difficulties*

Early language difficulties must persist for children to receive a diagnosis of DLD. Consequently, diagnoses are not usually given until at least four years, when practitioners can be more confident of the long-term nature of any difficulties (McKean et al., 2017). Within the CATALISE framework it is less clear how pre-school children should be labelled (Bishop et al., 2017). The target of this study are pre-school children with language skills falling below expected levels for their age, however such children would be unlikely to have a formal diagnosis. Children in this study are therefore described as having language difficulties and being at risk of DLD.

1.1.2.3 *The Early Years*

The definition of pre-school varies by country. In this thesis the terms ‘the early years’ and ‘the pre-school years’ are both used to define children who have not yet started school. In the United Kingdom, where the present research was conducted, children start Reception in the September after their fourth birthday and begin formal education in the September after their fifth birthday. While BEST is recommended for use with children between 3;0-6;0 years, the children included in the present study are a subsection aged between 3;5-4;5 years old and are described as being in the early years.

1.2 Language Difficulties in the Early Years

Robust early language skills are of crucial importance for later development and are associated with positive long-term social-emotional, mental health, academic and employment outcomes (Bryan et al., 2007; Cohen et al., 2013; Durkin et al., 2009; Hollo et al., 2014; Law et al., 2009, 2017; Snowling et al., 2006). If children’s language difficulties are not resolved by school entry, they are likely to persist until at least age 11 (McKean et al., 2017), therefore it is important to identify these children early and to provide timely educational support and intervention.

To illustrate the scale of language difficulties in the early years at a population level, several factors must be considered. These are prevalence, lack of stability, identification and risk factors, each of which are discussed in detail below.

1.2.1 Prevalence of Language Difficulties

Prevalence indicates the number of cases of language difficulties in a given population. Multiple prevalence estimates of language difficulties have been calculated for school age and pre-school children. A recent British epidemiological study found approximately 9.9% of four-to-five-year-old children presented with a language difficulty severe enough to negatively affect academic performance, which is roughly equivalent to two children in a classroom of 25. Approximately 7.6% of the identified children had no diagnosed co-morbidity which otherwise explained their language difficulties (Norbury et al., 2016). This estimate was obtained in an area of relative social advantage. Given that language difficulties follow a social gradient, with prevalence estimates as high as 50% in the most socially disadvantaged areas (Locke et al., 2002), Norbury et al. (2016) suggest that their estimate may underestimate the true population level prevalence of language difficulties for four-to-five-year-old children. Whilst it must therefore be interpreted with caution, Norbury et al's estimate aligns with previous studies. For example, Tomblin et al. (1997) calculated the first robust estimate of language impairment in five-to-six-year-old children of 7.4%, and a subsequent meta-analysis by Law et al. (2000) estimated prevalence in five-to-seven-year-olds between 2.02%-10.4%.

The target population for the present study is 3;5-4;5-year-olds, so prevalence estimates for this age group are of particular importance. As with school-age children, in the pre-school years there is also a high level of variation in prevalence estimates, both by the age of the sample and by the specific methodology of a given study. In a recent review of studies estimating prevalence for pre-school children, Law et al. (2017) reported that the median prevalence estimate for three year olds was 7%, while the median estimate for four-to-five year old children was 10.7%. Overall median estimates for 18 month-to-four-to-five-year-olds ranged between 7-14%, suggesting that the target population for the present study fell between the low-to-mid point of these estimates. This may be due to the high probability of resolution for 'late talkers' at around 24 months, discussed in more detail in 1.2.2 (Reilly et al., 2010). Clearly caution should be exercised when generalising prevalence estimates to specific populations due to the variation reported in the literature, however 7%-10% appears to be a reasonable estimate for the target age group.

1.2.2 Lack of Stability in the Early Years

The early years is a period of relative instability for children's language difficulties, which is reflected in the variance of prevalence estimates discussed above. An awareness of the lack of stability in the early years is important when researching interventions for this age group, since spontaneous within-child variation can occur, unrelated to the intervention being delivered. There are a wide range of individual differences in language abilities, with much fluidity and natural change occurring over the early years (Dollaghan & Campbell, 2009; Frans et al., 2017). Longitudinal studies have shown that some children who present with early language difficulties catch up to their peers and that the percentage of children with resolving difficulties reduces as the age of initial identification increases (Ghassabian et al., 2014; Henrichs et al., 2011; Law et al., 2012; Reilly et al., 2010; Zambrana et al., 2014; Zubrick et al., 2015).

Findings from the Australian cohort Early Language in Victoria Study (ELVS) suggest around 70% of children with language difficulties at age two, often described as 'late talkers', catch up to their typically developing peers by age four (Reilly et al., 2014). Conversely, there is another group of approximately 8% of children, who present as typical talkers at age two, but by age four show signs of language difficulty.

There is evidence that stability increases in school-age children. Children with receptive vocabulary difficulties at age seven are unlikely to resolve by age 11 (Law et al., 2008). McKean et al. (2017) demonstrate that children who reach age four without recovering from language difficulties are highly likely to experience persisting difficulties until at least 11 years of age, while Beitchman et al.

(2008) find a similar pattern between ages five and 25. Conti-Ramsden et al. (2001) followed-up a cohort of children who attended a language unit at age seven when they were 11 years old. Only 8% of children no longer showed evidence of language difficulties on any measures, while the remainder of children continued either to meet the criteria for Specific Language Impairment (SLI), or had language difficulties but had received another diagnosis, thus not being eligible for a diagnosis of SLI. Tomblin et al. (2003) also assessed children diagnosed with Primary Language Impairment (PLI) at age five-to-six and again at age nine-to-ten. There was little evidence of improvement in this time period, and language change was not statistically significant from baseline language measurements. Given the trajectories of children's language difficulties it is important that we accurately identify the children most in need of support.

1.2.3 Reliable Identification of Children with Language Difficulties in the Early Years

Reliable identification of children with language difficulties depends on accurate measurement techniques. Accurate diagnoses are essential for treating the correct children and for the effective utilisation of finite resources.

Strict cut-points in standardised tests are frequently used to identify children with language difficulties, although this method can be highly problematic. Measurement error close to thresholds poses a risk of mis-categorisation, with consequences for identification in clinical and research contexts (also see Le et al., 2020; McKean et al., 2017). Measurement error can occur when assessing young children due to a wide range of factors including environmental distraction, cognitive demands and the child's attention and behaviour, leading to unreliable identification (Burchinal, 2008).

Predicting the persisting difficulties necessary for a DLD diagnosis must also take into account contextual factors relating to the child. Identification of the child's language profile (see 1.2.3.1) and multilingualism (see 1.2.3.2) can be helpful in predicting expected long-term outcomes, although there is no infallible way of doing this.

1.2.3.1 Language Profiles

Language difficulties can be expressive, affecting a child's ability to use words and sentences; receptive, affecting a child's ability to understand what others say to them; or mixed, affecting both expressive and receptive skills. Children with expressive language profiles display greater progress with intervention than those with mixed expressive-receptive language difficulties (Boyle et al., 2010), and children with mixed language profiles and receptive-only profiles are at risk of poorer long term outcomes than expressive-only profiles (e.g. Law et al., 2004; Nickisch & von Kries, 2009). This evidence suggests that there is value in establishing a child's language profile early as it may provide an indication of the severity and longevity of difficulties and likelihood of spontaneous improvement.

1.2.3.2 Multilingualism

In a UK context, children who speak English as an additional language can appear to be behind their monolingual peers in the early years, likely due to reduced exposure, whilst not suffering from any long-lasting difficulties. Of course, it is also possible for multilingual children to have DLD, highlighting the need for assessment and intervention across all of the child's languages (Stow & Pert, 2015). Practitioners must also consider the degree of exposure to various languages when making judgements about DLD in multilingual children, since most will catch up to their monolingual peers without any need for SLT intervention (Stow & Pert, 2015).

McKean et al. (2015) demonstrated that while children who spoke English as an additional language (EAL) showed the largest English language disadvantage at age 4, they made faster gains than monolingual children before age 11.

Bilingualism is an important consideration for reliable identification of language difficulties and a multilingual child should be viewed in the context of all languages spoken when being assessed.

1.2.4 Risk Factors for Developmental Language Disorder

A number of environmental, genetic and cognitive factors are associated with increased risk of language difficulties (Conti-Ramsden & Durkin, 2016; Prathanee et al., 2007). It is important that teachers and practitioners are aware of these factors so that they can effectively monitor and support at-risk children.

Relative socio-economic deprivation and lower levels of maternal education are associated with lower language scores in the Early Years (Letts et al., 2013). Children of more highly educated parents are exposed to more words than children of parents with less education (Fernald et al., 2013; Hart & Risley, 1995), although the extent of the 'word-gap' is now debated (see Sperry et al. 2019).

Evidence that language difficulties have a genetic component comes from twin studies (Bishop, 2002) and sibling studies (Choudhury & Benasich, 2003). This research demonstrates a familial component to language difficulties, with an aggregation of 32% in relatives of people with a language disorder, higher than for a control group with no diagnosed impairment (Choudhury & Benasich, 2003).

Biological sex is associated with language difficulties, with a higher prevalence amongst boys (Choudhury & Benasich, 2003; Leonard, 2014) and girls displaying better communication skills and larger vocabularies at 24 months (Reilly et al., 2009) (although see Tomblin et al., 1996 and Yew & O'Kearney, 2013 for alternative explanations).

Cognitive difficulties including impaired short-term memory capacity (Gathercole & Baddeley, 1989, 1990), procedural learning (Evans et al., 2009) and procedural memory (Ullman & Pierpont, 2005) are also implicated in language difficulties. Cognitive factors are closely linked with genetic factors, representing the phenotypic expression of the language disorder genotype and therefore driving the behavioural and cognitive manifestation of language difficulties (Bishop, 2006).

1.2.5 Conclusion

There is much instability in language ability in the early years, however 7%-10% of 3;5-4;5-year-olds experience language difficulties. Early identification of language difficulties in the pre-school years enables the timely provision of early intervention to support children and mitigate potential poor long-term outcomes. Environmental, genetic and cognitive factors can also increase the risk of developing language difficulties.

Knowing when to intervene for children in the early years is a challenge and must also account for risk factors. From the review above, the case can be made that between 3;5 and 4;5 is an optimal time to intervene for children with language difficulties as this age avoids the high rate of spontaneous resolution seen at around 24 months, but is early enough that language trajectories are not necessarily fixed. This age group is therefore targeted in the present study.

1.3 What Interventions are Currently Available to Support Children with Language Difficulties?

In this section our attention turns to existing methods of addressing language difficulties in the pre-school years. As is discussed above, early intervention may help to mitigate difficulties in the early years. Children between three-and-six-years-old could particularly benefit from early intervention since language difficulties are more easily identified and trajectories are less 'fixed' at this age (McKean et al., 2015). Such interventions should be evidence-based (Klee et al., 2009) and underpinned by explicit theories of change (Law et al., 2017).

Here we focus on interventions delivered in the early years by educators or SLTs (Boyle et al., 2007a). There is a large body of evidence relating to Parent-Child Interaction Therapy (PCIT) in which the focus of the intervention is the interaction between the parent/carer and the child. There are many issues with equity of access to such interventions particularly with respect to socially disadvantaged families (Gibbard & Smith, 2016; Klatte et al., 2019; Klatte & Roulstone, 2016). Interventions delivered by SLTs or early educators in pre-school settings are therefore essential to ensure all children have access to necessary support and are the focus of this study.

This section reviews the evidence to determine whether effective interventions are currently available to treat severe language difficulties in pre-school children. It also aims to determine the level of available evidence for such interventions and identify which theories of language acquisition, if any, underpin them. This review does not consider interventions that are incomparable to the present study due to targeting different areas of language e.g. phonology or following highly different approaches. Interventions targeting children with diagnosed disorders other than DLD, or children older than six years are also not included.

1.3.1 The What Works Database

Several meta-analyses and systematic reviews have identified existing evaluated and peer reviewed interventions designed to treat language difficulties in the target population (Cable & Domsch, 2011; Gladfelter et al., 2011; Law et al., 2017; Law et al., 2004; Law et al., 2012). The What Works database compiled by Law et al. (2012) is a large, interactive, open-access database of information about evaluated interventions that support children's speech, language and communication development. As well as collating published evidence, the database integrates external practitioner views and experiences of delivering the interventions (Law et al., 2015). The database is updated annually by an advisory group who invite contributions to build upon this. What Works is therefore the most inclusive, comprehensive and up to date collection of evidence for language interventions, which is also freely available to practitioners and researchers. For these reasons it was used to identify interventions for this review.

The database provides a number of intervention characteristics including: 1) who delivers the intervention (specialist e.g. Speech and Language Therapist or Specialist Teacher, Teacher, Teaching or Educational/SLT Assistant or Other; 2) target age range (preschool, primary or secondary age children); 3) the area of need (speech, language or communication); 4) the format of the intervention (manualised intervention, approach or technique); 5) the focus of the intervention (universal, targeted or specialist) (described below); 6) assigned level of evidence (described below); and 7) effect size.

1.3.1.1 Response to Intervention (RTI)

The Response to Intervention (RTI) model (Fuchs et al., 2003) details three tiers of service delivery. Universal interventions are delivered to all children and often form part of typical classroom practice. Targeted interventions are delivered to sub-groups of children thought to be at risk, or in need of additional support, often by educational staff such as Teaching Assistants (TAs) as well as SLTs. Finally, Specialist intervention is delivered to children with persistent difficulties, usually by specialist practitioners such as SLTs or SENCOs (see Fuchs et al. (2003) for further discussion of the RTI model).

1.3.1.2 Level of the Evidence

Interventions in the database are assigned a level of evidence using a three-point rating system. Interventions are considered to have strong evidence if they are supported by at least one positive systematic review as well as individual trials. Moderate evidence consists of either a single or multiple randomised controlled trials (RCTs) or quasi-experimental studies. Indicative evidence includes multiple baselines, case-series and pre-and-post-test designs where statistically significant differences between treatment and control have been identified. Millard et al. (2009) provide detailed examples of each of the levels of evidence. Cohen's d effect sizes that describe the magnitude of impact are included in the database where possible (see Cohen, 1988).

1.3.1.3 Intervention Intensity

Although it is not quantified in the What Works database, this review also seeks to identify the quantitative measures of intervention intensity comprising: 1) dose; 2) dose frequency; and 3) total intervention duration (Warren et al., 2007). Dose is defined as the number of teaching episodes within a session, duration of intervention session and the distribution of teaching episodes over the session. Dose frequency refers to the number of sessions delivered over a given period of time and total intervention duration is the total time period over which the intervention is delivered. Few studies have examined its effects on language outcomes, however intervention intensity has implications for service delivery and funding and is an important component of intervention evaluation (Frizelle et al., in press). Further, limited findings suggest that higher doses may not always promote better outcomes (Schmitt et al., 2017). Indeed, a recent systematic review identified optimal delivery as either very short frequent sessions or longer, less frequent sessions for morphosyntactic and vocabulary outcomes (Frizelle et al., in press). No additional benefits were found for long, frequent sessions. For these reasons, intervention intensity is examined in the identified studies where possible.

1.3.1.4 Theoretical Underpinnings

As with intervention intensity discussed above, the What Works database does not explicitly report the theoretical underpinnings of the interventions it describes. Interventions should be motivated by theories of typical language acquisition and language disorder (Law et al., 2017), so that 'active ingredients' can be identified and tested. This also presents the opportunity to evaluate current theories on the basis of whether they are supported or contradicted by the evidence. This review therefore examines the theoretical underpinnings of identified interventions where available.

1.3.2 Searching the Database

The What Works database was searched to identify the current state of the evidence. The search aimed to identify evaluated targeted and specialist language interventions for children with language and communication difficulties in the pre-school years (i.e. below 5 years of age). These interventions could be delivered by specialists, teachers or assistants and could be in a manual, approach or technique format. Interventions with all levels of evidence ratings were included (see Appendix 1 for the full search criteria).

Initially thirty-eight interventions met the search criteria (see Appendix 2). The relevance of these interventions to the present study was considered. Five interventions warranted further investigation due to meeting the age requirements, specifically addressing language and being delivered by professionals. These are presented in Table 1. 1 below. Further information on interventions that did not meet the current search criteria are available in Appendix 2 and in the What Works technical annexe (Law et al., 2012) (see also Law et al. 2017).

Table 1. 1

Evidence-based interventions currently implemented to support children with language difficulties in the early years

Intervention and References	Description	Focus of the Intervention	Target Age Group	Level of Evidence	Recorded Outcome and Effect Size (if provided)[a]	Therapy Delivery	Treatment Intensity
Focused Stimulation (Girolametto et al., 1996; Wolfe & Heilmann, 2010)	Focused Stimulation supports language by drawing attention to particular vocabulary (Girolametto & Pearce, 1996) and grammatical elements. An experienced speaker uses the same construction repeatedly in sentences, varying its position, to highlight its use, stress and positioning. Within Focused Stimulation, techniques such as repetition, modelling and melodic intonation and stress are used (Gladfelter et al., 2011). Focused Stimulation is supported by moderate evidence.	Targeted/ specialist	Pre-school	Moderate	Vocabulary size 0.88	Specialists, teachers and assistants	16 sessions containing 10 words each (Wolfe & Heilmann, 2010)

Morphosyntactic intervention (Camarata et al., 1994; Cleave & Fey, 1997; Fey et al., 1993; Tyler et al., 2006)	Morphosyntactic Intervention supports language through auditory awareness, focused stimulation and elicited production activities which are implemented to promote syntax and morphology. The intervention promotes use of target constructions, including correct morphological endings, and the therapist provides correct recasting where necessary. It has also been suggested that Morphosyntactic Intervention may have a positive effect on phonology (Tyler et al., 2006). There is a high level of overlap with other interventions e.g. focussed stimulation. The practitioner reduces the level of support over the three cycles. Morphosyntactic intervention is supported by moderate evidence.	Targeted	Pre-school and primary	Moderate	Not provided	Specialists	Two 50-minute sessions per week for 12 weeks

Nuffield Early Language Intervention (Bowyer-Crane et al., 2008; Fricke et al., 2013, 2017)	The Nuffield Early Language Intervention (NELI) targets multiple facets of oral language including vocabulary, listening, phoneme awareness, grammatical competence, and narrative skills. The intervention is based on the Oral Language Programme intervention (Bowyer-Crane et al., 2008) (see below) and is supported by a moderate level of evidence.	Specialist	Pre-school and primary	Moderate	Oral Language Skill 0.21 (20 week programme)- 0.3 (30 week programme)	Teachers and assistants (following training from a specialist)	Three 20-minute sessions for 10 weeks in nursery; 57 30-minute group sessions and 37 15 minute one-to-one sessions over 20 weeks in reception
Oral Language Programme (Bowyer-Crane et al., 2008)	Utilised in the Nuffield Early Language Intervention, the Oral Language (OL) Programme covers a broad range of oral language skills. The activities are derived from established sources including Time to Talk (Schroeder, 2001) Black Sheep Press materials (Rippon, 2002), and Rhodes to Language (Rhodes, 2001). Children are required to listen to and remember instructions to aid language development. Children are taught vocabulary relating to the	Targeted	Pre-school and primary	Moderate	Vocabulary 1.15 Comprehension - 0.44	Teachers and assistants	Alternating daily one-to-one 20-minute and group 30-minute sessions over 20 weeks

	tasks and are encouraged to utilise what they learn in different contexts such as show and tell, storytelling and object description tasks. TAs model and recast constructions when children make mistakes. The OL Programme is supported by moderate evidence.						
Talk Boost KS1 (Lee & Pring, 2016)	Key Stage 1 Talk Boost highlights the importance of SLCN to school staff and encourages incorporating positive strategies into whole class teaching practice. The intervention embeds SLCN supporting practices into classrooms and enables teachers to identify children with language difficulties, aiming to close the language gap between children with and without difficulties. Key targeted areas include conversations, narratives, sentence building, vocabulary and listening skills. Talk Boost is supported by moderate evidence.	Targeted	Pre-school and primary	Moderate	Language 0.52 Grammar 0.1	Teachers and assistants	Thirty minutes three times per week for 10 weeks

a Cohen's d Standardised Mean Difference (SMD) effect sizes as they are reported in the What Works Database (J. Cohen, 1988). An effect size of 0.2 represents a small effect size, 0.5 represents a medium effect size and an effect size of 0.8 or above is considered large.

The interventions included in Table 1. 1 were selected for further examination because of their relevance to the present study evaluating a targeted/specialist intervention to support sentence building in the early years, delivered via specialists, teachers or assistants. Notably, only six interventions were identified for this age group and they tend to focus on vocabulary and expressive language outcomes. The remainder of this section presents a discussion of the interventions that have been identified for children within the target demographic and concludes by summarising the current state of the evidence.

1.3.3 What are the Characteristics of the Identified Interventions?

The Nuffield Early Language Intervention (NELI) (Bowyer-Crane et al., 2008; Fricke et al., 2013, 2017), which draws on the Oral Language Programme (Bowyer-Crane et al., 2008), targets children's expressive language and grammatical competence by supporting vocabulary, narrative skills, active listening and oral language skills.

NELI is delivered via a combination of group and one-to-one sessions with trained teaching assistants and in the study children either took part for twenty (reception only) or thirty (nursery and reception) weeks. The first 10-week segment of the intervention is delivered in nursery and consists of 10 hours of group therapy across three twenty-minute sessions per week. The following twenty weeks are delivered in reception and consist of fifty-seven thirty-minute group sessions and thirty-seven 15-minute one-to-one sessions, with a total of 9.25 hours of one-to-one and 28.5 hours group intervention in reception.

Morphosyntactic Intervention (Camarata et al., 1994; Cleave & Fey, 1997; Fey et al., 1993; Tyler et al., 2006) is delivered by specialists in group and one-to-one sessions over 12 weeks. The intervention focuses on three main areas. First, auditory awareness activities are implemented to increase the child's awareness of morphological items. Next focused stimulation exposes the child to a rich input, repeatedly demonstrating target constructions, to highlight the syntactic features of language such as morphological endings and sentence construction. As part of this, recasting and expanding are used to highlight correct morphological endings. Finally, elicited production tasks aim to elicit multiple uses of the target morphemes. The practitioner gradually withdraws support enabling the child to develop independence when using the target morphemes. Focused stimulation (Girolametto et al., 1996; Wolfe & Heilmann, 2010), is part of Morphosyntactic Intervention but has also been evaluated as a stand-alone intervention. It exposes the child to multiple incidences of the same target structure or morphological form to promote its correct use and withdraws support gradually to increase the child's independence.

Key Stage One Talk Boost (Lee & Pring, 2016) targets vocabulary, listening, storytelling and sentence building through activities designed to support vocabulary comprehension and use, narrative building, sentence construction, listening skills and social interaction. Intervention is delivered in groups of four children and is primarily delivered by Teaching Assistants. The 30 minute sessions are delivered three times per week over 10 weeks. A further aim of the intervention is to enhance teacher's understanding and awareness of SLCN, enabling them to better identify children with language difficulties and provide continued support.

Several other interventions which share relevant characteristics with BEST were identified during the present review but were not included as they are designed for school-age children. These interventions include Shape Coding (Ebbels, 2007), the Strathclyde Language Intervention (Boyle et al., 2007; McCartney et al., 2010) and Colourful Semantics (Bolderson et al., 2011; A. Bryan, 2008; Guendouzi, 2003; Spooner, 2002). Despite targeting older children, these interventions contain components that are relevant to BEST, specifically their support of morphological elements of language. The Colourful Semantics approach uses visual representations of underlying sentence structure and facilitates the association of syntax and semantics. Similarly, the Shape Coding system uses colours, shapes and arrows to explicitly highlight the grammatical rules of English. The Strathclyde Language Intervention provides children with strategies to enhance memory, vocabulary, comprehension and narrative skills. For grammar in particular, the intervention supports children to use and understand syntactic markers such as plurality, tense and sentence structure using an approach similar to that of Colourful Semantics. Further information about these interventions is available in Appendix 2.

1.3.4 Are there any Effective Language Interventions for Pre-school Children?

The intervention effect sizes, where provided, range from negligible to large with the largest effects seen for language and vocabulary measured across a variety of standardised and nonstandardised measures (see Table 1. 1). This variation highlights the lack of consensus about which measures are used, and results in studies that can be difficult to compare effectively. Further, few studies provide estimates of the effects of intervention on morphology and syntax. The majority of effect sizes reported indicate that the intervention is educationally useful (Kilburn et al., 2014; What Works Clearinghouse, 2007), however more evidence is clearly needed to substantiate these findings.

1.3.5 What Level of Evidence Supports the Interventions?

Table 1. 1 reports the assigned level of evidence which indicates the strength of the evaluation design (strong, moderate or indicative, see above and also Law et al. (2012); Millard et al. (2009)). All interventions examined in this review are supported by a moderate level of evidence; that is they have been evaluated with a single randomised-controlled trial or quasi-experimental trial (Law et al., 2012). Despite this, only the Nuffield Early Language Intervention reports substantial development, piloting and feasibility work prior to a definitive trial (see 1.7 for a discussion of phases of evaluation) (Bowyer-Crane et al., 2008; Fricke et al., 2013; Fricke et al., 2017). It is not clear whether this work has taken place for other interventions, but it is not sufficiently reported if this is the case.

1.3.6 What is the Nature of the Intervention Intensity?

Intervention intensity was reported for all interventions, although it varied across studies in some cases. Children who were involved in these studies received an approximate average of one hour and

20 minutes per week, or around 16 minutes per day over an average of 18 weeks. Only NELI was tested under different intervention intensity conditions (Bowyer-Crane et al., 2008) and conclusions cannot be drawn about the appropriateness of the prescribed intensities.

1.3.7 What are the Explicit or Implicit Theoretical Underpinnings of the Interventions?

None of the interventions identified for the target participant group and therapy context in this review report sufficient theoretical underpinnings which is a serious shortcoming of the pool of available interventions. There is also little mention in the studies of the hypothesised 'active ingredients' that are responsible for affecting change (i.e. the mechanisms implicated in informing or teaching children (Fey & Finestack, 2009)). Some of the interventions are implicitly underpinned by social-interactionist theories (e.g. Focussed stimulation and Morphosyntactic Intervention), however for others it is difficult to infer any theoretical motivation other than very general theories of learning. Notably, no study makes explicit reference to usage-based theories of language acquisition which is under examination in the present study (see 1.4.2).

1.3.8 Interventions Currently in Use: Conclusions

The interventions identified in this section are designed to support language development in children with language difficulties in the early years. Interventions are carried out in a variety of settings including mainstream and special school settings, and occasionally in clinics. Specialists, trained teachers and teaching assistants are all involved in the delivery of interventions, and one intervention, Talk Boost KS1 also directly targets teacher's awareness of SLCN. The interventions identified are all supported by moderate levels of evidence, although there is variation in the quality of this evidence (Law et al., 2017) and the Nuffield Early Language Intervention papers are the only ones that explicitly discuss sufficient piloting and feasibility work before the main trials (Bowyer-Crane et al., 2008; Fricke et al., 2013, 2017).

There is some evidence supporting language and communication interventions, with effect sizes ranging from negligible to very large (see Cohen, 1988). Effect sizes were not reported for Morphosyntactic Intervention. Overall, there is a dearth of theoretically motivated early-years interventions which have been robustly evaluated in accordance with a recognised framework (Law et al., 2017) (see 1.7.3 for a review of existing guidelines). This results in a limited evidence base for therapists and practitioners to refer to when making clinical and theoretical judgements about intervention delivery. More high quality evidence is required to increase the evidence-base and improve clinical decision making (Law et al., 2017; Law et al., 2012).

Given the lack of theoretically motivated interventions identified in this review, the following section presents contrasting theories of language development and their associated accounts of language difficulties from each theoretical perspective. It considers the extent to which various theories can account for language difficulties to determine which may be promising as the basis from which to develop language interventions. Emphasis is placed upon usage-based theory due to the number of potential 'active ingredients' and its transferability to components of language intervention.

Section 1.5.1 then describes the theoretical underpinnings and hypothesised 'active ingredients' of BEST. One such 'active ingredient' is the use of sign alongside speech to represent the content and morphology of target sentences in the visual domain, the evidence for which is considered in 1.6.

1.4 Theories of Language Development and Developmental Language Difficulties

We saw in the preceding review that there are rarely explicit theories underpinning language intervention that are reported in intervention studies. As a result, it is often unclear why elements of interventions were designed in the way that they were. Further, it is difficult to identify the 'active ingredients' that underpin change, as they are not related to current theories (Law et al., 2017). This omission is unlikely to be due to a paucity of theories, since several prominent proposals aim to explain the mechanisms underpinning language difficulties. Rather, there seems to be a general lack of application of theory to intervention development and testing, possibly because certain theories such as nativist accounts do not suggest clear mechanisms for the improvement of language difficulties. Clearly, there is a need to identify theories that are applicable to language intervention and then to develop and adapt interventions in line with their theoretical underpinnings.

There are two major 'families' of theory of the nature of typical language acquisition in current academic discourses (Ambridge & Lieven, 2011): Nativist linguistics (e.g. Chomskian linguistics (Chomsky, 1965; Chomsky, 2006)) implies innate linguistic knowledge (Chomsky, 1965; Chomsky, 2006) and in the case of nativist generativist accounts, innate knowledge of grammatical parameters. Language difficulties are typically explained by domain-specific deficits in linguistic knowledge or computational processes. Usage-based, or constructivist, theories on the other hand assume that children are born without innate linguistic rules- a 'tabula rasa' or blank slate (see Locke, 1954)- and instead acquire language through repeated exposure and a desire to communicate socially (Tomasello, 2003). Under such accounts, language difficulties may be explained by domain-general deficits such as impaired processing or working memory (WM). The following sections outline each of these accounts and provide evidence for their claims about typical language acquisition as well as language difficulties. We see that usage-based theories of language acquisition can be more readily applied to interventions, as they do not make claims about an innate facility for language learning which would naturally be difficult to treat, as well as postulating potential 'active ingredients' of intervention.

1.4.1 Nativist Theory of Language Acquisition

Traditional Chomskian linguistic theory (Chomsky, 1965; Chomsky, 2006) suggests that children are born with a 'language acquisition device' (LAD) which facilitates language acquisition. Chomsky argues that the LAD contains a 'Universal Grammar', or innate knowledge of language structures, consisting of phrase structures and language principles that apply across all known languages (e.g. the 'extended projection' principle which stipulates that every sentence requires a subject). Additionally, these accounts argue that children possess parameters for syntactic elements of language that vary cross-linguistically, that are activated once the child is exposed to a particular language in their environment (e.g. the 'head-direction' parameter which determines whether the order of words in a verb phrase matter in a specific language (such as in English and Korean) or not (such as in German)) (Chomsky, 1982). In addition to Universal Grammar, the 'Merge' function enables children to make use of hierarchical structures within language to form new linguistic structures by combining linguistic items (Chomsky, 1995; Yang et al., 2017).

1.4.1.1 Domain-General Accounts of Language Difficulties

A number of accounts within the generativist sphere argue that deficits in linguistic knowledge or computational processes cause children with language difficulties to make more errors with specific linguistic and grammatical features than typically developing (TD) children (Leonard, 2014).

One early linguisitc knowledge deficit account is the Extended Optional Infinitives (EOI) hypothesis, which suggests that during development, chidren enter a phase in which they optionally mark verb tenses (Wexler, 1994). This results in errors of tense and agreement when verb tenses are not marked. For example, in English the third person present singular is inflected with an [-s], e.g. [she run**s**]; [he cook**s**] (with the exception of auxillary modal verbs)**.** Omitting the inflection results in an ungrammatical utterance e.g. [she run*]; [he cook*]. The EOI hypothesis suggests that children with language difficulties spend longer in the phase of optional tense marking, producing more errors and falling behind their TD peers (Rice & Wexler, 1996a, 1996b; Rice et al., 1995).

An extension of the EOI hypothesis is the Extended Unique Checking Constraint (EUCC) hypothesis proposed by Wexler (1998, 2003). This hypothesis suggested that children with language impairment are limited to checking only one functional category (e.g. tense; subject agreement) for longer than TD children. Therefore in a sentence which requires checking of both the tense and subject agreement, e.g. [Last week she saw me], one will not be checked, producing either [Last week her saw me*] or [Last week she see me*].

Explaining broader grammatical deficits, the Representational Deficit for Dependent Relations hypothesis (RDDR) (van der Lely, 1998; van der Lely & Stollwerck, 1997), a computational deficit hypothesis, suggests that children have a deficit in a range of grammatical principles resulting in specific grammatical difficulties. According to this account, children struggle with features such as tense, subject-verb agreement and trace position of e.g. wh- questions to the correct word within a sentence.

The RDDR was extended to account for morphology and phonology deficits in the Computational Grammatical Complexity (CGC) account which focuses solely on children with Grammatical SLI (G-SLI)- a hypothesised subgroup of children with SLI who display purely grammatical language difficulties (van der Lely, 2004, 2005). This account suggests a limitation in children's computations of hierarchical, structurally complex forms in linguistic components such as morphology and phonology, as well as syntax, due to the complexity of morphological and phonological cues.

It has also been suggested that children struggle to preserve the thematic roles of elements during their movement, with evidence from studies in Hebrew, a language that contains homographs (words with the same orthography but differing meaning and pronunciation). Hebrew speaking children with SLI struggle to correctly interpret the meaning of object relative sentences containing ambiguous homographs. The homograph provides two interpretations with only one correctly relating to the trace object (Friedmann & Novogrodsky, 2004, 2007; Novogrodsky & Friedmann, 2006; Sukenik & Friedmann, 2018), indicating that children with DLD have difficulty with long-distance dependencies when it is necessary to retain the trace object.

Other accounts have sought to explain additional errors that are more typical in children with DLD such as the Grammatical Agreement Deficit Account (Clahsen, 1989, 1991) and the Agreement/Tense Omission Model (Schütze & Wexler, 1996). The key point here is that these models imply deficits in specific linguistic or computational knowledge in the errors exhibited by children with DLD. They are generally unable to account for broader deficits commonly seen in children with DLD which are discussed below, and they generally fail to identify mechanisms which could be readily exploited to develop theoretically motivated interventions.

As such, we move on to look to alternative explanations that can better explain language difficulties and provide greater opportunities for to support children with DLD through language intervention (MacWhinney, 2014).

1.4.2 Usage-Based Theories of Language Acquisition

Another prominent group of theories of typical language acquisition are usage-based theories (e.g. Tomasello, 2003) which draws on numerous previous constructivist and social-pragmatic accounts (e.g. Bowerman, 1975; Braine & Bowerman, 1976; Bybee, 1995; Croft, 2001; Dąbrowska & Lieven, 2005; Goldberg, 1995; Pine & Lieven, 1993). In contrast to nativist theories of language acquisition, Tomasello's usage-based theory argues that language is learned through: 1) exposure to the input provided by parents and caregivers; 2) a desire to communicate socially; and 3) social-cognitive and pattern-finding skills. Usage-based theory is concerned principally with the grammatical dimension of language acquisition.

Under this account, at around 12 months old children develop the social-cognitive skill of intention reading, bringing about a greater level of social understanding. There is extensive empirical research examining the role of social-cognitive skills (e.g. skills that enable children to understand and participate in the social world (Forrester, 1992)) on general learning processes. Fourteen-to-18-month-olds who watched an adult try but fail to achieve a goal through an action reproduced the action and the goal, suggesting that they were not only imitating the adult, but understood the intentions behind their actions, even when they adult did not succeed (Meltzoff, 1995). Additionally, when 16-month-olds watched a paradigm where an adult sometimes succeeded and sometimes failed to achieve a goal through an action, they imitated more of the successful outcomes- the intentional actions- than the accidental ones (Carpenter et al., 1998).

Intention reading consists of three subskills which facilitate the understanding of symbolic communication: 1) the ability to attend to joint attentional frames with others; 2) the ability to understand the communicative intentions of others; and 3) the ability to implement 'role reversal imitation'- a form of social, cultural learning. Two other essential skills, pattern-finding and statistical learning, which are present from earlier than one year old, begin to support the construction of grammar once the child starts to understand linguistic symbols.

Joint attentional frames are a form of triadic attention used to establish common ground between, in a prototypical example, a child, a caregiver and a referent. When part of a triadic joint attentional frame, the child is well prepared to understand the common representational format of both their own and the adult's role from an 'outside perspective'. This enables the child to establish mutual understanding with their caregiver and creates the ideal environment for 'understanding communicative intentions' (Ambridge & Lieven, 2011; Tomasello, 2003; Tomasello & Farrar, 1986; Tomasello & Todd, 1983). It must be noted however that evidence suggests cross-cultural variation and differential outcomes for children with particular developmental disorders (Akhtar & Gernsbacher, 2007).

The second skill, 'understanding communicative intentions', is hypothesised to take place within the 'joint attentional frame' discussed above. There is a distinct focus on the communicative intentions, as opposed to just attending to an object, whereby the child recognises that the adult is communicating with the specific intention of them understanding something (Ambridge & Lieven, 2011; Tomasello, 2003).

The third skill under the umbrella of intention reading, 'role-reversal imitation' is a form of cultural, or social learning. Pre-linguistic children, before nine months of age, can understand that other people have intentional relationships towards the world, and therefore carry out actions with intentionality to achieve goals (Tomasello, 2003). Because of this, the child will pay attention to adults when they understand that the adult is trying to achieve a goal and will often imitate the adult. At around nine to 12 months when they acquire the ability of triadic attention this is taken further; children learn to use communicative symbols towards others in the same way others use them towards the child. In this way symbols become socially shared (Tomasello, 2003). This is important for the child to become a productive user of language, rather than just an imitator.

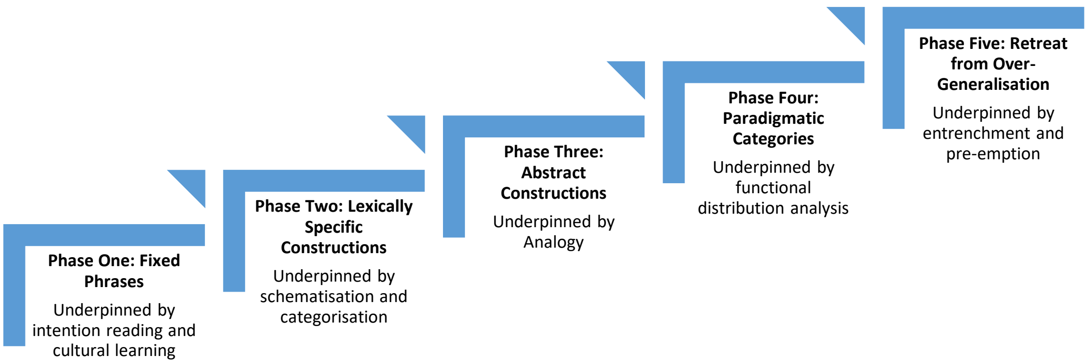
As well as the intention reading umbrella of social-cognitive skills discussed above, the skills of pattern-finding and statistical learning are also crucial for early word and construction learning under the usage-based account. These, like the social-cognitive skills, are domain-general and are present pre-linguistically (discussed further below) (Haith & Benson, 1998; Kirkham et al., 2002; Saffran et al., 1999). When children develop the ability to understand linguistic symbols at around one year of age, these skills are harnessed to facilitate learning about which parts of constructions vary, and in what ways, based on statistical probabilities of previously heard exemplars (discussed further below).

1.4.2.1 Phases of production and the underlying mechanisms

The social-cognitive and pattern-finding skills outlined above are hypothesised to support the child's ability with linguistic constructions to move through five distinct stages. This moves the child from using fixed phrases, which are rote-learned, inflexible associations with cultural routines and pragmatic functions, to progressively more abstract phrases, enabling eventual arrival at the fully abstract adult end state. It is at this point that the speaker can produce correct, novel utterances without producing over-generalisation errors (ungrammatical utterances caused by an over-generalisation of a learned rule e.g. [you cried her*], meaning [you made her cry], a transitive causative over-generalisation of [she cried] (Bowerman, 1988)) (Ambridge & Lieven, 2011; Tomasello, 2003). This process is not linear but varies across languages and individuals with children moving through the stages at different rates for different constructions. The stages are as follows: 1) using fixed ‘frozen’ phrases; 2) using lexically specific constructions; 3) using abstract constructions; 4) using paradigmatic categories; and 5) a retreat from over-generalisation errors to arrive at an adult like end-state. Each of the stages are underpinned by specific mechanisms which are discussed in more detail below. Figure 1. 1 shows these stages and highlights the key mechanisms underpinning each stage.

Figure 1. 1

The five stages children’s abilities move through and the proposed mechanisms underpinning them (Tomasello, 2003)



The first three stages 1) fixed phrases; 2) lexically specific constructions; and 3) abstract constructions are of particular importance in an early years context, since children do not complete the retreat from over-generalisation errors for a number of years (Ambridge & Lieven, 2011). These three stages are implicated in BEST and are therefore the focus of this section (see Tomasello, 2003 for an account of the final two stages).

1.4.2.1.1 Phase One: Fixed ‘Frozen’ Holophrases

Fixed, ‘frozen’ phrases, also known as holophrases, are inflexible words or constructions associated in some way with a referential event or object. A child may form an association between the phrase ‘get-dressed’ and the action of being dressed by a caregiver. Although this action is paired with a cultural routine and the pragmatic function of dressing, the association is broad and therefore inflexible. The child may not say ‘get-dressed’, for example, if they see their sibling being dressed, because the phrase at this point applies only to their own perspective of being dressed. As such, fixed phrases are like single lexical items in that they are deployed as ‘wholes’ and the words within them are not used or combined in other flexible ways.

Fixed phrases are initially learned through processes of intention reading and cultural learning. The umbrella of intention reading skills (joint attentional frames, understanding communicative intentions and role-reversal imitation/cultural learning) enable children to understand that the ‘noise’ adults produce has communicative intent, and that the adult wants the child to attend to referents triadically with them. Tomasello (2003 p. 23) represents this in the following model:

You intend for [me to share attention to [X]]

This model demonstrates that the child understands the caregiver’s intention for the child to share their attention between the caregiver and the appropriate referent within the joint attentional frame (Tomasello, 2003).

Once a child is part of a triadic joint attentional frame with their caregiver and is sharing communicative intention, they must use pattern-finding and statistical learning (see 1.4.2.1.5) to successfully segment the speech and the part of the attentional frame that the speech is referring to, in order to understand it. Finally, the child can store the newly learned holophrase in their long-term memory for retrieval when needed. It is through this process that children begin to learn linguistic symbols, forming the foundation of language.

1.4.2.1.2 Phase Two: Lexically Specific Constructions

Lexically specific constructions are less concrete than frozen phrases, but they are not yet fully abstract. These constructions consist of two components: firstly, fixed ‘frames’, which remain constant across construction use such as ‘He’s ACTIONing it’ and ‘X gone’. Additionally, the frames have room for flexible ‘slots’ which vary depending on the communicative intent, for example if a child wanted to say ‘He’s EATing it’ or ‘DADDY gone’.

Children typically hear many examples of the same word or phrase applied in differing contexts, leading the formation of ‘schemas’, or knowledge about how that word is used across contexts. The slots and frames are represented as schemas, generated through a process of ‘schematisation’ based on previous exposures. Categorisation also takes place at this stage whereby children group the frames and slots into groups of like actions or objects. Table 1. 2 displays frames and slots for two example constructions, and possible groups that the child might categorise them into.

Table 1. 2

The five stages children’s abilities move through and proposed underpinning mechanisms (Tomasello, 2003)

Example number	Frame	Example Frame Category	Slot	Example Slot Category	Example slots
1	‘**He’s** ACTION**ing** **it**’	a group of events in which somebody manipulates an object	**‘ACTION’**	A group of actions that can be performed on the object	Eat, play, move
2	‘X **gone**’	A group of events in which something that was visible before is no longer visible	**‘X’**	A group of objects that can disappear or no longer be seen	Food, toy, Daddy

*Note. Relevant items for each example are displayed in **bold***

1.4.2.1.3 Phase Three: Abstract Constructions

Abstract constructions are flexible, nonconcrete representations. Beginning with constructions that they have already developed, children determine the appropriate roles and use the mechanism of analogy to abstract across to new constructions with similar function or structure. A functional example of this process is a child taking the known construction ‘He’s ACTIONing it’ (see Table 1. 2 above) and establishing its semantic categories (e.g. AGENT + ACTION + PATIENT). The mechanism of analogy then allows the child to apply this abstract construction to new scenarios and across different verbs, such as (e.g. EATER + EATS + THING EATEN). Syntactic analogy can be applied in the same way, for example if the child hears '**The** man walk**s**', they may then create the abstract representation of 'DETERMINER + AGENT + VERB' and use this to produce '**The** lady jump**s**'. The development of abstract constructions also promotes more effective learning of similar novel constructions, thus further supporting the child's language development (Langacker, 2000).

1.4.2.1.4 Pattern-Finding and Statistical Learning

Children's skills in pattern-finding in non-linguistic auditory and visual stimuli from a young age have been well documented (Haith & Benson, 1998; Kirkham et al., 2002; Saffran et al., 1999). Application of pattern-finding skills to linguistic auditory stimuli are particularly relevant for early speech segmentation. Pre-linguistic infants who are exposed to a continuous stream of trisyllabic nonsense words are able identify those which they have previously heard based on the transitional probabilities, i.e. the likelihood of sounds co-occurring (Kirkham et al., 2002; Saffran et al., 1996). There is also evidence that children can utilise the pattern structure of trisyllabic nonsense words (e.g. ABB pattern, such as 'wididi'; 'delili'), even when the test stream does not contain the same words and the transitional probabilities therefore do not help (Gomez & Gerken, 1999; Marcus, Vijayan, Bandi Rao, & Vishton, 1999). This evidence supports the role of statistical learning and pattern-finding for segmenting and processing language.

1.4.2.2 Domain-Specific Accounts of Language Difficulties

There is extensive evidence for domain-general deficits affecting general cognitive processes such as processing and working memory, which in turn result in compromised language development. Leonard et al. (1992) suggest that general perceptual or resource/capacity limitations could be caused by observed difficulties in children with SLI, and Norbury et al. (2001) argue that limited capacity could cause morphosyntactic difficulties. A number of studies support these claims: Bishop, (1994) identified children who optionally omitted inflectional morphology in some, but not all, utterances. This indicated that reduced processing capacity and rate, especially for multiple operations being handled simultaneously, were responsible for the omission of grammatical morphemes. This was supported by the fact that ommissions tended to occur towards the end of sentences, suggesting capacity had been reached.

Kail (1994) found group mean response times (RTs) to be longer for children with language difficulties on a variety of tasks, suggesting slower processing times. These findings can explain many of the linguistic and non-linguistic impairments associated with DLD, however they generally fail to account for more specific patterns of morphological errors discussed in the section above (e.g. Rice & Wexler, 1996a; van der Lely, 1998).

There is reasonable evidence that children who meet the criteria of DLD/SLI have domain-general perceptual deficits (Bishop & McArthur, 2005; McArthur & Bishop, 2004; Shafer et al., 2007). It is less clear whether these affect domain-specific linguistic processes. Drawing on the theoretical work of Lapointe (1983) and Pinker (1979, 1984), Leonard (1989) proposed the Surface Account to explain children with SLI's difficulty with grammatical morphology. This account suggests children have reduced capacity for processing taxing morphological roles, sometimes resulting in compromised morpheme processing, while greater exposure is needed before they can become established in the child's syntactic abilities.

Further evidence for domain-general limitations, and against linguistic and syntax specific deficits, comes from studies investigating working memory demands in sentence comprehension. Montgomery & Evans (2009) studied two core components of working memory (phonological short-term memory storage and attentional resource capacity/allocation) and their association with complex sentence processing. They demonstrated that children with SLI showed weak attentional resource capacity/allocation, which was associated with poor sentence comprehension. They also found that simple sentence comprehension placed greater demands on phonological short-term memory for children with DLD than for younger TD children matched on working memory abilities.

During the 1990s a prominent temporal processing deficit theory suggested language impairments are the product of difficulty in recognising and sequencing the spectrotemporal structure of speech (time and frequency related signals (see Atencio & Schreiner, 2008)), and that temporal properties of speech could be gradually manipulated via computer assistance to improve temporal processing ability. Temporal processing deficit hypotheses and the associated FastForword intervention (Tallal, 1980, 1990, 2004) have been generally discredited (Bishop et al., 1999; Gillam et al., 2013). It is instead argued that demands on working memory may have led to observed difficulties with speech perception in previous studies (Coady et al., 2005).

Whilst we saw in 1.4.1.1 that hypotheses focussing on specific grammatical or linguistic deficits cannot always explain the broader difficulties observed in children with DLD (Leonard, 1996, 2014; Ullman & Pierpont, 2005), domain-general hypotheses conversely struggle to account for the specific syntactic errors observed in children with specific grammatical difficuties (Leonard, 2014).

More recently, neurobiological theories have provided evidence for the origin of language impairment, building upon domain-general cognitive hypotheses outlined above. The Procedural Deficit Hypothesis (PDH) proposed by Ullman & Pierpont (2005) describes the hypothesised neurobiological basis for previously discussed cognitive deficits. The PDH suggests abnormal development of brain structures that are part of the procedural memory system, including the caudate nucleus in the basal ganglia and Broca's area in the frontal cortex, cause language impairment. These areas support learning of broad cognitive and motor rules, skills and procedures, as well as linguistic abilities such as grammar, working memory, processing and word retrieval. On a processing level, it is hypothesised that rule-based tasks such as grammar and morphology depend on the procedural memory, and are thus impaired, whilst 'idiosyncratic mappings' and the mental lexicon depend on declarative memory and are thus relatively spared (Ullman, 2001a, 2001b).

1.4.2.3 Usage-Based and Emergentist Theories of Language Disorder

Usage-based theories suggest that children require sufficient and appropriate processing skills to effectively analyse and process the language they are exposed to. As previously discussed, in order to learn language, this account stipulates that children must pair the pragmatic functions they witness with the language input they are exposed to. Children must also be able to successfully identify patterns within language, both at the phonological and the construction level (Tomasello, 2003), and must be interested in the activities of others around them (see also Shonkoff & Phillips, 2000).

Recent Emergentist accounts of language difficulties place greater emphasis on the learning mechanisms underpinning language acquisition, and argue that the child's unique interaction with their environment (e.g. social, acoustic and linguistic cues and communicative interaction) facilitates the *emergence* of language (Evans, 2001). This account sees language as a dynamic state evolving as a function of acquisition, whereby children move between stable and unstable states to facilitate development (e.g. one-word utterances which the child is familiar with, to two-word utterances which are less stable). Through exposure, language develops via a process of statistical learning and probabilistic inference, rather than through the application of abstract rules as in the constructivist account discussed above. Evidence that impaired phonological processing networks can result in specific grammatical deficits (Joanisse, 2004, 2009; Joanisse & Seidenberg, 2003; Karmiloff-Smith & Thomas, 2003; Karmiloff-Smith, 1998; Thomas & Karmiloff-Smith, 2002) and working memory deficits (Joanisse & Seidenberg, 2003) lend support to the emergentist account.

Chiat's (2001) 'mapping theory' suggests that poor phonological processing results in a reduced ability to parse individual words in a stream of speech, causing an impairment with mapping 'form chunks' and 'meaning chunks'. Difficulties are particularly apparent for processing phonemic contrasts when sounds are presented rapidly, or in long strings. This causes a difficulty mapping items onto their lexical forms, resulting in a smaller vocabulary and a difficulty learning new words. Emergentist accounts therefore predict less impairment with directly observable verbs than temporal forms, and hypothesise that words which have multiple contexts should show the same pattern across uses (Evans, 2001).

Whilst strikingly similar to the constructivist account outlined above (see chapter 5 for a discussion), emergentist accounts are distinct in that they give a specific account of SLI whereby the child with phonological processing difficulties is following the least effortful developmental trajectory, resulting in potentially maladaptive communication strategies that may, on the surface, present like domain-specific impairments (Evans, 2001). They advance the emphasis on *how* children learn language, providing insight into *how* therapy should be addressed and which cognitive mechanisms might be harnessed to support the learning process (Chiat, 2001; Evans, 2001). Despite their clarity however, emergentist and usage-based theories still fail to define exactly how a phonological processing impairment results in the phenotypic expression of language impairment (Mckean et al., 2013).

Few researchers have thus far attempted to explain language disorders from a usage-based perspective. Usage-based, constructivist accounts of language intervention have conceivable potential for development into interventions due to explicit cognitive mechanisms underpinning each stage of language development. In the BEST manual, McKean et al. (2013) suggest an application of usage-based theory for underpinning interventions to treat DLD. They posit that young children with DLD may remain in the stage of producing fixed phrases rather than moving on to producing lexically specific, and then abstract, constructions (see discussion of usage-based theories above), explaining the lack of productivity in their sentences. This is supported by the lack of flexibility and productivity of sentence structures produced by children with DLD (Bishop & Mogford-Bevan, 1993; Conti-Ramsden & Windfuhr, 2002; Thordardottir & Weismer, 2002).

1.4.3 Conclusion

Application of usage-based linguistics to the development of language intervention is both warranted and feasible. Usage-based theories arguably align more closely with current understanding of language acquisition than nativist theories (MacWhinney, 2014). Further, usage-based theory offers more specific accounts of cognitive mechanisms that may be harnessed by interventions. It could also explain the successful outcomes of current interventions implicitly underpinned by social-interactionist theories (see 1.3.7). One of the key aims of this thesis is to test the efficacy of a usage-based intervention to determine its theoretical and clinical implications.

The following section of this chapter discusses the specific applications of usage-based theory to Building Early Sentences Therapy, demonstrating the ways in which hypothesised mechanisms are harnessed to support children's language learning.

1.5 Usage-Based Language Intervention

As outlined in the previous section, the cognitive mechanisms posited in usage-based theories are readily applicable to language intervention. Despite this, to the Author’s knowledge there is only one preliminary study reporting an intervention informed by usage-based theory for older children with DLD (Riches, 2013); although see also Bruns et al. (2019) for a recent application of usage-based linguistics to adults with Aphasia).

Riches (2013) tested a usage-based intervention targeting passive constructions (e.g. The cat was chased by the dog) in two eight-year-old children with SLI. The intervention was designed to harness implicit mechanisms posited to underpin acquisition of the passive construction (these mechanisms: ‘constructional grounding’ and ‘construction conspiracy’ (see Figure 1. 1) are beyond the scope of this thesis as they are not applicable in the early years). Although the evidence is extremely limited, Riches suggests that further research exploring interventions underpinned by usage-based theory is warranted given the promising outcomes.

1.5.1 Building Early Sentences Therapy

Building Early Sentences Therapy (BEST) (McKean et al., 2012) is one such theoretically motivated, complex language intervention designed for three-to-six-year-old children with severe language difficulties. A complex intervention is one containing multiple interacting components (Craig et al., 2008), for example BEST employs a systematically varied input, joint-action routines, a signing system representing content and morphology, and toys (discussed below).

BEST aims to harness the cognitive mechanisms implicated in the first three phases of language development: 1) intention reading and cultural learning (implicated in phase one); 2) schematisation and categorisation (implicated in phase two); 3) analogy (implicated in phase three) (Tomasello, 2003). BEST also supports the broader skills of mapping (Hirsh-Pasek et al., 2000) and retention (Ambridge et al., 2006; Janiszewski et al., 2003) which are hypothesised to support language development (see 1.5.1.4.6 below). These skills are harnessed to support children's ability to understand and produce two, three, and four clause element sentences (see Table 1. 3) effectively by moving through the stages of acquisition from 'frozen phrases' to increasingly abstract constructions.

Table 1. 3

Examples of two, three and four clause element sentences

Number of Clauses	Example sentence[1]
Two clause element sentence	The lady is walking
Three clause element sentence	The man is washing a cup
Four clause element sentence	The girl is giving the banana to the boy

Note: [1] Individual clauses are underlined

1.5.1.1 Sentence structures and Morphology Goals

BEST targets two specific language goals: 1) to develop the range and complexity of simple sentence structures the child uses in terms of the roles of the main sentence elements e.g. [Agent] [Verb] [Patient] (see emboldened text in example (1) below); and 2) to develop the consistent use of grammatical morphology (e.g. verb endings [-ing; -ed], determiners [a, the], auxiliaries [is]) (see emboldened text in example (2) below).

Example Constructions

(1)	The	**teddy**	is	**wash**ing	the	**banana**
	Det 1	**Agent**	Aux	**Verb**(+ing)	Det 2	**Patient**

(2)	**The**	man	**is**	smell**ing**	**the**	cake
	Det 1	Agent	**Aux**	Verb **Inflection(ing)**	**Det 2**	Patient

Throughout this thesis, sentence structure is used to refer to the roles of the main sentence elements and how they are combined (as in example 1) and morphology is used to refer to the determiners, auxiliaries and verb endings (as in example 2).

1.5.1.2 How is BEST Delivered?

BEST consists of 16 therapy sessions and corresponding homework booklets. Therapy is delivered in groups of one to six children in school or clinic settings. Sessions are delivered twice per week for eight weeks by a trained practitioner (henceforth referred to as Adult One) and a second adult, usually a Teaching Assistant (TA) known to the children (referred to as Adult Two) (variability to this role is discussed in detail later). Alternatively BEST may be delivered once per week for 16 weeks which should not affect the outcome (Bellon-Harn, 2012).

Over the course of a typical BEST intervention session, after a warm-up activity, children are presented with a series of verbs in simple sentences of varying clause-element and construction types (see Table 1. 4 below). For each session, there are three sets of model and target sentences. Each session consists of an input and an output phase.

Table 1. 4

Examples of input and output phases for 1, 2, and 3 argument predicate argument structures. Reproduced with permission from the BEST Manual (McKean et al., 2013, p. 13)

Number of Arguments	PAS	Input	Output
1	SET A: Agent + Action	1 The baby is laughing 2 The woman is laughing 3 The boy is laughing 4 The girl is laughing 5 The teddy is laughing 6 The man* is laughing	1 The man is **sitting** 2 The woman is **sitting** 3 The boy is **sitting** 4 The girl is **sitting** 5 The teddy is **sitting** 6 The baby is **sitting**
2	SET C: Agent + Action + Patient	1 The man is eating an apple 2 The man is eating an orange 3 The man is eating a banana 4 The man is eating a carrot 5 The man is eating a lolly 6 The teddy* is eating a banana	1 The **teddy** is **washing** an apple 2 The **teddy** is **washing** an orange 3 The **teddy** is **washing** a banana 4 The **teddy** is **washing** a carrot 5 The **teddy** is **washing** a spoon 6 The **teddy** is **washing** a cup
	SET G: Agent + Action + Patient + Locative	1 The baby is putting a spoon on the table 2 The baby is putting a cup on the table 3 The baby is putting a flower on the table 4 The baby is putting a key on the table 5 The man* is putting a phone on the table	1 The **man** is putting a spoon on the **bed** 2 The **man** is putting a cup on the **bed** 3 The **man** is putting a flower on the **bed** 4 The **man** is putting a key on the **bed** 5 The **man** is putting a phone on the **bed**
3	SET H: Agent + Action + Patient + Locative	1 The man is pouring milk into a cup 2 The man is pouring milk into a shoe 3 The baby* is pouring milk into a box	1 The **baby** is pouring **juice** into a cup 2 The **baby** is pouring **juice** into a shoe 3 The **baby** is pouring **juice** into a box
	SET I: Agent + Action + Patient + Locative	1 The man is putting a spoon on the bed 2 The man is putting a cup on the bed 3 The man is putting a flower on the bed 4 The man is putting a key on the bed 5 The Baby* is putting a phone on the bed	1 The **baby** is **pouring** juice in the shoe 2 The **baby** is **pouring** juice in the box 3 The **baby** is **pouring** juice in the cup

*Note. Variation is underlined; Contrast is in **bold**; Linking agents between input and output phases are marked with a * (see below for an explanation)*

During the input phase Adult One Models the sentences orally and signs the sentences (use of sign is discussed in detail in 1.6). Adult Two acts out the sentence with the BEST toys. Other verbal input such as praise is avoided so that all spoken input maps directly onto the visual input.

In the output phase sentences vary by one element. To elicit the target sentences, Adult Two acts out the target sentences with the toys and Adult One asks each child in turn 'What's happening, [*child's name*]?'. The child's response is recorded and Adult One produces the correct target sentence and signs. Section 2.8 provides a detailed description of intervention delivery procedures as they applied to this study.

1.5.1.3 Who Delivers BEST?

In the early years, intervention programmes to support children with language difficulties are often delivered by educational practitioners supported by SLTs. Children with more severe difficulties may receive direct intervention by SLTs, but for many children with milder difficulties, access to SLT services may only occur if their difficulties do not resolve after receipt of targeted support by educational staff. In line with this response to intervention model (Bayat et al., 2010), BEST is designed to be delivered in partnership with early-years school staff who have received training from a specialist, but it can also be delivered by an SLT, for example in a clinic setting.

1.5.1.4 BEST and Usage-Based Theory

BEST is designed to support children to move through Tomasello's first three phases of language development: phase 1) fixed phrases; phase 2) lexically specific constructions; and phase 3) abstract representations (see Figure 1. 1).

The cognitive mechanisms thought to underpin each of the respective stages are: 1) intention reading and cultural learning; 2) schematisation and categorisation; and 3) analogy. Additionally, BEST supports the learning mechanisms: 4) mapping; and 5) retention which are heavily implicated in language learning.

Key features of BEST support children to harness the cognitive mechanisms listed above, in line with Usage-based theory. These features are: 1) joint action routines and role reversal; 2) manipulation of the input; and 3) multiple exposures to analogous constructions. These features, and how they are hypothesised to support language acquisition are discussed below with examples of how they occur in BEST. Also see the BEST manual on the LIVELY website (LIVELY Research Group, 2019; McKean et al., 2013) for more detail on the mechanisms and their underpinnings.

1.5.1.4.1 Joint-Action Routines and Role Reversal

The first stage of Tomasello's model (see Figure 1.1) is the development of fixed phrases. This stage is underpinned by the cognitive mechanisms of intention reading and cultural learning discussed in 1.4.2.1.1. Joint-action routines support the child to infer (intention read) the communicative intention of the adult and therefore to identify referents that are being referred to, and for what communicative reasons. This process scaffolds the child's inferences about others' communicative intent. Role reversal allows the child to understand that, when they communicate with others, the roles are reversed, compared to if somebody communicates with them. This is essential for producing symbolic representations through cultural learning (Ambridge & Lieven, 2011; Tomasello, 2003) and is also promoted by engagement in joint action routines.

BEST supports these mechanisms by engaging children in structured joint action routines as part of the input and output phases of intervention sessions. These give children the opportunity to correctly infer the adult's intention for the child to attend to the referents through joint attentional frames and to understand role reversal. The combination of intention reading and role reversal allows the child to learn the meaning and the structure of the new sentences they hear in the routine (Tomasello, 2003).

1.5.1.4.2 Manipulation of the Input

The second stage of language development (see Figure 1.1) is the use of lexically specific constructions (Tomasello, 2003). The cognitive mechanisms of schematisation and categorisation underpin this stage and are promoted by BEST through the manipulation of the input the child hears. Schematisation is the process through which children recognise the predicate argument structure 'frames' which items can be inserted into, such as 'He's ACTIONing it' and 'X gone' (see 1.4.2.1.2).

This process produces lexically specific constructions that can be used with corresponding categories. Categorisation is the sorting of referents into categories of items that can be inserted into the 'frames', such as 'actions that can be done' 'inanimate objects that can be thrown'.

The BEST input was constructed to facilitate schematisation and categorisation. This input consists of multiple sentences with the same predicate argument structure where one element systematically varies (*either* the agent, action, patient or benefactive is varied per set, underlined in Table 1. 4 above). Examples include ['the <u>man</u> is walking', 'the <u>lady</u> is walking', 'the <u>girl</u> is walking'] and 'the man is eating an <u>apple</u>', 'the man is eating an <u>orange</u>', 'the man is eating a <u>cake</u>'. These items rotate throughout each sentence set and each session.

Variation of the agents, actions, patients and benefactives within the sentence sets supports children to form structures, or 'frames' through the process of schematisation. The variation between items also helps children to determine 'what kind of objects do X', facilitating categorisation into groups of words or 'slots' that can be used in particular positions in sentences. Children combine these newly identified 'slots' and 'frames' to produce lexically specific constructions (Gomez, 2002; Tomasello, 2003).

During each BEST session, the same words and constructions are modelled many times, both by the practitioners and other children in the group. The input is always presented without the practitioner delivering praise or providing additional verbal input, so that lexical forms and structures can be clearly mapped onto referents without the competing processing demands of other spoken language.

Many high quality exposures to the input are thought to be necessary for children with language difficulties to effectively map lexical forms onto semantic and pragmatic functions (Fey et al., 2003; Gray, 2003, 2004; Riches et al., 2005). Tomasello (2003) argues that multiple exposures help facilitate schematisation which is necessary to produce lexically specific constructions (see 1.4.2), while Leonard et al. (2019) demonstrates that opportunities for continued retrieval during word learning activities also enhance retention compared to situations where this is not facilitated. In this way, both the input and output phases may support mapping through repeated exposure and opportunities for retrieval.

It has also been demonstrated that repetition of grammatical, well-formed sentences can facilitate better sentence recall in sentence repetition tasks than other linguistic properties such as frequency (Polišenská et al., 2015).

1.5.1.4.3 Multiple Exposures to Analogous Constructions

The final stage of Tomasello's model targeted in BEST is the development of abstract constructions which is underpinned by the cognitive mechanism of analogy.

BEST supports children to harness the cognitive mechanism of analogy by providing multiple exposures to analogous constructions through the manipulated input described above. Analogy allows children to produce new sentences using the functional categories identified across sentences.

The input and output phases of BEST (see Table 1. 4 above), provide consistent variation and contrast between sentences. The predicate argument structure remains the same, meanwhile the grammatical morphemes facilitate bootstrapping of commonalities across the contrasts through identification of functional categories e.g. agent, verb, patient (Gentner & Markman, 1995, 1997; Gentner & Medina, 1998). In particular the use of sign, which is discussed in detail in 1.6, is hypothesised to support this element because it provides an additional visual representation to the oral language highlighting the morphemes, as grammatical morphemes pose a particular difficulty for children with language difficulties (Leonard, 2003; Leonard, 2007; Rice et al., 2004). Children are exposed multiple times to each contrast, and this consistency is hypothesised to support them to develop abstract semantic representations of the sentence structures (e.g. 'agent + action + patient' analogises to 'the boy is eating the lolly'). During this process children analogise from what they hear in the input to produce abstract representations e.g. 'the X is Ying the Z' in the target sentences (Childers & Tomasello, 2001; Gentner & Medina, 1998).

Unlike in some interventions, mastery of one sentence structure is not necessary for the child to move on to the next; rather, the structures are rotated over the intervention sessions. Van Horne et al. (2017, 2018) provide evidence that starting with more complex exemplar structures can be beneficial for accuracy on trained and untrained items and for accurate generalisation of trained past tense morphemes. Distributed exposure is also argued to increase the learning of abstract syntactic constructions compared to massed exposure (for example if items were all presented at the same time, rather than over a number of sessions), or learning in control groups (Ambridge et al., 2006). The distributed learning effect finding is also robust for other cognitive domains (see Janiszewski et al. 2003 for a review).

1.5.1.4.4 Promotion of General Cognitive Processes: Mapping and Retention

Mapping and retention are further key processes for the stages of language development outlined above. Mapping a lexical form or construction to its meaning relies on children being able to acquire new meaning-construction links quickly, and is essential for learning words and constructions and their meanings (Hirsh-Pasek et al., 2000). Future use of the newly learned constructions relies on the child's ability to retain the item and its meaning in their long-term memory (Ambridge et al., 2006; Janiszewski et al., 2003). For children with language difficulties, mapping and retention can be more challenging, due to underlying processing and memory deficits posited in 1.4.4. BEST therefore provides many examples of each construction in use which has been shown to facilitate mapping (Fey et al., 2003; Gray, 2003, 2004; Riches et al., 2005). The BEST input is presented through distributed exposure over 16 sessions (discussed above), which has also been shown to promote retention (Riches et al., 2005).

1.5.1.5 Summary

BEST aims to model how two, three, and four clause element sentences are formulated through repeated systematically varied and distributed exposure to the input, harnessing the cognitive mechanisms outlined in the section above. BEST aims to support the key skills implicated in language learning: 1) the ability to attend to joint attentional frames with others; 2) to understand the communicative intentions of others; 3) to implement ‘role reversal imitation’- a form of social, cultural learning; 4) pattern-finding; and 5) statistical learning (see section one for a discussion of each of these skills).

The intervention also utilises tactile and visual reinforcement with toys and a signing system. As is discussed in detail in the following section, BEST was designed to be delivered with a signing system which is deployed to represent the sentence content and syntactic structures visually (McKean et al., 2013). The hypothesised benefits of this are threefold. Firstly, McKean et al. argued that representing vocabulary with sign would support semantic mapping and consequently facilitate learning of predicate argument structure. They further argued that signing morphological items (e.g. ‘the’, ‘a’ and ‘-ing’) supports the mapping of the morphological frame, which supports children to bootstrap abstraction of semantic roles, which is outlined above. Finally, they also suggested that signing morphological items would support children to use and understand grammatical markers by drawing children’s attention to the grammatical ‘frame’ of the sentence (e.g The X is Y-ing the Z), and the ‘slots’ which are inserted (see above for more detail).

BEST has undergone initial piloting work in a series of Undergraduate student dissertations (see 2.2). The study did not address the role of sign, however it did identify sign as a potential area of concern with regards to acceptability to parents, for reasons which are discussed in 1.6.5 below.

The next section reviews the current evidence regarding the potential benefits of sign, specifically as an adjunct to BEST.

1.6 The Role of Sign as an Adjunct to Language Intervention

As described in 1.5.1.5, BEST uses a sign system alongside verbal input as an additional cue to support the use of specific cognitive mechanisms thought to underpin language learning. There is some evidence to suggest that sign and gesture can support children's oral language development, but very little research has been carried out into the effect of signing systems on language acquisition and specifically as an adjunct to language intervention.

Signing systems are not without their controversies; there is anecdotal evidence, which can be observed on online parenting websites, suggesting parents are concerned that the use of a signing system may delay oral language. Unfortunately, there is little robust evidence to counter such claims. It is therefore important to determine whether signing systems may support language development when implemented as part of a usage-based intervention and, if so, increase the evidence base pertaining to the role of sign as an adjunct to language intervention for pre-school children with language difficulties. One of the key aims of this thesis is to establish the role of signing as an adjunct to BEST and to contribute to the broader understanding of sign so that parents can accept signing as part of BEST if it is found to be beneficial (Capone & Mcgregor, 2004).

1.6.1 Differentiating Between Signing Systems and Iconic Gesture

The terms *signing systems, gesture* and *Sign Languages* cover broad topics within the speech, language and hearing literature. The focus of this study is on formal *signing systems*, which are symbolic communication systems used to support the learning and understanding of spoken language and are used in the same order as the spoken words.

Signing Systems consist of standardised hand movements and shapes designed to be used alongside speech to provide a symbolic representation of what is being said (Bickford-Smith et al., 2005; Grove & Walker, 1990; Lal, 2010; Poncelas & Murphy, 2007; Rowe, 1981; Volpato et al., 1986) (also see

Abbott & Lucey, 2005 for a review of Signing Systems; Gustason, 1990 for a review of Signing Exact English (SEE); Seal & De Paolis, 2014 for a discussion of Baby Sign which is not relevant here).

Signing systems, as opposed to Sign Languages e.g. British Sign Language (BSL), do not have their own grammar and are not formally recognised languages (Perniss et al., 2015); they are instead used to support spoken communication, often with children who are acquiring language or people who have developmental or learning disabilities.

Makaton, developed in the 1970s, is one of the most popular signing systems in use today (Abbott & Lucey, 2005). Makaton signs are derived from BSL but are used alongside spoken English and generally only key elements of the sentence are signed. Makaton is widely used in mainstream and special schools and with children and adults with language impairments and learning disabilities (Abbott & Lucey, 2005; Bickford-Smith et al., 2005; Grove & Walker, 1990; Lal, 2010; Poncelas & Murphy, 2007). Despite this, little formal research has been carried out into the efficacy of Makaton for supporting children with language impairments (Law et al., 2012). The research that has been completed mainly focuses on children with Autism Spectrum Disorders (e.g. Lal, 2010).

Paget Gorman Signed Speech (PGSS) (Rowe, 1981) is a lesser known system used to represent spoken language. Like Makaton it has standardised signs for vocabulary. Unlike Makaton, it provides more scope for representing morphological and syntactic elements of words (for example, the Makaton vocabulary has no sign for morphological verb endings such as the present tense regular verb ending '-ing'; one must sign the individual letters 'i', 'n', 'g' which would not be practicable in typical use). PGSS allows more syntactic information to be communicated via individual signs (for example, it has a separate sign for the verb infinitive and the present tense '-ing' ending).

Similarly to signing systems, gesture is described as the movement made by the arms and hands which is intrinsically linked with communication (Kendon, 1980, 2004). Iconic gesture is one category of gesture which conveys information about a physical aspect of the subject (McNeill, 1992; Wagner et al., 2014). In contrast to signing systems, Novack et al. (2015), argue that iconic gesture is produced spontaneously and as such is not in any way standardised or replicable. Despite this, many studies examining the role of gesture deploy 'standardised iconic gestures' as part of the experimental paradigms. Gesture (e.g. Hodges et al., 2018; Lüke et al., 2020) and sign (van Berkel-van Hoof et al., 2019) therefore often describe what is arguably the same type of cue consisting of repeated, standardised signs (for example the flapping of the hands to represent a bird (Hodges et al., 2018)). Indeed Kendon (2004) highlights the 'common ground' between gesture and sign, noting that unstable gestures that are not shared with others are rapidly transformed into symbolic signs. Further, there appears to be more literature examining gesture than explicit signing systems. For these reasons the review below includes studies of symbolic gesture as well as signing systems.

1.6.2 The Role of Iconic Gesture and Signing Systems on Oral Language Development Typically Developing Children

Gesture and language are highly associated (Iverson et al., 1999; Kendon, 1980; Lüke et al., 2020; Wagner et al., 2014). It has been demonstrated that children as young as three years old can interpret and benefit from iconic gestures with some success (Marentette & Nicoladis, 2011; Namy et al., 2004). Novack et al. (2015) found children as young as two years old can access semantic information about how to solve a problem from iconic gesture. Actional iconic gesture (i.e. gesture representing an action) appears to be interpreted earlier than gestures referring to characteristic attributes, suggesting that actional iconic gestures are easier for children comprehend at an earlier age (Hodges et al., 2018).

For typically developing children the usefulness of gesture in interpreting meaning appears to diminish as the child develops, in favour of using lexical cues (Capone & Mcgregor, 2004). At around 18 months, children spontaneously interpret gesture as they do lexical forms for object naming. By 26 months however, children rely more heavily on spoken cues and appear to have surpassed the need to rely on gestural information (Namy & Waxman, 1998). Botting et al. (2010) found that typically developing children relied more heavily upon spoken cues when comprehension broke down than did their peers with SLI. When compared with non-iconic gesture such as pointing, iconic gesture improved word naming, comprehension and linking gesture to lexical forms in typically developing children, however it did not increase the encoding of semantic information (Vogt & Kauschke, 2017a, 2017b).

Ting et al. (2012) found that TD pre-linguistic infants' ability to segment speech in the presence of background noise was impaired by the simultaneous presentation of spoken English and SEE (Signed Exact English, see Gustason (1990). Ting et al. suggest that this may be a result of competition between sign and oral language on limited processing resources within the child's working memory. This finding also aligns with the limited processing capacity theories discussed in section three. Other studies have found no detrimental effect, but also no benefit of sign and spoken English presented together (Seal & De Paolis, 2014).

These findings suggest that young, typically developing children may utilise cues encoded in gesture to aid comprehension at early ages but seem to outgrow the need to rely on such cues by around 2;5 years and instead rely on lexical cues in the input, although this evidence remains inconclusive.

1.6.3 The Role of Iconic Gesture and Signing Systems on Oral Language Development in Children with Language Difficulties

Studies have shown that children with language difficulties benefit more from gesture, particularly iconic gesture, than typically developing peers, and that language is more closely related to gesture for language impaired children (Botting et al., 2010; Marton, 2009). In the study discussed above, Botting et al. (2010) found that children with language impairment utilised gestural cues to scaffold meaning when comprehension broke down.

As is the case for typically developing children, iconic gesture has also been shown to improve word naming and comprehension in language impaired groups, when compared to non-iconic gesture (Vogt & Kauschke, 2017a). Iconic gesture promoted stronger links to lexical forms in, and also promoted richer encoding of, semantic information for children with SLI (Vogt & Kauschke, 2017b).

No effect of word class (nouns and verbs) was found, suggesting this finding holds across classes regardless of actionality, unlike for TD children who benefit more from more actional gestures (Hodges et al., 2018). A further recent single-cases study of four late-talkers suggested that co-speech shape gesture for object names facilitated better name learning (Capone et al., 2020). Recent research by Lüke et al. (2020) suggests that observed differences between TD and language impaired children are the result of a time delay, while their underlying integrated speech-gesture communication systems remain similar. This finding supports the ideas presented above that TD children 'outgrow' their reliance on gestural cues in favour of lexical cues while children with language difficulties continue to benefit from these cues for longer.

Positive effects of sign for children with DLD continue to be observed into the older school years. Van Berkel-van Hoof et al. (2019) found that children with DLD aged between seven and nine-years-old benefited from sign-supported speech on measures of vocabulary, whilst no benefits were demonstrated for TD children.

It may be that children with DLD in particular benefit from non-verbal information encoded in iconic gesture, perhaps due to their less sophisticated interpretation of spoken cues relative to age peers. The evidence suggests that children with language difficulties continue to exploit cues from gestures for a longer time than their typically developing peers. It is not clear when or if they outgrow this benefit. The role of gesture should therefore be considered as a possible tool for diagnosis and intervention for children with language difficulties (Capone & Mcgregor, 2004).

1.6.4 Signing Systems as an Adjunct to Language Intervention in the Early Years

The What Works database (Law et al., 2012) was searched for interventions specifically consisting of or incorporating signing systems using the same criteria in section 1.3.1 above. Only two interventions included the use of sign to support language difficulties: 1) Makaton (Bickford-Smith et al., 2005; Grove & Walker, 1990; Lal, 2010; Poncelas & Murphy, 2007); and 2) Visual Approaches to Support Speech and Language (Archibald & Gathercole, 2006; Lal & Bali, 2007). The Hanen It Takes

Two to Talk intervention (Girolametto & Weitzman, 2006) also places an importance on non-verbal communication including gesture, however does not explicitly involve a signing system and refers only to naturalistic gesture.

As outlined above, Makaton is a popular signing system used in schools and by SLTs. Despite its popularity, few studies have sought to evaluate its efficacy and underlying mechanisms, instead focusing on clinician and teacher attitudes towards sign (Abbott & Lucey, 2005).

The limited research outlined above suggests that the use of a signing system alongside spoken language could be beneficial to children with language difficulties as an adjunct to language intervention. Further research in this area would be beneficial both to establish whether children's language outcomes could be improved by the inclusion of sign, and also to test the underlying theoretical explanations of language difficulties discussed in section 1.4.

Finding that sign is beneficial as an adjunct to language intervention would support the ideas of Tomasello (2003) outlined in 1.5.1.4. This would indicate that sign highlights the morphological 'frames' implicated in the promotion of analogy (Tomasello, 2003). It could be argued that establishing a joint attentional frame with a sign referring to the spoken word at the centre of it also provides the child with two complementing exposures to the same input, thus harnessing mechanisms underpinning language acquisition.

1.6.5 The Use of Sign as an Adjunct to BEST

The evidence outlined above suggests that the use of co-speech iconic gesture or sign may support young children to learn the meanings of new words (Capone & Mcgregor, 2004; Capone et al., 2020; Novack et al., 2015; Shimpi & Huttenlocher, 2007; Vogt & Kauschke, 2017a, 2017b). However, no previous research has evaluated the role of signing systems such as Makaton and Paget Gorman Signed Speech as an adjunct to language intervention in the early years. Previous research also focuses mainly on the learning of words and their meanings (e.g. cat, jump, shoe) (Capone &

Mcgregor, 2004; McGregor et al., 2002) and has not examined other important facets of acquisition such as morphosyntax.

In BEST, a signing system is deployed alongside oral language which represents both the sentence content and syntactic structure in the visual domain. BEST was originally designed to use Paget Gorman Signed Speech (PGSS), although this study employed a combination of PGSS and Makaton due to children's familiarity with the Makaton vocabulary (see chapter two for the full intervention methodology). When designing BEST, McKean et al. (2013) hypothesised that signing vocabulary may support mapping of semantics and therefore the learning of predicate argument structure. They also hypothesised that signing morphological items (e.g. 'the', 'a' and '-ing') could support the mapping of the morphological frame (e.g. The X is Y-ing the Z), which could in turn support children with bootstrapping abstraction of semantic roles. Both of these may support language development, particularly in the domains of sentence content and sentence morphology.

The hypothesised benefit of sign as an 'active ingredient' of BEST has not yet been empirically tested. Testing the role of sign is important given that there is a potential benefit suggested by the sign and gesture research. There are also potential disadvantages to the use of sign; it may be time consuming and difficult to learn and practitioners who are inexperienced with signing systems could inadvertently provide inconsistent signs, thus affecting treatment fidelity. It is therefore necessary to establish whether signing is an 'active ingredient' of BEST, or whether the intervention can be delivered equally effectively without sign, so that training and delivery of BEST is time and cost effective.

In addition, the role of sign on morphology learning is specifically addressed. Most signing systems (e.g. Makaton) do not specifically support morphological constructions, however BEST uses morphological signs from PGSS, which could be important for promoting morphological development. This research will contribute to the literature on the usefulness of signed speech for language and morphology development. The role of cognitive mechanisms that may be supported by signed speech under the constructivist account of language acquisition is considered.

1.6.6 Controversy Surrounding the Use of Signing Systems

As mentioned at the beginning of this section, there is some controversy surrounding the use of sign with non d/Deaf children (terms which are used to represent people with hearing impairments, and those who are pre-linguistically Deaf respectively). Scepticism about the use of signing emerged in the 1980s, as some practitioners were unsure whether signing might delay typical language acquisition (see Sheehy & Rolph, 2004; Sheehy & Duffy, 2009). Teachers surveyed in the 1980s held mixed views towards signing, with some stating that Makaton was detrimental to speech, although others felt it was helpful (see Sheehy & Duffy, 2009). A more recent survey of providers of special education found that teachers' attitudes towards the use of sign have improved greatly (Abbott & Lucey, 2005).

There is no known research addressing current parental attitudes towards the use of sign, however searches of online parenting forums suggest that some parents and carers have concerns that signing may delay the onset or development of oral language, especially for children who do not have a diagnosed learning difficulty. Some researchers have sought to promote the message that gesture and sign enhances, rather than hinders, language development, and practitioners' attitudes towards signing are positive (Abbott & Lucey, 2005), but further work is clearly required, firstly to determine the efficacy of sign and then to promote its use to parents and carers if it is found to have a positive effect.

1.6.7 Use of Sign: Conclusion

The research outlined above suggests that children with DLD may benefit more than their typically developing peers from sign or gesture presented alongside oral language in vocabulary and comprehension outcomes. Research into standardised signing systems is therefore warranted since these present one viable method for incorporating sign into interactions and language intervention. Despite its potential use, no known work has specifically addressed the role of a signing system as an adjunct to language intervention or examined its theoretical underpinnings with regards to promoting bootstrapping of abstraction and mapping. Understanding the nature of the mechanisms underpinning sign is important not only due to its potential clinical implications, but also due to the insight into language acquisition that it may provide. This must also be interpreted in light of the time and costs involved in training practitioners and teaching staff to deliver signing systems effectively.

1.7 Frameworks for Evaluating and Reporting Complex Language Interventions

We have seen from the review of existing interventions that there is a paucity of robustly evaluated language interventions for pre-school children with severe language difficulties (Law et al., 2010). Compounding the problem, the reporting of trials is often inadequate (e.g. Eldridge et al., 2016; Schulz et al., 2010), leading to difficulty in interpreting treatment effects (Moher et al., 2010). It is therefore essential that interventions are evaluated to the appropriate standard (Craig et al., 2008; Eldridge, Chan et al., 2016; Eldridge et al., 2016) and that studies are reported transparently so that the reader can make appropriate judgements about possible biases affecting outcomes (Des Jarlais et al., 2004).

1.7.1 Complex Language Interventions

Whilst past models of intervention evaluation have focused mainly on simple biomedical interventions, specifically drug trials, there is now a greater acknowledgement of the need to rigorously evaluate non-pharmacological complex interventions. These are interventions consisting of more than one key component, which require differing approaches to evaluation than simple interventions due to their multiple components (Campbell et al., 2000; Craig et al., 2008).

1.7.1.1 The Need to Evaluate Complex Interventions

With the emerging movement for evidence-based practice within the Speech and Language Sciences (Ebbels et al., 2019; Gladfelter et al., 2011; Klee et al., 2009; Stringer, 2010a, 2010b, 2010c, 2010d, 2010e, 2010f, 2010g; Zipoli & Kennedy, 2005), it is important that researchers undertake and report intervention studies in order to increase the evidence base for practitioners and researchers.

Randomised controlled trials are widely regarded as the “gold standard” of intervention research in terms of level and strength of evidence (Schulz et al., 2010). Whilst this method is essential for establishing the effects of an intervention on a wide scale, there is an increasing understanding that, for complex interventions, moving too soon to an RCT can be problematic. The importance of the piloting and feasibility work, which is integral to the success of the trial, is often overlooked, leading to poorly conducted trials and erroneous conclusions being drawn. Ultimately such studies waste time and resources (Craig et al., 2008; Levati et al., 2016). It is also unethical to waste finite resources as well as participants’ time on poorly conducted or misinformed trials.

As discussed above, careful development, piloting and feasibility work is needed prior to a definitive trial. If the piloting and feasibility work suggests intervention efficacy, there is then a need to determine whether the intervention would work in a ‘real world’ setting and whether it is cost effective to deliver. Of the interventions reviewed in section two, only the Nuffield Early Language Intervention is explicitly described as undergoing all phases of evaluation (Bowyer-Crane et al., 2008; Fricke et al., 2013, 2017; Law et al., 2017). It is possible that the majority of available interventions have not undergone the full process of evaluation; certainly, early phases of development are rarely reported.

1.7.1.2 Current Difficulties Faced by Researchers Evaluating Complex Interventions

Researchers face two problems when attempting to evaluate complex language interventions. Firstly, there is a lack of consensus on the terminology used to define early stages of trial evaluations, namely the piloting and feasibility phases. Since differing sets of guidelines use differing terminology, it is essential that these terms are properly defined and used consistently in the literature.

Secondly, there are a number of contrasting frameworks for the evaluation of complex interventions available to researchers. Fey and Finestack (2009) argue that this may underpin the lack of robust intervention trials, as well as having replication, theoretical, methodological, empirical and clinical implications. This raises the question of which set of guidelines researchers should follow, particularly as the most widely cited of these, the Medical Research Council (MRC) Guidelines for Developing and Evaluating Complex Interventions (Craig et al., 2008), are criticised for their brevity (Eldridge et al., 2016).

Seeking to address these issues, this section critically evaluates the available frameworks for the evaluation of complex interventions in order to: 1) determine the appropriate terminology to describe the present study; 2) highlight the contrasting guidelines available to researchers; and 3) determine the most appropriate method of evaluation for the present study. This will ensure that the study not only addresses the appropriate questions, but also that the present research can inform latter stages of the evaluation process.

1.7.2 Terminology Describing the Early Stages of Intervention Evaluations

As discussed above, there are two pairs of analogous terms which must be defined to ensure that research can be interpreted in line with other similar studies. These are *efficacy and effectiveness*, which pose little controversy, and *piloting and feasibility* which are more widely debated. The terms and their meanings as employed in the rest of this thesis are defined below.

1.7.2.1 Efficacy and Effectiveness

Robey and Schultz (1998) describe two distinct clinical trial outcomes: efficacy and effectiveness. According to their 1988 aphasiology model, efficacy is the “magnitude of the therapeutic effect” (p.

790) under ideal conditions. They define effectiveness as the "realised value of an efficacious treatment in a typical field or context" e.g. how well the intervention works in a clinical context under normal conditions (p. 791). These definitions are consistent within the literature (Flay, 1986; Glasgow et al., 1999, 2003). Importantly, Robey and Schultz note that efficacy is a relative rather than absolute measure, since aphasia is not 'cured' in the way that some other biomedical conditions are.

The above definitions proposed by Robey & Schultz are used in the present study since the model they are derived from was the first SLT specific model to differentiate between the two study types in detail.

1.7.2.2 Piloting and Feasibility

Even amongst key sources there is a lack of consensus on the definitions of pilot and feasibility studies; the MRC state that "A pilot study need not be a 'scale model' of the planned mainstage evaluation" (Craig et al., 2008, pp.10), while the NIHR definition appears to contradict this: "A version of the main study that is run in miniature to test whether the components of the main study can all work together" (NIHR, 2018). There is also a lack of consensus on whether pilot and feasibility trials were separate types of study. The terms 'pilot' and 'feasibility' are used interchangeably to define a broad range of scoping studies (Lancaster, 2015) and to describe both stand-alone trials and 'internal' trials conducted as part of a larger-scale RCT (Eldridge et al., 2016). The NIHR provide a separate definition of feasibility studies: "Pieces of research done before a main study in order to answer the question "Can this study be done?. They are used to estimate important parameters that are needed to design the main study..." (NIHR, 2018), however the MRC do not clearly define feasibility studies independently of pilot studies (see Appendix 3 for full MRC and NIHR definitions of piloting and feasibility).

Eldridge et al. (2016) addressed this gap in the literature by introducing a conceptual framework for defining pilot and feasibility studies. Through a Delphi Method study, they reached consensus on the following definitions: "a feasibility study asks whether something can be done, should we proceed with it, and if so, how". "...a pilot study is a study in which a future study or part of a future study, is conducted on a smaller scale to ask the question whether something can be done, should we proceed with it, and if so, how" (pp. 8). These definitions are used for the present study since they were produced via a vigorous consensus method unlike other definitions.

Until recently, pilot and feasibility trials were rarely published or registered, although the introduction of dedicated journals (Eldridge et al., 2016; Lancaster, 2015; Lancaster et al. 2004) and pre-registration of clinical trials (Nosek et al., 2018) is beginning to address this.

1.7.3 Frameworks for the Evaluation of Complex Interventions

Since the 1980s a number of frameworks for evaluating complex interventions have been proposed (for a review pre-dating these see Robey & Schultz, 1998). Several systematic reviews have sought to describe the broad range of available frameworks for developing and evaluating complex interventions (Colquhoun et al., 2017; Levati et al., 2016; O'Cathain et al., 2019). There is an acknowledgement of the large number of frameworks for evaluation and the difficulties faced by researchers when attempting to select the appropriate framework for their study (Levati et al., 2016).

Of key interest to the present research is The Medical Research Council (MRC) Framework for the development and evaluation of complex health interventions (Craig et al., 2008) since it is the most popular and well regarded framework currently available in the literature. Despite this, it is not without its issues (Eldridgeet al., 2016). The MRC guidelines are discussed below, followed by the other notable frameworks which may clarify unclear aspects of the MRC guidelines.

1.7.3.1 Medical Research Council Developing and Evaluating Complex Interventions

The Medical Research Council (MRC) provide a major framework for the development and evaluation of complex health interventions (Craig et al., 2008). The framework comprises four stages: 1) development; 2) feasibility/piloting; 3) evaluation; and 4) implementation. The stages are non-linear but considering them as distinct stages can help to contextualise the overall process. The authors argue that each stage is necessary for the overall success of a complex intervention.

The development stage of the MRC framework involves identifying the current evidence base and appropriate underlying theory. Process and outcomes are modelled to identify the anticipated outcomes, effect sizes, and potential costs of the intervention to indicate whether a full-scale trial is feasible and justified.

The feasibility and piloting stage focuses on defining research questions, testing procedures and determining sample size and point interval effect sizes. This stage is important for anticipating methodological problems that could occur in larger and more costly evaluation trials. In particular this phase aims to inform design so as to avoid issues with sample size, recruitment and retention, acceptability and compliance in subsequent trials. Changes to the intervention protocol are implemented following careful reflection by the Author throughout the study. Multiple studies may be warranted at this stage to systematically improve the design.

The evaluation stage of the framework relies on appropriate methodological choices and design implementation for the research to be effective. This stage assesses the effectiveness of the intervention, including its ‘active ingredients’, treatment fidelity, understanding processes and cost-effectiveness.

The implementation stage involves dissemination of findings to stakeholders, service holders, academic communities and other relevant bodies. Service, monitoring, and long-term follow up are also included in the implementation stage.

The nonlinear nature of complex intervention evaluation is evident from the position of the current research within the MRC Framework, and elements from multiple stages inform the design and research questions for the present study.

Although it outlines a much needed process, the MRC framework has been criticised for its brevity (Eldridge et al., 2016). It lacks sufficient detailed guidance about how to conduct each stage of the process that it describes and does not make reference to efficacy and effectiveness which are two important clinical outcomes for intervention research (Robey & Schultz, 1998). The most problematic element of the MRC guidelines is the lack of reference to efficacy. It is important that efficacy is assessed prior to effectiveness (Eldridge 2016) however the MRC guidelines do not include this as a specific step.

It is therefore assumed that efficacy falls under the 'piloting and feasibility' stage of the guidelines, since effectiveness falls under the subsequent 'evaluation' phase.

Within the context of the present study, the framework also does not address issues specific to SLT research such as appropriate effect sizes and outcomes, as well as the heterogeneity of symptomology. For this reason, it is important to look to alternative frameworks to gain clarity on aspects of conducting trials. These are divided into two categories: frameworks for the evaluation of complex interventions in fields other than Speech and Language Therapy and frameworks that specifically speak to SLT research.

At the time of writing an updated version of the MRC guidance is anticipated, so criticisms may be addressed in the subsequent version.

1.7.3.2 Other Frameworks for the Evaluation of Non-SLT Complex Health Interventions

The Scientific Approach to Cancer Control is an early model for the evaluation of cancer treatments designed by the National Cancer Institute (NCI) (Greenwald & Cullen, 1984). This systematic approach was introduced to ensure that sufficient initial research was conducted before wide-scale interventions were implemented, ensuring optimal outcomes for public health, scientific rigour and cost-effectiveness. The original model consisted of five phases: 1) hypothesis development; 2) methods development; 3) controlled intervention trials; 4) defined population studies; and 5) demonstration and implementation.

Phase one of this model involves conducting small-scale, small-n studies to develop hypotheses, establish that the treatment is safe, and determine whether further research is warranted. *Phase two* involves development of study procedures, refining the treatment protocol and optimising dosage and the assessment battery. Efficacy is also partially tested in phase two. *Phase three* studies test the hypotheses which were developed over phases one and two, using larger sample sizes to robustly test efficacy. If efficacy is determined, *phase four* of the model tests efficacy for specific subpopulations, and begins to assess effectiveness, testing in 'real-world' scenarios such as service-delivery settings. Finally, *phase five* continues effectiveness testing in the appropriate clinical setting. Following this, cost-effectiveness is also ascertained.

Flay (1986) extended Greenwald & Cullen's (1984) model to produce 'Efficacy and Effectiveness Trials (and Other Phases of Research) in the Development of Health Promotion Programs'. This model was more comprehensive and applicable to other areas of health research. The model included the following stages: 1) basic research; 2) hypothesis development; 3) pilot applied research; 4) prototype studies; 5) efficacy trials; 6) treatment effectiveness trials; 7) implementation effectiveness trials; and 8) demonstration studies. Flay placed an emphasis on the initial scoping phase, noting the importance of carrying out initial work before larger efficacy and effectiveness trials, which is widely recognised as essential today (Craig et al., 2008).

Several important advances were made in Flay's (1986) model, particularly the clearer distinction between efficacy and effectiveness studies and the applicability of the framework to other domains of health research. Despite this, the early phases are inadequately defined and therefore difficult to distinguish from one another. As is the case for the MRC framework, these models also lack sufficient information about undertaking each stage of the process.

1.7.3.3 Models for the Evaluation of Speech and Language Therapy Interventions

Adapted from the NCI Cancer Control Research Phases (Greenwald & Cullen, 1984), Robey & Schultz' (1998) model is specifically adapted for SLT interventions. This refined model was the first to distinguish between efficacy and effectiveness clinical outcomes from the communication disorder perspective. This notion of 'efficacy-then-effectiveness testing' argues that efficacy testing (under ideal conditions) justifies effectiveness testing (in 'real-world' settings) (Robey & Schultz, 1998, p. 791). The retention of many of the facets of the NCI model also mean that evaluation outcomes are easily communicated to healthcare professionals from other fields.

Whilst the model considers outcomes in the context of Speech and Language interventions and omits irrelevant oncological considerations from the NCI model such as toxicity and pharmacology, it does not add a great deal to understanding of SLT intervention development. The model itself does not address many of the issues specific to language research such as the heterogenous presentation of a language disorder in different patients and selecting appropriate outcomes to measure progress. Additionally, it does not provide sufficient detail to implement the stages reliably. Ultimately it does not differ greatly from the NCI model in what it offers the present study.

Another more in-depth model is The five-phase model from Fey & Finestack (2009). This model is the most specific framework for evaluating language interventions. The five phases: 1) pre-trial studies; 2) feasibility studies; 3) early efficacy studies; 4) later efficacy studies; and 5) effectiveness studies do not differ greatly from previous frameworks, however the authors provide a comprehensive commentary of how each is specifically applied to language intervention research.

Pre-trial studies often consist of correlational or observational non-clinical trials. They are designed to understand the nature of specific language difficulties and generate hypotheses about methods for addressing these.

Feasibility studies are early studies used to test hypotheses developed in the pre-trial studies and to establish the feasibility of the proposed interventions. Questions addressed at this stage might include whether the intervention is acceptable to children and parents and what is required to ensure successful delivery. These studies may be carried out with typically developing children and may not use control groups, although there is a high degree of variability at this stage and some feasibility studies may also be considered clinical trials.

Early efficacy studies compare a treatment or set of intervention procedures with another treatment or no treatment. Designs may be experimental or quasi-experimental and steps may be taken to reduce costs, although this will often result in reduced generalisability. Early efficacy studies should aim to describe the intervention and assess treatment fidelity, outcome measures and internal validity.

Later efficacy studies assess whether there is a cause and effect relationship between the intervention and positive outcomes for children's language. To answer this question, studies aim to compare the recipients of intervention with control groups or other, existing interventions. These studies should be generalisable to wider populations. Outcomes should be clinically meaningful progress on language outcomes, and studies may seek to address whether results generalise to non-target outcomes.

The final stage of evaluation, *effectiveness studies*, should be completed only when appropriate efficacy studies have demonstrated that further research is warranted. Such studies evaluate whether the same efficacious outcomes found in early and later efficacy studies are still present when the intervention is conducted under typical clinical conditions. Effectiveness studies can also consider cost-benefits of the intervention. Since clients involved in effectiveness research are likely to have a broader range of difficulties the studies may have to account for more than one set of goals in some circumstances.

This model is specifically adapted to language intervention research and considers questions that should be addressed at each stage of the research process in more depth than other models including Robey & Schultz's (1998) model. Whilst this model has many merits, the descriptions of each phase do not align well with previously outlined definitions of efficacy and effectiveness. For example, the third phase, early efficacy studies, describes assessing treatment fidelity, outcome measures and internal validity, which would fall under piloting and feasibility work in most other accounts.

1.7.4 Situating the Present Research within the Frameworks for Evaluating Complex Interventions

The above review of frameworks for the evaluation of complex interventions demonstrates the variation which is described as a key issue facing researchers at the beginning of this section. In accordance with the MRC guidelines, the present study is situated between the *feasibility and piloting stage* and the *evaluation stage*, since some initial pilot and feasibility work had been carried out to inform the present study (see chapter two) and because the study aimed to capture learning about evaluating BEST to inform a future large-scale evaluation.

According to Robey & Schultz' model, the present research evaluates efficacy outcomes in phases two and three of the model, as the primary aims of this study are to test hypotheses, refine the treatment protocol, optimise dosage, and test efficacy. The present research initially sought to address specific subpopulations, as well as to determine whether underlying baseline characteristics affect response to intervention, which would fall under phase four.

In line with Fey and Finestack's (2009) model, the present research sits between early and later efficacy studies. The present research is carried out under ideal conditions and seeks to refine the intervention and assess treatment fidelity and outcome measures (early efficacy) whilst also seeking to ascertain a causal relationship between intervention and outcomes, compared to a treatment as usual condition (later efficacy).

The overlap between phases in these models supports the non-linear nature of intervention evaluation described by Craig et al. (2008) and also suggests that there is a high level of variability between frameworks that should be addressed by subsequent research (Levati et al., 2016)

Although criticised for its brevity, the MRC framework does not suffer from the lack of clarity between phases identified in the other frameworks. The present study therefore follows the MRC guidelines for developing and evaluating complex interventions to ensure transparency of the research.

This study is situated between the *piloting and feasibility* and *evaluation* stages of the MRC framework, with a focus on defining research questions, testing procedures, anticipating methodological problems and determining sample size and point interval effect sizes. The study also seeks to evaluate the 'active ingredients' and ensure high treatment fidelity. The work also reports barriers encountered during the study and makes recommendations for subsequent evaluations based on these.

Despite this, important clarifications can still be made by observing the other frameworks, particularly with regards to efficacy and effectiveness outcomes. The study therefore seeks to measure efficacy under ideal study conditions (Robey & Schultz, 1998) and could be considered a *later efficacy study* under Fey and Finestack's model.

1.8 Research Questions

We saw in the last section that following the development of an intervention, it must undergo a rigorous process of pilot and feasibility testing, followed by evaluation (Craig et al., 2008). In the early stages of evaluating an intervention it is necessary to conduct small scale efficacy studies to identify potential therapeutic effects and estimate effect sizes (Craig et al., 2008; Fey & Finestack, 2009). The literature reviewed in this chapter emphasises the need for theoretically motivated language interventions that have undergone robust evaluation. Building Early Sentences Therapy is underpinned by usage-based theory and has undergone initial pilot and feasibility work, which is discussed in chapter two.

In line with the MRC Guidelines (Craig et al., 2008), this study takes the next step in the evaluation of BEST. Through a quasi-experimental matched group design this research aims to estimate the presence and magnitude of efficacy of BEST compared to treatment as usual (TAU). It also aims to test the role of sign as an 'active ingredient' and to produce evidence-based recommendations for a future fully powered trial.

The research questions are divided into two strands: 1) examining efficacy; and 2) informing a future fully powered trial. The quantitative results of strand one are presented in chapter three. The qualitative results of strand two are presented in chapter four.

The research questions are as follows:

1.8.1 Strand 1: Determining the efficacy of Building Early Sentences Therapy

1A) Is an intervention underpinned by usage-based theory (BEST, with and without additional sign) more effective than treatment as usual (TAU) for improving language and associated functional communication outcomes in pre-school children with language difficulties?

1B) Is signing a necessary 'active ingredient' of BEST, and does it differentially affect progress on different language goals (sentence structure and morphology)?

1C) What are the estimated effect sizes across outcomes for BEST with and without Sign, relative to Treatment as Usual?

1.8.2 Strand 2: Informing a future fully powered trial

2) What can we learn regarding the study processes and acceptability which will inform further development of BEST and contribute to knowledge about school-based speech and language intervention trials?

1.8.3 Hypotheses

Combining the above literature and research questions the following a-priori hypotheses were made:

1A) BEST will be effective for supporting generalised and targeted language structures in children with language difficulties compared to Treatment as Usual.

1B) The inclusion of a signing system in BEST will provide additional support for generalised and targeted language structures compared to BEST without sign.

1C) There are no a-priori hypothesised effect sizes based on limited previous evidence for interventions of this nature.

2) There will be greater acceptability to parents and teachers due to the quasi-experimental design employed in the present research, compared to the student randomised controlled pilot study (see section 2.2).

Chapter 2

Methods

2.1 Introduction

Chapter one discussed the rationale for the present research and outlined the research questions addressed in this study. This chapter presents the methodology used to address the research questions. The study design, recruitment procedure, participants, measures, intervention, reliability, and statistical analysis for evaluating Building Early Sentences Therapy (BEST) are reported.

2.1.1 Research Questions

As set out in the previous chapter, this study has two strands of enquiry: the first quantitative strand evaluates the efficacy of BEST compared with Treatment as Usual. It seeks to determine the role of a signing system as an adjunct to the BEST intervention to determine whether sign is an active ingredient that promotes language development. This strand also provides point and interval estimates of effect sizes across a range of outcomes. Strand two informs the design of a future full-scale efficacy trial in terms of power, design, and processes through an acceptability and feasibility study carried out on stakeholder consultation data.

The methods for addressing the quantitative research questions 1A-C are described in this chapter and the results are reported in chapter three. Chapter four addresses research question 2 through analysis of stakeholder consultation with school staff. The methods, results and conclusions for this qualitative work are reported in chapter four.

2.1.2 Reporting Intervention Trials

Despite the shift towards evidence-based practice in healthcare, including Speech and Language Therapy (Zipoli & Kennedy, 2005), trial reporting remains poorer for complex, non-drug based intervention trials compared to drug trials (Glasziou et al., 2008). To counter these issues, guidelines have been developed for the effective designing and reporting of complex interventions (Campbell et al., 2000; Des Jarlais et al., 2004; Robey & Schultz, 1998; Schulz et al., 2010). The Consolidated Standards of Reporting Trials (CONSORT) guidelines are internationally accepted as the gold standard for trial reporting, and ensure that study findings are accessible to clinicians and researchers to support evidence-based decision making (Schulz et al., 2010). They aim to provide researchers with a framework for reporting to avoid problems of bias and inaccuracy arising from inadequate reporting.

More recently extensions have been provided to describe reporting procedures for different types of trials. The CONSORT Extension to Randomised Pilot and Feasibility Trials (Eldridge et al., 2016) provides flexibility at the early piloting and feasibility stages to address efficacy and effectiveness in small scale studies. The CONSORT checklist and Extension to Randomised Pilot and Feasibility Trials guidelines can be found in Appendix 4. The reporting of the present study follows the CONSORT Extension to Randomised Pilot and Feasibility Trials guidelines where possible, as these are the most appropriate guidelines available. The matching procedure is reported instead of a randomisation procedure since the study is non-randomised. Due to its piloting and feasibility nature, there is also a section of important changes to the study protocol that were introduced after the study began (see 2.15).

2.2 Prior Development of Building Early Sentences Therapy

BEST underwent some initial evaluation prior to the present study. The hypothesised underlying mechanisms for change are described in the BEST manual and include categorisation, schematisation and analogy (McKean et al., 2013). Following theoretical work undertaken to identify these potential mechanisms, a prototype of BEST was developed and applied across Heywood, Middleton and Rochdale (HMR) children’s SLT services (Pennine Care NHS Trust) (McKean et al. 2013).

2.2.1 Service Evaluation

Repeated cycles of small-scale implementation and adjustment, informed by discussion between SLT service managers, SLTs and researchers led to the development of the current BEST manual and standardised resources. Examples of modifications which occurred over this period include changes to toys and picture materials to ensure children could clearly differentiate between agents in the sentences, introduction of eligibility criteria to support SLTs to choose children who may benefit, and development of therapy recording forms to promote greater adherence to the protocol.

McKean et al. (2012) presented preliminary data from a service evaluation for 18 monolingual and bilingual children receiving BEST. Children's use of morphology and argument structure was measured at four time points using the BEST Assessment Progress Checker (McKean et al., 2013). All children made significant progress; 15 on both morphology and argument structure, and three on either morphology or argument structure. The service evaluation also gathered qualitative data indicating a high level of acceptability of BEST to SLTs, assistants and parents. These preliminary findings suggested that further research was warranted and resulted in a student-led randomised controlled pilot study.

2.3.2 Pilot and Feasibility RCT

Following the service evaluation development work detailed above, and prior to the present study, a small-scale randomised controlled pilot study was implemented by Speech and Language Therapy Undergraduate Students to gather preliminary data (Davis, 2016; Flintoff, 2016; Harwood, 2016; Isherwood, 2016; Standbridge, 2016; Topham, 2016). This student-led study compared progress made by children in BEST and Treatment as Usual (TAU) conditions on language and social, emotional and behavioural measures. Recruitment and retention data were recorded, and acceptability of the research process to teachers was explored through semi-structured interviews to inform future study design. Throughout this thesis this study is referred to as the student randomised controlled pilot study.

2.2.1.1 Progress on targeted language structures

Children in BEST and TAU groups made progress on standardised language tests and showed decreased difficulties scores on the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). All children made progress on targeted language structures, while progress was statistically significant for children receiving BEST.

Unexpectedly, negative effect sizes between the groups suggested that children in the TAU arm made more progress on standardised expressive language measures and on a measure of functional communication (Focus on the Outcomes of Communication Under Six (FOCUS)) (Thomas-Stonell et al., 2010; Washington et al., 2013). This may have been due to lack of control in the TAU group, as some children in the TAU condition received more intervention than children in the BEST condition. Additionally, despite the randomisation employed in the design, children in the BEST condition scored lower on baseline language assessments than the children in the TAU condition, suggesting that children in the BEST group had more severe impairments on average.

Results suggested positive benefits of receiving BEST, however many language interventions are associated with small gains of this nature on targeted structures (e.g. Ebbels, 2014). There was less evidence of generalised gains across broader language abilities and the power was not sufficient to draw conclusions. Despite this, there was adequate evidence to warrant further evaluation of BEST under more controlled conditions. The present study is the next step in the BEST development.

The Reynell Developmental Language Scales (Edwards et al., 2011) and Focus on the Communication of Children Under Six (FOCUS) questionnaire (Thomas-Stonell et al., 2012) were recommended for use in future studies due to their reliability and feasibility (Harwood, 2016).

2.2.1.2 Recruitment and Retention

In the student randomised controlled pilot study, 36 families were initially approached. Consent was given by 86% of families (n=31), and these children were tested on the baseline battery of assessments. Sixty-seven percent of children who were assessed were eligible to take part (77% of children whose families provided consent) (n=24). There was an attrition rate of 6% (n = 1) during the study as a result of a child's family moving away from the area. Sixty-one percent of children who were initially approached therefore completed the study and were included in the final analyses (n=22). Data gathered about recruitment and retention were used to inform sample size calculations and to estimate anticipated percentage of eligible children and rate of attrition for the present study.

2.2.1.3 Acceptability

Semi-structured interviews were conducted to assess the acceptability of the research to parents and teachers. Parents of children in the BEST condition reported positive experiences of the study and felt that communication with the Author and Research Assistants was good. They reported that their children were excited to take part in the intervention sessions and complete the homework booklets. They understood the benefits of the underlying principles such as modelling language and the use of repetition.

Conversely, parents in the TAU condition found the randomisation procedure highly unacceptable. They reported anger and frustration that their children had been identified as having language difficulties but did not receive any support. A lack of communication was reported between the Author and Research Assistants and parents of children in the TAU group, but ultimately parents' frustrations seemed to stem from lack of treatment for their children.

Teachers' attitudes towards the intervention were positive, although they reported feeling that communication with the Author and Research Assistants was insufficient. This was exemplified when one teacher did not know that the children had been receiving homework booklets throughout the study, suggesting that they could have encouraged parents to complete them had they known.

Teachers also reported being unhappy that children were taken out of their regular activities to receive intervention. These findings highlight the importance of effective communication and timetabling throughout the present study.

2.2.1.4 Summary

The findings of the student randomised controlled pilot study highlight key areas of learning that directly influence the design and execution of the present study. The issues surrounding randomisation, communication with schools and parents, and timetabling that emerged from this study are all addressed through design and study processes of the present research.

2.2.2 Current Implementation of BEST

BEST is currently implemented across multiple Local Authorities, primarily in schools and hospital outpatient clinics, by SLTs and other relevant trained professionals. No formal data is available about current implementation methods. Due to time and economic constraints on therapy provision in the NHS, it is likely that dosage and other delivery factors vary considerably across settings. Further work is required to establish the current situation with regards to BEST delivery.

2.2.3 Informing a Future Trial

The initial research outlined above suggests that BEST could be an effective intervention for children with language difficulties. Issues regarding acceptability, feasibility and the randomisation procedure have been identified. The findings of the present study will establish the efficacy of BEST and determine whether a future fully powered trial is feasible and warranted, in line with the MRC framework (Craig et al., 2008). Research questions 1A-1C examine the evidence to determine whether BEST is efficacious and are addressed in chapter three.

Research question two specifically concerns the qualitative data gathered about study processes and is addressed in chapter four.

2.3 Ethical approval

The Author and supervorsy team sought guidance from the Research Ethics Committee (REC) through the Integrated Research Application System (IRAS) on whether NHS ethical approval was required for the present study. Guidance from the REC suggested NHS ethical approval was unnecessary, as participants were not recruited by virtue of being on NHS caseloads. Consultation with local SLT services resulted in an approach whereby the Author would inform SLTs of which schools were participating in the study, to ensure a balance between collaborative partnerships and maintenance of confidentiality.

An application for ethical approval was made to the Newcastle University Faculty of Humanities and Social Sciences (HaSS) Ethics Committee on 21/11/17. Following minor clarifications, the project was approved on 13/12/2017. The reference number for this project is 1395.

Compliance with General Data Protection Regulations (GDPR) became a legal obligation on 25/5/2018. Information sheets and consent forms for the first data wave had already been sent to parents, so a further form was produced to inform parents of GDPR data procedures. This additional form was approved by the Newcastle University HaSS Ethics Committee on 25/04/2018. All information from the additional form was incorporated into the existing information sheets for the subsequent data waves. This amendment was approved on 26/06/18.

Copies of information sheets and consent forms for children in the BEST with sign treatment arm can be found in Appendix 5 and 6. All other information sheets and consent forms were a variation on these. See Appendix 7 for the additional GDPR form.

2.4 Design

2.4.1 Rationale for Choice of Design

In designing the study, the Author drew on the frameworks discussed in 1.7.3 to meet the goals of determining efficacy. The student randomised controlled pilot study found that randomisation at the level of the child was unacceptable to parents. Parents in the TAU arm reported frustration that their children were identified as having language difficulties but did not receive intervention. A matched groups quasi-experimental design was therefore the most effective and ethical design choice to address the research questions, allowing parents to be approached to take part in a specific treatment arm from the outset. The design also prevented parents within the same school making comparisons between children receiving different interventions, which may have exacerbated parents' negative perceptions in the Student randomised controlled pilot study.

A quasi-experimental design is the optimal choice for the present stage in the intervention development process. BEST sits between the *piloting and feasibility* and *evaluation* stages of development within the MRC framework (Craig et al., 2008), since it has undergone some pilot and feasibility work. The sample size in the student randomised controlled pilot study was such that the results could not be generalised to the wider population and results were inconclusive. Non-randomised designs are proposed in the 'Framework for Defining Feasibility and Pilot Studies' (Eldridge et al., 2016) as the appropriate method to test an intervention before planning a large-scale trial. This design is an economical method to gather further evidence about the efficacy of BEST, before deciding whether a costly fully powered evaluation is warranted.

The scale of the present study is such that random allocation may have introduced bias, rather than eliminated it, due to the modest number of schools and their differing environments and demographic characteristics. Random assignment could have resulted in heterogeneous treatment arms; schools were instead matched on Income Deprivation Affecting Children Index (IDACI) derived from the school postcode (Department of Communities and Local Government, 2015) and

Communication Supporting Classrooms Observation Tool (CSCOT) scores (Dockrell et al., 2015). These proxy measures of social disadvantage, which are described in 2.11.1, account for the school demographics and the degree to which the classroom environment contains factors likely to promote the children's language and communication development respectively. This method assigned schools to three equivalent treatment arms allowing for a robust comparison of children's progress across the study.

The quasi-experimental design also allowed schools to be matched accounting for their geographical location, reducing travel times so that the Author could deliver BEST to the maximum number of participants within each data wave, thus increasing statistical power. Due to the well documented difficulties working in schools due to timetabling and limited resources, this was a pragmatic choice allowing the Author to be flexible around the needs of the school, thus enhancing participation and minimising attrition (Bartlett et al., 2017).

2.4.2 Study Design

The study reported in this thesis is a non-randomised pilot study in line with the Conceptual Framework for Defining Feasibility and Pilot Studies (Eldridge et al. 2016). It sits between the piloting and feasibility stage and the evaluation stage of the MRC framework for evaluating complex interventions (Craig et al., 2008) and it measures efficacy outcomes as the study is carried out under ideal conditions (El Dib et al., 2015; Fey & Finestack, 2009; Robey & Schultz, 1998).

2.5 Materials

A range of paper-based assessments and intervention materials were used in the study. These are listed below with references to the corresponding assessments and appendices as appropriate.

- Information sheets and consent forms (Head Teacher and parent/carer versions; BEST with Sign, BEST without Sign, Treatment as Usual) (see Appendix 5 and 6 for examples for children in the BEST with Sign arms; all other versions are a variation on these)
- Printed assessment materials for:
  - Communication Supporting Classrooms Observation Tool (CSCOT) (Dockrell et al., 2012; 2015) (See Appendix 8)
  - Strengths and Difficulties Questionnaire (SDQ) (2-4 version) (Goodman, 1997) (See Appendix 9)
  - Focus on the Outcomes of Children under Six (FOCUS) (36 Item Version) (Thomas-Stonell et al., 2012) (See Appendix 10)
  - New Reynell Developmental Language Scales (NRDLS) (Edwards et al., 2011)
  - Targeted BEST Assessment (See Appendix 11)
  - Generalised BEST Assessment (See Appendix 11)
  - Alternative Picture Version Assessment (see RA Manual in Appendix 13)
- BEST Language intervention materials including therapy recording forms, BEST toys, BEST homework booklets, BEST sticker wall chart (LIVELY Project Group, 2019)
- Manual for Teaching Assistants (TAs) (See Appendix 12)
- Manual for Research Assistants (RAs) (See Appendix 13)
- Sony ICD-SX2000 audio recorder

2.6 Participants

2.6.1 Schools

2.6.1.1 School Eligibility and Exclusionary Criteria

To be eligible to take part, schools were required to support recruitment and provide space for assessments to be conducted. Additionally, schools in the BEST arms were asked to provide a teaching assistant to support delivery of the intervention twice per week, and an appropriate room or quiet space within the school setting to be used for BEST sessions.

Schools were not approached to take part if they were taking part in other known Speech and Language Therapy intervention studies to avoid contamination effects.

2.6.1.2 School Recruitment

The present study recruited schools from one Local Authority (LA) in the United Kingdom (see 2.7.1 for more information about the demographics of the area).

The LA encompasses 59 state primary schools. Recruitment targeted schools with high levels of SLCN to increase potential recruitment. Twenty schools were participating in the Nuffield Early Language Intervention (NELI) trial (Fricke et al., 2013) and so were not approached.

As detailed below, NHS SLTs acted as gatekeepers and approached Head Teachers (HTs) or Special Educational Needs Co-ordinators (SENCOs) of the 39 schools not involved in the NELI trial. With consent they passed the contact details of the interested schools to the Author, who arranged a face-to-face meeting.

The staff who attended the initial meeting with the Author varied, although either the HT or SENCO was always present. The Author supplied detailed information leaflets to explain the study and address any questions or concerns. If satisfied, the HT completed the consent form, giving the Author permission to recruit children from that school.

In the second and third data waves, schools who had already taken part were invited to continue into the following waves, while schools that had not initially responded were approached again.

2.6.1.3 NHS SLT Gatekeepers

There are several barriers to conducting school-based research, including gaining entry to schools, constraints on time, staff and resources and building relationships with schools and families. Using a school-based liaison or gatekeeper can facilitate entry to the school, increase engagement with the target population, and increase participation in the project (Bartlett et al., 2017). The Author utilised an existing relationship with the SLT services to introduce the project to the SLTs serving the Local Authority schools. The SLTs acted as gatekeepers who contacted schools directly, utilising already established relationships. This strategy was adopted to facilitate connections with the senior leadership team and gain trust within the schools. The approach contrasts to a 'cold-calling' recruitment strategy which is not advised as it can be time consuming and ineffective (Roschelle et al., 2014).

2.6.1.4 School Matching and Treatment Arm Allocation

Schools were matched to produce three equivalent groups that did not differ significantly on baseline demographic characteristics such as classroom oral language environment and socio-economic status (Bray et al., 2009).

Matching was carried out based on the following characteristics:

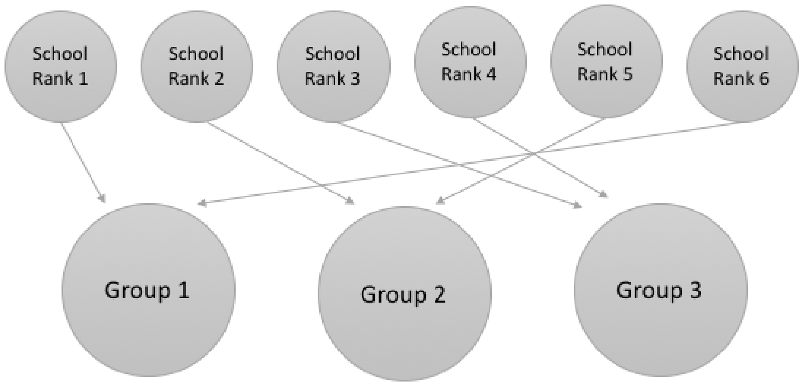
1) I–CAN Communication Supporting Classrooms Observation Tool (CSCOT) (Dockrell et al., 2015)
2) Income Deprivation Affecting Children Index (IDACI) rating, based on the school postcode (Department of Communities and Local Government, 2015)

These measures are discussed further in the measures for matching schools section below.

The protocol for assigning schools to treatment arms was as follows: First schools were ranked from highest to lowest according to Income Deprivation Affecting Children (IDACI) score. Schools were then assigned to three groups, maximising homogeneity as demonstrated in figure 2.1.

Figure 2. 1

Example of the assignment of ranked schools to matched groups.



Next, the CSCOT total scores were ranked from highest to lowest for each school, and the minimum number of schools were moved between arms for optimum homogeneity across arms on the IDACI and CSCOT scores. Groups were adjusted for class size and number of nursery and reception classes per school, so each contained an equal distribution of similar nursery and reception classes, whilst maintaining homogeneity across IDACI and CSCOT scores.

This protocol was intended to be followed exactly, but due to an error in the matching process, the treatment arms were based solely on CSCOT and IDACI ratings (see 2.15.2).

The groups were then manually randomly assigned to one of three intervention arms by a member of the supervisory team who had no knowledge of the school characteristics. The treatment arms were: 1) BEST with Sign; 2) BEST without Sign; and 3) Treatment as Usual (TAU).

The Author provided and calculated all of the matching variables. Matching was carried out by the first thesis supervisor to avoid any potential bias associated with the Author assigning groups to intervention arms, having already spent substantial time in each school.

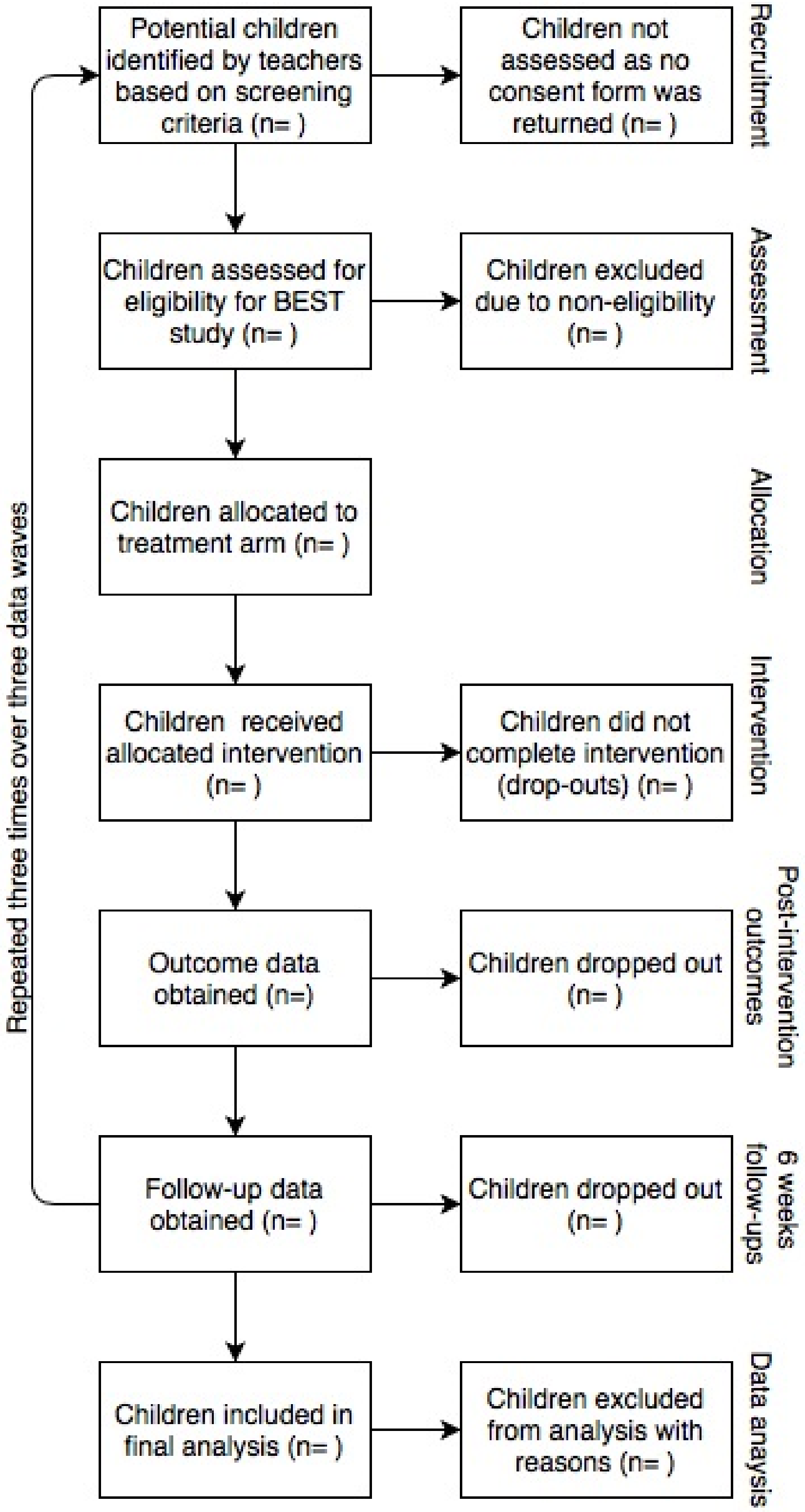
2.6.2 Children

Following the allocation process schools were informed of the treatment arm they had been assigned to. Recruitment then commenced, with schools and parents having full knowledge of which intervention arm they would be part of.

Figure 2. 2 below shows the journey of participants through the study from recruitment to follow-up measures. The process was repeated three times over three data waves with three groups of children. This approach optimised recruitment of sufficient participants while ensuring the Author could deliver all BEST sessions in schools around the LA.

Figure 2. 2

Participant flow-chart for the BEST nonrandomised pilot study



Note: see figure 3.1 for recruitment results with ns

2.6.2.1 Child Eligibility and Exclusionary Criteria

2.6.2.1.1 Teacher Screening Criteria

Teachers were initially asked to identify children based on meeting the following descriptive screening criteria:

- Aged between 3;5-4;5 years
- Monolingual speaker of English, or home environments where the family speak to the child almost exclusively in English
- No diagnosed learning disability
- The child was not meeting the language and/or communication expectations on the Early Years Foundation Stage Profile, or not meeting other expected targets or milestones, or teachers were concerned for other reasons relating to the child's language development
- The teacher felt that the child was able to cope with the nature of the assessments and language intervention

2.6.2.1.2 Eligibility Criteria

Children whose parents or carers provided informed consent were measured on a number of assessments described in the measures section below to determine formal eligibility based on the following criteria:

- Aged between 3;5–4;5 years old
- Monolingual speaker of English, **or** almost exclusively exposed to English, based on parent/carer self-report of percentages of English and other languages spoken in the home. In unclear cases (e.g. where a small percentage of another language/languages were spoken), children were not be immediately assessed, and their case was discussed with the supervisory team before a consensus was reached using RCSLT guidance (Royal College of Speech and Language Therapists, 2007)

- New Reynell Developmental Language Scales (NRDLS) scores falling on or below the 16th centile on expressive and/or receptive language (see below for rationale)
- BEST assessment score demonstrates that the child would benefit from language intervention. This was determined on a case-by-case basis by the supervisory team based on the inflectional morphology and vocabulary difficulties indicated by BEST assessment score suggesting the child was not already using all of the target structures consistently.
- Displays triadic attention, symbolic play and imitation in play during the assessments, to ensure children have sufficient nonverbal skills and Theory of Mind to access the intervention. Determined by teacher judgement and observed performance during standardised baseline tests.

The eligibility assessment scores were used as baseline scores for eligible children.

As stated above, the 16th centile on the NRDLS comprehension and/or production subscales was used as the cut-off criterion in the present study. Children therefore presented with mild to severe difficulties (e.g. Wiig et al., 1992). BEST was originally designed for children with severe language difficulties (McKean et al., 2013), however this decision was made to acknowledge that not only the most severe children are in need of, or could benefit from, intervention (Law et al., 2017). Evidence from epidemiological studies shows that children who score below the 16th centile remain at risk of poorer outcomes (Beitchman et al., 2001). The 16th centile has been used for clinical identification of language impairment (Conti-Ramsden, 2003) and as a cut-off for intervention studies (Pawlowska et al., 2008; Rice et al., 2010). Due to the nature of this piloting and feasibility study, the criteria for inclusion was also intended as a way to establish which children would benefit from BEST, as well as being a pragmatic way of maximising sensitivity and ensuring sufficient power within the constraints of the study.

Children were excluded from participation if they met any of the following criteria:

- Dominant/preferred speaker of a language other than English, or a child for whom BEST should be delivered in a language other than English (see above)
- Sensorineural hearing impairment
- Severe visual impairment
- Diagnosed Learning Disability (LD) or Autism Spectrum Disorders (ASD)
- NRDLS scores falling above the 16^{th} centile on both expressive and receptive language

2.6.3 Child Recruitment

Teachers initially identified children according to the descriptive screening criteria above. The screening criteria were designed to be more inclusive than the eligibility criteria to minimise the number of eligible children who were not put forward.

Information sheets and consent forms were sent home to parents/carers of the children identified by teachers. Parents were encouraged to contact the Author to ask questions. A pragmatic approach to recruitment was taken to best suit the needs and timetables of individual schools and face-to-face meetings with parents were offered to all schools. The Author obtained fully informed consent from parents/carers only when satisfied that questions had been resolved.

2.7 Procedure

2.7.1 Study Context

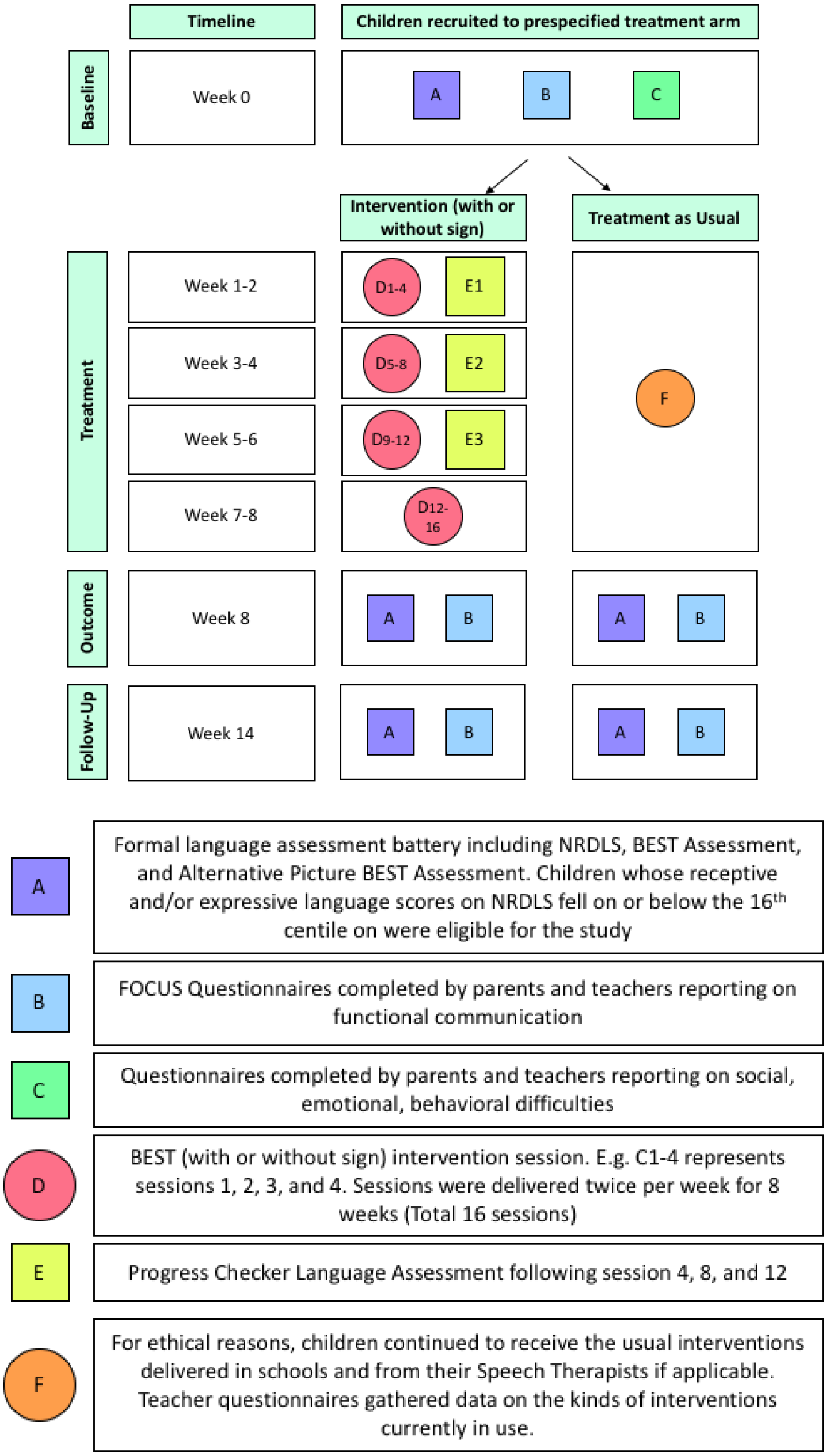
The study was conducted across one Children's Services Directorate in a Local Authority in England with a strong reputation for its work addressing speech, language, and communication needs. The 82 km^2 site serves 59 primary stage schools and approximately 42,700 children under the age of 18. In

2014 a review of the same Local Authority found 30% of the areas within the LA were ranked within the 25% most deprived areas in England (McKean et al., 2017). All schools were part of the LA School Improvement Service, and the site had an effective history of NHS and LA Children's Service collaboration, as well as being a Bercow Pathfinder site (Bercow, 2008, 2018; Gascoigne, 2012).

The study took place between April 2018 and July 2019. Data collection took place over 3 data waves to effectively distribute time and resources while including maximum participants (see Figure 2. 3). The data waves ran from: 1) April 2018 – September 2018; 2) September 2018 – January 2019; 3) February 2019 – July 2019.

Figure 2. 3

Perera Diagram displaying overall study processes. The cycle is repeated 3 times over 3 consecutive school terms.



2.7.2 Study Processes

2.7.2.1 Schools

Figure 2.3 outlines the study procedures. NHS SLTs covering the Local Authority schools acted as gatekeepers to schools. SLTs approached the 39 schools not participating in another language trial inviting them to take part in the study. The Author contacted Head Teachers who expressed interest and organised a meeting to discuss the study and obtain informed consent to recruit from within the schools. The Author completed the Communication Supporting Classrooms Observation Tool (CSCOT) (Dockrell et al., 2012) in each participating school as a measure of the oral language environment. Class teachers completed a qualitative questionnaire which was used to understand current language interventions and classroom practices. Schools were matched and assigned to one of three treatment arms (BEST with Sign, BEST without Sign, Treatment as Usual) based on the CSCOT scores and Income Deprivation Affecting Children (IDACI) (Department of Communities and Local Government, 2015) scores (see Figure 2. 1). Schools were informed which treatment arm they had been assigned to.

The Author trained the group of five RAs who each had experience of working with children to conduct the assessments and reliably score and record the assessment measures. The Author was blinded to assessment results and the RAs were blinded to intervention arm to maintain integrity and avoid any biases (Eldridge et al., 2016) (see 2.12 for more information on the blinding procedure).

2.7.2.2 Children

Teachers recommended children for the study based on the teacher screening criteria (see 2.6.1.1). Consent forms were sent home to parents or carers inviting their child to take part in the specified treatment arm.

Once informed consent was obtained, trained Research Assistants administered the assessment battery (see 2.11.2). Teachers and parents also completed SDQ and FOCUS questionnaires for each child. The research assistants visited each school in pairs. All RAs were trained to deliver all measures, but for ease one RA delivered the BEST eligibility assessment (see 2.11.2.1) and NRDLS (see 2.11.2.2) while the other delivered the Targeted BEST Assessment (see 2.11.2.3), Generalised BEST Assessment (see 2.11.2.4) and Alternative Picture BEST Assessment (see 2.11.2.5). The NRDLS was split over two days to maintain children's concentration and participation, with the comprehension subscale being completed first, followed by the production subscale in accordance with the manual.

Children's eligibility was determined based on the assessment battery scores. Parents and teachers were informed whether children were eligible. For the eligible children the assessment battery scores formed the baseline assessment data. Eligible children then received 2 sessions of BEST (with or without sign) per week for eight weeks or received Treatment as Usual for 8 weeks. All intervention sessions were delivered by the Author and in partnership with school TAs where possible.

As soon as possible following the 8-week treatment period, RAs administered the outcome assessment battery. The same assessments and questionnaires excluding the eligibility assessment and SDQ were carried out. The follow-up assessments were completed six weeks after the outcome assessments. Children's participation in the study was then complete, and parents and teachers were informed and thanked for their participation.

2.8 The BEST Intervention

This section summarises the manualised intervention (presented in table 2.1). Section 2.8.4 then describes how the intervention was delivered in the present study, highlighting the ways in which delivery deviated from the manual.

The protocol is as follows: sessions are delivered in groups of one-to-six children in quiet room or area in the school. Two adults: The Author (henceforth Adult 1), and a member of staff from the school (henceforth Adult 2), deliver the sessions. Sessions take place either sitting at a table or on the floor. The adults and children sit in a circle so that all participants can see each other. Sessions last for 15-20 minutes depending on the number of children in the group. Children receive a total of 16 sessions.

The therapy recording form lists the three sets of model and target sentences for each session. The sentences are designed to provide repetition, variation and contrast to support children to bootstrap learning (see 1.5.1 and Table 1. 4). Corresponding toy bags for each set contain all of the toys necessary to act out the sentences for the session.

2.8.1 Phase One

As set out in Table 2. 1, during a BEST session Adult 1 models three sets of model sentences orally and with sign (in the BEST with Sign arm). Adult 2 acts out the sentence with the BEST toys. The model sentences are simple sentences of varying clause-element and construction types containing verbs.

Other verbal input such as praise is avoided so that all spoken input can be mapped directly onto the visual input. Emphasis is placed upon the input and children do not have to take an active role unless they wish to.

2.8.2 Phase Two

Adult 2 acts out one of the target sentences with the toys and Adult 1 asks each child in turn 'What's happening, [*child's name*]?'. The child's response is recorded in the therapy recording form and Adult 1 recasts the correct target sentence (and signs, when applicable) regardless of the child's response.

Corresponding homework booklets are sent home for each session. To complete the homework, parents are asked to look at the booklets with their child every day and talk about what is happening in each picture. Parents are told that children do not have to say anything if they do not want to, as the most important thing is for the child to be exposed to the input provided (also see section 1.5.1).

Table 2. 1

The manualised BEST Procedure, Reproduced with Permission from the BEST Manual (McKean et al., 2013 pp. 6).

Step	Adult 1	Adult 2
Step 1: Introduction of Noun Vocabulary	1. Take the toys out of the bag and place them on a table in easy reach of the child saying "I've got a…" 2. If the child doesn't name after a short gap, then name the toy. Do not ask "What's this?" 3. Repeat with all objects for the set. Note: it is not essential for every child to name every object. If any child clearly labels an object, repeat this to reinforce and move onto the next object.	
Step 2: Input (phase 1)	1. Say: "Let's see what the people are doing". 2. Carry out the action with the toys and give a verbal model at the same time. 3. Repeat until all of the sentences have been presented to the children.	Provide a supporting visual model using PGSS while the action is taking place.
Step 3: Output (phase 2)	1. Carry out the first action with the toys. 2. Help Child 1 to carry out that action with the toys. Child 1 is not required to speak at this point. 3. Help Child 1 to repeat/continue the action while Child 2 is attempting to describe the action. 4. Repeat for each of the output sentences in the given set with the children's roles moving clockwise around the semi-circle of children (i.e. For sentence one: child 1 carries out the action while child 2 gives output. For sentence two: child 2 carries out the action while child 3 gives output).	1. While Child 1 is carrying out the action ask Child 2 "What's happening?" 2. However the child responds, this attempt is accepted by the adult who recasts this attempt – "Yes……the teddy is washing the apple". **The child is not asked to repeat this recast**. 3. Record the child's first response in the recording form. (Spontaneous responses of other children in the group should also be recorded.) Note: Do not give forced alternatives, scaffolded models or demand an imitated response or eye contact from the child. Additional verbal praise is also not necessary.
	Repeat the 3 steps above for the remaining 2 sets required for this session	

2.8.3 Signing System

As discussed in detail in 1.6.5, the rationale for testing the efficacy of a signing system as an adjunct to BEST is twofold. Firstly, there is a broad need to establish whether the use of sign could support language development and improve outcomes for children with language difficulties. Secondly, the rationale for testing this specific signing system as an adjunct to BEST is to determine whether the benefits outweigh the costs of training teachers to deliver sign. The BEST manual stipulates that Paget Gorman Signed Speech (PGSS) should be used for all signing, but discussion with LA SLTs and teacher questionnaires revealed that many teachers used Makaton in the classroom. For this reason, a combination of PGSS for inflectional morphology and Makaton for vocabulary were used, which was approved by the BEST authors. The signing manual containing all the signs can be found in Appendix 14.

2.8.4 Intervention Arms

This study compared three intervention arms: Best with Sign, BEST without Sign and Treatment as Usual. This comparison was made to determine whether there is evidence for the efficacy of BEST over Treatment as Usual. The following section describes the intervention delivery for each treatment arm in the present study.

2.8.4.1 BEST Intervention with sign

The Author (Adult 1) worked in partnership with a trained TA from each school, although sometimes this was not feasible given school timetabling issues and staffing shortages, in which case the Author conducted the sessions without additional support (also see 2.15.4). The sessions began with a song where everybody introduced themselves to build rapport.

2.8.4.1.1 Phase One

Adult 1 said and signed the model sentences (n = 3-6). Adult 2 acted out the sentences with the BEST toys. During the presentation of the model sentences, all other verbal input such as praise was avoided where possible.

If a school staff member was not available to fill the role of Adult 2, the Author acted out the sentences with the toys and then signed and said the sentences simultaneously (see 2.15.4 for further detail). In both cases the sentence was spoken once only.

2.8.4.1.2 Phase Two

Following presentation of the model sentences, the target sentences were acted out by Adult 2 using the BEST toys. Adult 1 asked each child in turn 'What's happening, [*child's name*]?' and the child was given the opportunity to describe the action. Regardless of the child's response, Adult 1 recast and signed the correct sentence in full. Responses were audio recorded and transcribed onto the Therapy Recording Form immediately after the session.

If a school staff member was not available to fill the role of Adult 2, Adult 1 acted out the sentences with the toys before asking the child 'What's happening' (see 2.15.4).

2.8.4.1.3 Homework Booklets

Homework booklets were sent home following each session. Parents were asked to look at the pictures with their child and discuss what was happening in each, to continue the child's exposure to the target sentences. Parents were told not to attempt to sign the sentences as this may have introduced bias. They were also told that their child did not have to speak unless they wanted to and the parents should not try to encourage this.

All children continued to receive all their usual classroom provision throughout the study (see the Treatment as Usual section below for detail).

2.8.4.2 BEST Intervention without sign

The BEST without sign intervention arm was identical to the BEST with Sign arm, other than the omission of signing. In BEST without sign, Adult 1 said the sentences while Adult 2 acted out them out with the toys.

If Adult 2 was not available, Adult 1 acted out and said the sentences simultaneously (see important changes to methods in section 2.15.4 on this).

2.8.4.3 Treatment as Usual (TAU)

The withdrawal of other treatments that children may already be receiving can be considered unethical in intervention studies. The Treatment as Usual (TAU) alternative is often implemented instead, whereby children continue to receive their usual provision including intervention, strategies and other support (Craig et al., 2008).

In the student randomised controlled pilot study, some children in the TAU condition reportedly received more intervention than children receiving BEST (determined by time spent receiving intervention). It was therefore important to sufficiently document TAU provision so that the results could be interpreted in light of this.

During the eight-week intervention phase of the study, children in the TAU arm received no input from the Author but continued to receive the interventions and support that their school and SLT (if applicable) would usually deliver. Children in the BEST arms also received their usual provision throughout the intervention phase.

Teachers completed a questionnaire so that the Author could gain a qualitative understanding of the universal and targeted interventions which were being delivered in classrooms (Bayat et al., 2010). Teachers were also asked about what input SLTs had in the classroom. Appendix 15 summarises the main interventions, activities and strategies used by teachers to support language in classrooms involved in the present study. Teacher questionnaires suggested that children received frequent, high-quality language intervention, strategies and activities to promote oral language development. Based on the teacher questionnaires, children were receiving a significant amount of support with many teachers reporting one-to-one or small-group intervention daily, or on most days, for approximately 5-10 minutes. As well as this, teachers also supported children via strategies and activities including visual supports such as symbols and signs, storytelling, circle time sessions and teaching meaning and checking understanding.

2.9 Intervention Fidelity

2.9.1 Avoiding Biases in Intervention Research

Bias in intervention research can be caused by numerous factors such as lack of blinding, lack of treatment fidelity, allocation bias and observer bias. Intervention fidelity is vital for a successful trial as it increases the validity and reliability of the intervention under examination, and thus the confidence that researchers can have in trial results (Bellg et al., 2004). The results of inadequately implemented intervention trials will not represent the true outcomes of interventions as they were intended (Mars et al., 2013). In line with the CONSORT statement (Eldridge et al., 2016; Schulz et al., 2010) in this study intervention fidelity is promoted through robust study design, high quality training for those involved in intervention delivery, and manualised assessment and scoring procedures. The steps taken to implement these are described in the following sections.

2.9.2 BEST Manualisation

BEST is a manualised intervention (McKean et al., 2013). The manual details the theoretical rationale for the intervention and identifies potential underlying mechanisms for the process of change. The intervention is scripted and involves standardised toys and delivery. The Author developed a protocol for scoring the assessments using the standardised therapy recording forms and distributed standardised homework booklets. The manual was strictly adhered to during the present study to increase intervention fidelity and validity (Mars et al., 2013). Any deviations from the manual are clearly reported in the Important Changes to the Protocol section 2.15.

2.9.3 Author's Training to Deliver the BEST Intervention

Prior to training, the Author met with Dr Sean Pert, co-author of BEST, at Manchester University, who introduced the intervention, its delivery and its theoretical underpinnings. He identified several key principles of BEST that should be adhered to, including the avoidance of verbal praise, adhering to the manualised script and children not being obliged to talk until they wish to do so.

The Author observed three videos of BEST sessions taken during the student randomised controlled pilot study and reflected upon them with the supervisory team.

The Author then visited a hospital in the Midlands eight times over two months to deliver BEST alongside a Speech and Language Therapist. BEST had already been implemented at the site for over a year, and the SLT was experienced in its delivery. The Author played the role of Adult 2 delivering the sessions, and followed the procedure detailed in the BEST manual (also see Table 2. 1). Throughout the training the Author received constructive feedback from the SLT and received accreditation to deliver BEST.

2.10 Teaching Assistant BEST Delivery Training

In the study the Author played the role of Adult 1, while a trained teaching assistant from each school fulfilled the role of Adult 2. The Author delivered training to each TA prior to the study, consisting of an overview of BEST and the principles underpinning its hypothesised mechanisms. The Author described key elements of BEST such as not giving verbal praise during the modelling phase and children not having to respond until they feel ready. The reasons behind these principles were explained. The Author and TA carried out a role play session during the training and the Author provided constructive feedback to the TA and answered any questions. A manual was developed for TAs to refer to (see Appendix 12).

2.10.1 Intervention Fidelity

Throughout the intervention sessions the Author video recorded the third BEST session for every BEST group (with and without sign). These videos were rated by Sean Pert using a treatment fidelity rating tool, so that any inconsistencies in delivery could be addressed. Table 2.2 below outlines the rating criteria. Ratings were high for every observation and Sean Pert did not express any concerns regarding intervention fidelity. One minor change to the protocol occurred here, in that the Author recorded responses using a dictaphone and transcribed them immediately following the sessions (see the final item under Output Phase in Table 2.2). This was not deemed to be problematic by the Authors of BEST. To ensure reliability the Author rechecked five session recordings at the end of data collection and found 100% of responses to be accurately transcribed.

Table 2. 2

Rating criteria for videoed BEST sessions

Step	Item[1]
Introduction of Noun Vocabulary	All items are shown to the child one-by-one and the child is given an opportunity to name the item
	All items are labelled correctly by the child spontaneously naming the items or the adult names the item or the adult correctly names an item when the child has incorrectly named an item
Input Phase	Toys complete the action with the verbal model given simultaneously
	Makaton/PGSS is provided at the same time as the verbal models
Output Phase	The child completes the action with the toys accurately or is facilitated to do so by the adult
	The child is prompted with ‘What’s happening?’ or similar
	The adult gives a recast of the target spoken sentence
	No repeat of the recast is requested of the child
	No verbal comments or praise are used, e.g. ‘Good try!’, ‘Well done’ etc.
	The child’s first spoken sentence is written down accurately in the recording booklet or recorded using a recording device

Group Working	Group Model or One-to-one model
	For each sentence one child (in turn) is given the toys and told “Now you do it”. The child makes the toys complete the action and is facilitated by the adult if they are not able to complete the action themselves. The researcher then repeats the sentence correctly once more.

[1] Ratings were based on a 5 point Likert scale (Never, Almost Never, Occasionally/Sometimes, Almost Every Time, Every Time), apart from the ‘Group model or one-to-one’ item for which the appropriate answer was selected based on how many children were in the BEST group.

2.10.2 Development of a 'Database of Ambiguities' for Scoring All Language Assessments

During all components of the assessment battery children's responses frequently deviated from examples given in the scoring protocol and were therefore difficult to score. Items were deemed ambiguous for the reasons set out in Table 2. 3 below. Resolutions to these were agreed upon between the Author and thesis supervisors.

When RAs encountered an ambiguity, they checked the database in the first instance. If a resolution had already been reached for that ambiguity then it was applied, maintaining scoring reliability. Where the ambiguity had not already been encountered, the RA discussed the response with the Author, maintaining blinding at all times. Once agreed, the resolution for how to score the item was recorded in the Ambiguities Database so it could be applied to future occurrences.

The RA Protocol and Ambiguities Database can be found in Appendix 13.

Table 2. 3

Assessment scoring ambiguities and agreed resolutions

Identified Ambiguity	Agreed Resolution
Children's speech sound difficulties made responses difficult to understand	The RAs used their knowledge of the individual child, speech sound difficulties and the context of the target item to interpret the sentence. The RAs consulted with a qualified Speech and Language Therapist when necessary. If no consensus could be reached the item was scored as incorrect, although in practice this was extremely rare.
Children did not produce the target agents, actions or patients, either omitting them or using other language to label them	Acceptable determiners for the agent were 'that', 'the' and 'a'. Acceptable determiners for the patient were 'the', 'a', 'his/her'. Where a noun argument was marked but the wrong lexical item was used e.g. 'dog' for 'cat' this was scored as 'signalled' but not 'correct'. Where a subject pronoun 'he or she' was used for the agent this was accepted as 'signalled' and 'correct'. Responses not containing target vocabulary, judgements were made about whether the vocabulary in the child's response was semantically similar or different to the target sentences. For example, when the target agent was 'The Lady', but a child produced 'The Mummy', this would be accepted since 'lady' and 'mummy' would be pictorially represented the same way. If the child produced 'Queen' for 'Lady' for example, this would not be accepted as the pictorial representation does not contain the common semantic features of a 'Queen', such as a crown.
Children did not describe the picture but produced another utterance	Correct items were marked in the normal way. Off-topic, irrelevant and other such spoken utterances were scored as incorrect.

Children used sentence structures other than the target structures	Correct items were marked in the normal way. Grammatically correct utterances were scored as signalled but not correct. In practice this was extremely rare. Off-topic, irrelevant and other such spoken utterances were scored as incorrect.
Children self-corrected their initial response	Where a child made a false start and self-corrected part way through an utterance the complete corrected sentence was analysed e.g. ‘the boy… **the man is eating the apple**’. Where a child makes made attempts or produced separate clause elements as separate utterances, only the first attempt/utterance was analysed.

2.11 Measures

The assessment battery was designed to measure progress on children's outcomes to determine whether there was evidence of intervention efficacy. The battery included standardised and non-standardised measures, measuring targeted and generalised language.

Standardised omnibus tests such as the New Reynell Developmental Language Scales (NRDLS) are robust measures of change that have been applied to the normal population. These tests measure a wide range of skills and progress on such measures would indicate a change across multiple systems or the ability to learn new language structures.

Omnibus language tests should be interpreted with caution, as their applicability to a particular child will depend on the population it was standardised for (Dockrell & Marshall, 2015).

In intervention studies children frequently make progress only on targeted language structures that do not extend to broader components of language (see Ebbels, 2014) and generalisation across structures is often not measured.

This study therefore utilises other, non-standardised tests of target structures and their generalisation to other sentences with the same predicate argument structure.

2.11.1 Measures for Matching Schools

2.11.1.1 CSCOT

The Communication Supporting Classrooms Observation Tool (CSCOT) (Dockrell et al., 2012; 2015) measures classroom oral language environment. The tool was funded by the UK Government following the first Bercow report (Bercow, 2008) and was developed as part of the Better Communication Research Programme (Lindsay et al., 2012). The CSCOT measures three language learning dimensions: environment, interactions and opportunities and it yields a numerical score for each dimension. The CSCOT dimensions are based on extensive evidence and are aligned with UK

SLCN policies. Feasibility testing has shown the CSCOT to have good inter-rater reliability (Dockrell et al., 2015).

The Author completed a CSCOT observation in each classroom involved in the study. In accordance with the instructions each observation lasted one hour and was carried out during a morning classroom session beginning with the register. The Author received training before completing the CSCOT observation (see CSCOT Training and Reliability section below).

2.11.1.2 Income Deprivation Affecting Children (IDACI)

The Income Deprivation Affecting Children (IDACI) is a supplementary extension to the Index of Multiple Deprivation (IMD) (Department of Communities and Local Government, 2015). IDACI describes the number of 0-15-year-old children living in income deprived households by measuring the proportion of the population living in economic deprivation. IDACI scores were generated based on school postcodes and used as an indicator of deprivation affecting children (calculated using the English Indices of Deprivation online database (2015 version)) (Ministry of Housing Communities and Local Government, 2015). The more deprived a postcode area, the higher the IDACI score. It must be noted that IDACI score was calculated based on the school postcode. Socioeconomic status of individual families likely varied.

2.11.1.3 Teacher Questionnaires

Semi-structured questionnaires were completed by the class teachers (or by the Author during a conversation with the teacher). Questions were simple and open-ended, aiming to gain a qualitative understanding of what interventions and support were being delivered in schools. The results did not form part of the official matching criteria, instead this understanding was crucial for contextualising the findings. The questionnaire and summary of responses can be found in Appendix 15.

2.11.2 Direct Measures

2.11.2.1 BEST Eligibility Assessment

The Eligibility assessment is part of the standardised package of published resources available with BEST (see Appendix 16). To benefit from the BEST intervention, children need to demonstrate symbolic play, triadic attention and imitation in play. They must also demonstrate significant language difficulties, be aged 3;5-4;5 and a monolingual/almost exclusive speaker of English. The criteria were also applied to children in the TAU arm. Research Assistants completed the eligibility assessment based on their interactions with the child when conducting the other measures.

2.11.2.2 New Reynell Developmental Language Scales (NRDLS) (Edwards et al., 2011)

The NRDLS is a manualised, standardised, omnibus Language assessment for children aged between 2;0-7;05. The NRDLS comprises of seven expressive and receptive language sections and is administered through play-based assessments.

Scores for expressive and receptive language abilities were generated and used to determine eligibility (i.e. on or below the 16th centile for receptive/expressive/both subscales), identify language profiles (expressive/receptive/mixed expressive-receptive) (see Table 2.4) and to measure change across time points for participants. Both raw and standard scores were calculated and recorded to assess change in absolute terms, as well as relative to the child's peers.

Table 2. 4

Language profiles categorised by performance on the NRDLS Expressive and Receptive subscales

		Receptive	
		Below 16th percentile	Above 16th percentile
Expressive	Below 16th percentile	Mixed Profile	Expressive Profile
	Above 16th percentile	Receptive Profile	Ineligible for study (scoring above 16th percentile for both measures)

2.11.2.3 Targeted BEST Assessment (McKean et al., 2012)

The Targeted BEST Assessment is a non-standardised picture description task included in the BEST resource pack which is designed to measure expressive language. The items are a sample of the BEST sentences used in the main intervention input and were selected by the BEST intervention Authors (McKean et al., 2013). They are representative of the varying complexity of sentence structures used in BEST (see table 1.4). The assessment measures children's use of predicate argument structure (PAS) and grammatical morphology when describing the pictures that are presented to them. Examples of predicate argument structure which, necessarily encompasses vocabulary, that are probed in the assessments are agents (e.g. man, lady, boy, teddy), verbs (e.g. walk, run, jump, wash), patients (e.g carrot, banana, lolly, key) and locatives (e.g. shoe, box, cup, bed). Predicate argument structure and vocabulary are referred to as *content* for the remainder of this thesis, since together they constitute the sentence content.

The BEST assessments also measure grammatical morphology, specifically the presence and correctness of the appropriate determiners (e.g. the, a, an), auxiliaries (e.g. is), and verb inflection (e.g. -ing) for each sentence. The assessments capture both correct and incorrect attempts made for each item allowing for a more nuanced view of the child's language than simply scoring items as correct or incorrect would provide. Appendix 13 provides further information on the assessments and scoring procedures.

When administering the assessment, children are shown each picture in turn and asked to describe "what's happening here". Their first spontaneous response is recorded on the assessment recording form and using a dictaphone. Since no formal manual existed, the Author created standardised protocols for conducting and scoring the BEST assessment (see Appendix 13).

2.11.2.4 Generalised BEST Assessment

The BEST intervention is designed to support children to generalise across exemplar constructions to produce novel, flexible constructions in line with the constructivist account of language acquisition (Langacker, 2000; Tomasello, 2003, 2009a). The Targeted BEST Assessment described above measured progress on target BEST sentences exactly as they appear in the BEST intervention. The Author therefore designed this novel assessment to assess whether children were making meaningful progress on the underlying sentence predicate argument structure and morphology measured in BEST, as opposed to simply becoming familiar with the BEST sentences due to continuous exposure from the intervention sessions.

The Generalised BEST Assessment measures whether any progress observed on the BEST Assessment generalises to other sentences describing differing agents, actions, patients and locatives. The generalised BEST assessment constructions therefore followed the same predicate argument structure (e.g. agent, action, patient) but contained different verbs and nouns which were matched with the BEST sentences on typical age of acquisition (Morrison, Hirsh, & Duggan, 2003).

Like the Targeted BEST Assessment, the Generalised BEST Assessment was designed to test whether progress with content and morphology BEST target items generalised to sentences with the same predicate argument structure but different content.

2.11.2.5 Alternative Picture Version (APV) BEST Assessment

A further generalisability assessment, the Alternative Picture BEST Assessment, was created to examine whether there were any measurable practice effects for children receiving BEST because of the same artwork being used in the BEST homework booklets and the BEST assessments. The APV was developed after commencement of the study and 50% of children therefore did not complete it.

The Alternative Picture BEST Assessment mirrored the BEST structure and sentences completely but used pictures redrawn by a different artist in an attempt to mitigate the effect of picture familiarity on performance.

This version was piloted with two typically developing children aged five and seven, who scored above 85% (above mastery score (McKean et al., 2013)), to ensure pictures were still recognisable.

2.11.2.6 BEST Progress Checker

The BEST Assessment Progress Checker (McKean et al., 2013) is a shortened version of the Targeted BEST assessment (8 items), which follows the same picture description task structure outlined in 2.11.2.3. The progress checker was administered by the Author following BEST session number four, eight and 12 to monitor children's progress over the course of the intervention (see Table 2.5). To avoid practice effects there were two alternating versions of the progress checker which were administered at alternate times. Children in the TAU intervention arm did not receive the progress checker assessments during the TAU period.

Table 2. 5

Administration of the BEST Progress Checker at each time point

Progress Checker Time point	**Progress Checker Assessments Delivered (BEST with/without Sign treatment arms only).**
Following session 4	BEST Assessment Progress checker B
Following session 8	BEST Assessment Progress checker A
Following Session 12	BEST Assessment Progress checker B

2.11.3 Parent/Teacher Reported Measures

2.11.3.1 SDQ

The Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997) was used to measure Social, Emotional and Behavioural Difficulties (SEBD). The SDQ consists of 25 items on five subscales. Four of the subscales signify emotional, conduct, hyperactivity/inattention and peer relationship difficulties. These difficulties are measured by the total difficulties score. The final subscale measures prosocial behaviour.

In the student randomised controlled pilot study, parental questionnaire response rate was poor, so parents and teachers were both asked to complete the questionnaires in the present study.

The SDQ was not used as an outcome measure but was completed at baseline to characterise the sample of children with respect to the important comorbidity of social, emotional and behavioural difficulties experienced by many children with language difficulties (Botting & Conti-Ramsden, 2000).

2.11.3.2 FOCUS

The Focus on the Outcomes of Communication Under Six (FOCUS) (Thomas-Stonell et al., 2010; Washington et al., 2013) is a questionnaire designed to measure activity and participation domains defined in the International Classification of Functioning Disability and Health (World Health Organization, 2001, 2007). It provides a valid measure of real-world communication outcomes (Thomas-Stonell et al., 2010). The FOCUS 34 is a shorter version of the original 50 item questionnaire, that has been shown to reliably measure the same constructs with less items (Oddson et al., 2019). The shorter version was selected for this study in an attempt to avoid the poor questionnaire response rate observed in the student randomised controlled pilot study. There are two versions of the FOCUS 34 which probe identical constructs; one designed for parents and one for clinicians, which was given to teachers. Correspondence with the FOCUS authors confirmed that teachers who knew the child well could complete the clinician version of the questionnaire. As with the SDQ, parents and teachers were both asked to complete questionnaires to improve response rate.

2.11.4 Measure Delivery

Table 2.6 depicts when each measure was carried out during the present study.

Table 2. 6

Measures administered at each timepoint

Assessment Time point	Measures Administered
Pre-study	• Teacher questionnaire • CSCOT • Calculation of IDACI Score
Baseline	• Eligibility Assessment • NRDLS • Targeted BEST Assessment • Generalised BEST Assessment • Alternative Picture BEST Assessment • SDQ (Parent and Teacher versions) • FOCUS (Parent and Teacher versions
Following BEST Session #4	• BEST Assessment Progress checker B
Following BEST Session #8	• BEST Assessment Progress checker A
Following BEST Session #12	• BEST Assessment Progress checker B
Outcome	• NRDLS • Targeted BEST Assessment • Generalised BEST Assessment • Alternative Picture BEST Assessment • FOCUS (Parent and Teacher versions)
Follow-Up	• NRDLS • Targeted BEST Assessment • Generalised BEST Assessment • Alternative Picture BEST Assessment • FOCUS (Parent and Teacher versions)

2.11.5 Reliability of Measures

2.11.5.1 Research Assistants' Training

Research Assistants (RAs) were Undergraduate and Masters level Speech and Language Therapy students and Psychology placement students. Two of the RAs did not have a background of working with children and therefore first completed a placement at a local nursery to gain experience of working with children and of Early Years educational settings. All other RAs had extensive experience of working with children and training in administering assessments. RAs were responsible for administering the language assessments at baseline, outcome and follow-up to maintain the blinding that was essential to the integrity of the study.

The Author delivered comprehensive training to all RAs through a series of lectures and workshops covering the background of the study, principles underpinning assessment, how to deliver and score each outcome measure and the importance of maintaining blinding and treatment fidelity. RAs developed experience in delivering the assessments through structured paired role play, in which they took turns playing the role of assessor and the child. Once familiar with the assessment protocols, the RAs videoed mock role play assessments and critically reviewed the recordings in groups and with the Author. Other areas were covered in the training including when to discontinue an assessment (for example if a child became distressed), managing professional relationships with school staff, and the importance of accommodating the needs of the schools.

The Author also created a manual detailing the background and rationale of the study, as well as the assessment and scoring procedures for each assessment (see Appendix 13). This was designed to increase the reliability and consistency of the assessments. Whenever any issues of inconsistency or ambiguity arose, these were clarified using a consensus approach and added to the manual for future reference.

The Author regularly met with the RAs throughout the assessment periods to discuss any issues and to provide ongoing supervision. Blinding of the assessments was maintained at all times and any queries were anonymised so that the Author was not made aware of any specific child's score and the treatment arms were not revealed to the RAs.

SLT students were able to use this experience and portions of the data for their own Undergraduate dissertations, for which the Author was a co-supervisor.

2.11.5.2 Reliability of Face-to-Face Measures

Assessments were audio recorded using a dictaphone to increase reliability. Every day after completing assessments in schools, the RAs returned to Newcastle University and scored the assessments using the appropriate manuals (see Appendix 13 for scoring protocol). Ambiguities that could not be resolved by the manuals occasionally arose during scoring and the RAs reached a consensus on these items before proceeding. The research assistants checked all of their own assessments against the audio recordings and rectified errors or omissions. They then cross-checked 20 percent of the other research assistant's assessments. Cronbach's alphas were above eighty percent for all assessments, although had they not been, further steps would have been taken to rectify this. Research assistants then entered data into a secure online assessment spreadsheet which the Author did not have access to in order to maintain blinding.

2.11.5.3 CSCOT Training and Reliability

The Author received CSCOT training from an RA who worked on the feasibility study (Dockrell et al., 2015) and was experienced in its use. The Author first piloted the CSCOT in the Newcastle University Paediatric Speech and Language Therapy Clinic. Whilst the CSCOT was not designed for clinical observations, this ensured that the Author became familiar with the items in the tool. The Author further piloted the CSCOT in classrooms in two Local Authority schools. As part of a separate student project co-supervised by the Author, the CSCOT reliability was checked on a subset of CSCOT observations carried out by both the Author and a fully trained RA at the same time. Cronbach's Alpha of >0.8 were found for all facets of the CSCOT.

2.12 Blinding

Blinding involves preventing certain members of the research team from accessing certain information, such as intervention arm and children's test scores. This method increases robustness by removing biases known to occur when the researcher is aware of which treatment arm participants are assigned to (Nunan & Heneghan, 2018; Schulz et al., 2010).

RAs were blinded to treatment arm for the duration of the study. The following steps were taken to ensure blinding was effectively implemented and maintained:

- Initial RA training session which detailed the scientific importance of blinding for reliable outcomes.
- Each RA created an email signature to remind teachers that they were blinded to the school treatment arms.
- Teachers were asked to only discuss treatment arms with the Author and reminded of this in all email correspondence.
- RAs entered all data into a shared spreadsheet using participant codes which contained no indication of treatment arm.
- Documents detailing the assignment to treatment arm were stored securely in a locked filing cabinet that only the Author had access to, minimising the risk of accidental unblinding.

2.12.1 Blinding of Author to Assessment Data

The Author was blinded to assessment data, which was gathered by RAs, until the end of the study. Assessment data was stored securely in separate computer folders to avoid accidental unblinding. Research Assistants addressed assessment-based queries with members of the supervisory team, rather than the Author, if anonymity could not be maintained. These steps ensured that no biases were introduced by the Author knowing children's baseline scores on the language measures.

2.13 Sample Size

A power calculation was carried out to determine the required sample size for the research. The 2012 'What Works' review as part of the Better Communication Research Programme (BCRP) (Law et al., 2012) included 54 intervention trials in meta-analyses, nine of which showed positive outcomes for expressive vocabulary (effect size 0.7; 95% confidence interval [CI] 0.05– 1.25; P = .04) and syntax (effect size 0.6; 95% CI 0.15–1.95; P = .01). Effect sizes were obtained from expressive language outcomes due to a lack of available studies examining receptive outcomes. This was used to inform the current power calculation for the present study.

The power calculation for a two tailed test at $p = <0.05$ significance level (Howell, 2013) stipulated a minimum of 25 participants for each of the three intervention arms (total n = 75) for the study to be sufficiently powered.

In the randomised controlled pilot study approximately 60% of children who were approached were eligible and went on to complete the study. Based on this calculation, to achieve the desired sample size of 75 children approximately 125 children needed to be approached.

2.14 Effect Size

Cohen's D effect sizes were used to interpret the present study so that findings could be compared across outcomes and with other studies. Effect sizes were calculated in line with Cohen (1988), whereby 0.2 represents a small effect size, 0.5 represents a medium effect size and an effect size of 0.8 or above is considered large.

2.15 Important Changes to Methods

As is common for pilot studies, a number of refinements and changes were made to the trial protocol before the study began or once it had commenced. As recommended in CONSORT guidelines these changes are described below. All changes were discussed and approved by the supervisory team and the co-authors of BEST and were not deemed to affect overall study outcomes. These are discussed below and at relevant points throughout the chapter.

2.15.1 Eligibility Criteria

Initial eligibility criteria included 'The child must have been enrolled in nursery for at least one term' to exclude children who may be struggling with adjusting to nursery and being away from their caregivers for prolonged periods for the first time. The criterion was eventually amended as it posed difficulties for recruiting many otherwise eligible children that had recently started nursery, particularly for the second data wave which began in September 2018. Instead, teachers were asked to only suggest children who they believed were well adjusted enough to take part. The same ethical procedures were followed throughout the study, whereby if any child became distressed or did not want to take part they would be returned to their class and withdrawn from the study. During the study there were no children who were unwilling to take part.

2.15.2 Matching Approach

The initial matching protocol stipulated that the schools were to be matched on oral language environment (CSCOT) and income deprivation affecting children (IDACI). Due to an error in the school matching process, IDACI score was not taken into account when matching the second wave of schools (see 2.6.1.4). Despite this, the groups did not have statistically significantly different IDACI scores overall, so the change to the protocol was not considered to have affected the study findings. Furthermore, the analytical methods accounted for any clustering within the treatment arms (see 2.17.1.1.3).

2.15.3 Treatment Waves and Allocation to Treatment Arms

In the first data wave, schools were matched based on CSCOT scores among other characteristics. Despite efforts to maintain groups that did not differ statistically significantly at baseline, attrition of several schools before the study began meant that the groups differed.

In the second wave the procedure was therefore altered slightly to account for this. Schools who chose to continue in the study remained in their original treatment arm to avoid contamination effects and were not re-evaluated on the CSCOT.

New schools were evaluated on the CSCOT and then added to the existing groups on a case by case basis to maintain the homogeneity of CSCOT scores and IDACI scores across the three arms.

Following completion of the second data wave there were very few children in the TAU arm. This was likely due to school attrition before the study began and the perceived lesser benefit to taking part if the children were not receiving therapy. In an attempt to retain overall homogeneity of groups, as well as even numbers of children across the three treatment arms, it was decided that the final wave would be solely TAU. In this wave therefore, schools were approached specifically to take part in the TAU wave. CSCOT score and other demographic information were still obtained so that similarity at baseline of groups could be analysed. These results are presented in chapter three.

2.15.4 BEST Delivery when a TA was not Present

Whilst schools were requested to provide a TA to support intervention sessions, this was sometimes not possible due to school timetabling or staff shortages. The intervention fidelity ratings carried out by BEST co-author Dr Sean Pert identified a minor difference between how BEST with Sign was delivered when there was a TA present compared to no TA present. The signing was presented at a different time depending on whether the TA was present. The deviation is described in the table 2.7.

This issue affected approximately half of the schools, and some schools were never able to provide a TA whilst others could either sometimes or all of the time.

Table 2. 7

Discrepancy identified for BEST delivery when conducted with and without a TA present for BEST with and without Sign treatment arms

	TA present	TA not present
BEST with sign	The Author said and signed the sentence while the TA acted it out with toys (all simultaneously)	The Author said and acted out the sentence simultaneously, then signed the sentence directly afterwards
BEST without sign	The Author said the sentence while the TA acted it out with toys simultaneously	The Author said the sentence while acting it out with the toys simultaneously

Note: Target sentences were said once only regardless of condition and presence of TA.

There was no protocol in place for how to deliver BEST without the support of a TA as the issue was not anticipated prior to the study and not detailed in the manual. It was therefore agreed that the method identified in table 2.7 was satisfactory since it did not affect the input received during the intervention.

2.15.5 Shortening BEST sessions to increase children's attention

The BEST manual protocol stipulates that each child should have a turn manipulating the toys following each of their target sentences, so that every child acts out every sentence. This method resulted in long sessions and children's attention was poor, particularly with the largest groups.

It was therefore decided that one child would act out the first sentence, the second child would act out the second sentence, moving around the group so that each sentence was acted out by only one child and all children had several opportunities to manipulate the toys.

It was also decided that to maximise the children's exposure to the toys there would be five minutes of free play following each session. To avoid differential exposure to input at this point, the Author did not comment on what the children were doing during the free-play sessions.

2.16 Data Analysis

2.16.1 Data Entry and Cleaning to Ensure Integrity and Reliability

The data were managed, entered and cleaned in such a way as to ensure the integrity and reliability of the results. Assessment data were entered into the assessment spreadsheet by research assistants directly following testing sessions in schools and were not accessible to the Author to maintain blinding. The Author entered the progress check and intervention response data into a separate spreadsheet. Following completion of the research, the Author merged the two spreadsheets and made checks against the original datasets to ensure accurate merging.

Any duplicate data were removed and missing numerical data was labelled with NA. Syntax and structural errors such as inconsistent entry of dates of birth and ages and inconsistent capitalisation were rectified.

2.16.2 Analysis software

The data were initially inputted in Microsoft Excel and following cleaning were imported into R in CSV (comma-separated values) format. Data analysis was carried out using R (R Core Team, 2019).

2.17 Analytical Methods

Baseline comparisons were first conducted to assess the similarity of the groups before the study began. Outcome measures assessing expressive and receptive language (NRDLS), targeted sentence content and morphology (Targeted BEST Assessment) and generalised sentence content and morphology (Generalised BEST Assessment) were used to address research questions 1A-1C about efficacy and the role of sign. Longitudinal multi-level modelling was used to account for clustering and lack of independence in the data. This method was also supported by the implementation of ANCOVA to provide simple estimates of group differences. Chapter four addresses the qualitative methods employed to answer research questions 2.

2.17.1 Research Questions

2.17.1.1 Is an intervention underpinned by usage-based theory (BEST, with and without Sign) more effective than treatment as usual (TAU) for improving language and associated functional communication outcomes in pre-school children with language difficulties?

To address this research question analyses were conducted using the following outcomes: NRDLS Production score, NRDLS Comprehension score, Targeted BEST content percentage correct, Targeted BEST Morphology percentage correct, Generalised BEST content percentage correct and Generalised BEST Morphology percentage correct.

The three intervention arms were compared to establish whether children in each of the groups made differential progress. Scores pre, immediately post and six weeks post intervention were compared.

2.17.1.1.1 Baseline Comparisons

Analysis of Variance (ANOVA) baseline comparisons were conducted to assess the success of the matching procedure in creating three equivalent groups. Comparisons were conducted for the following characteristics: sex, language profile, age, level of deprivation, classroom oral language and dosage. Tukey's HSD post-hoc testing was used to establish where any significant differences lay.

2.17.1.1.2 Multi-Level Modelling

Multi-level modelling (MLM) is a statistical technique used to account for the hierarchical structure of data where there is clustering, for example in studies examining groups of children within multiple schools (Finch et al., 2019). It avoids the higher likelihood of intra-cluster correlation and type I errors. Although the matching procedure sought to avoid these issues, this method provides additional robustness in this respect. MLMs can also cope with missing data unlike traditional statistical approaches such as ANOVA, meaning that incomplete cases could still be analysed and complex procedures of data imputation could be avoided.

Longitudinal MLMs apply many of the same principles as traditional MLMs, but are carried out on person-period (long-format) data whereby time variant factors are stacked on top of each other using the stack() command in R (R Core Team, 2019), while time invariant factors are duplicated for each participant (Finch et al., 2019).

2.17.1.1.3 Multi-Level Model Employed in this Research

The maximal parsimonious model contains the following variables as fixed effects: time point, intervention arm, SDQ score, dosage, language profile, age at entry, an interaction between intervention arm and language profile, and an interaction between intervention arm and SDQ score. The model also contains random slopes and intercepts for school and participant. The model was built to account for the theoretically important data and to account for clustering of repeated measures and at the level of the individual and the school. The code for the maximal theoretical model can be found in Appendix 17. A list of all variables and their descriptions can be found in Appendix 18 and the rationale for the inclusion of each variable is available in Appendix 19.

2.17.1.1.4 Non-convergence of multi-level models

Although MLMs are powerful statistical models that hold many advantages over traditional Analysis of Variance approaches, they can face convergence issues (Finch et al., 2019). A protocol for the systematic simplification of the models was devised to counter these issues following an a-priori principled method if they arose. Non-essential variables, defined as those that were not central to the research questions, were removed. High levels of missing data also made variables candidates for exclusion if non-convergence occurred. The procedure is summarised in table 2.8 and reporting in chapter three makes it clear when variables were omitted.

Table 2. 8

Priority for exclusion of variables and reasons for potential exclusion

Priority for Exclusion (highest to lowest)	Variable	Rationale
1	Fixed effect: SDQ score	Teacher response rates were generally poor for the SDQs, hence the amount of missing data relative to the small sample size may be a contributing factor towards the lack of convergence. The model can still address the majority of research questions without taking into account the SDQ score
2	Fixed effect interaction: Intervention Arm : SDQ Score	For reasons described above the interaction between intervention arm and SDQ score may contribute to the model failure to converge
3	Fixed effect interaction: Intervention Arm : Language Profile	While possible, interactions may cause difficulties for the model given the small sample size
4	Random slopes for school: SDQ Score	SDQ score is less likely than other variables to converge for the reasons outlined previously in this table. It is also not involved in the main research questions (1A & 1B)
5	Random slopes for school: Language Profile	Language profile is less important than time, dosage and age because it is not involved in the main research questions (1A & 1B)
6	Random slopes for school: Age in months at entry	Although more important than SDQ score and language profile, if necessary, age in months at entry could be removed as it is also a fixed effect and children's ages varied by no more than a year, hence age is not expected to impact greatly on performance
7	Random slopes for participant: SDQ Score	As with SDQ score as a random slope for school, SDQ score as a random slope for participant is less likely than other variables to converge for the reasons previously discussed. It does not address the main research questions (1A & 1B)
8	Random slopes for participant: Language Profile	As with language profile as a random slope for school above, language profile as a random slope for participant would be removed before time, dosage and age because it is not implicated in the main research questions (1A & 1B)

9	Random slopes for participant: Dosage	Although implicated in the main research questions, the question of dosage can be separated from that of efficacy, hence it will be removed if necessary

2.17.1.2 Is signing a necessary 'active ingredient' of BEST, and does it differentially affect progress on different language goals (sentence structure and morphology)?

The methodology described for question 1A above was also used to assess whether signing was an 'active ingredient' of BEST. Coefficients and effect sizes were compared for outcomes across the three intervention arms allowing direct comparison of BEST with and without sign.

2.17.1.3 What are the point and interval estimates of effect sizes across outcomes?

In line with current reporting standards (Crawford et al., 2010), point and interval effect sizes were calculated yielding *d* statistics (J. Cohen, 1988; Nakagawa & Cuthill, 2007). Effect sizes were interpreted using Cohen's categories (1988), whereby 0.2 represents a small effect size, 0.5 represents a medium effect size and an effect size of 0.8 or above is considered large. Standardised betas are also provided for all outcomes so that the relative magnitude of effects can be compared.

2.17.1.4 What theoretically motivated recommendations can be made from applying Behaviour Change Theory to stakeholder consultation outcomes, to increase acceptability of the research and therefore inform the next stage of BEST development?

A qualitative approach was taken to triangulate the findings of this study. Stakeholder consultation data gathered from teachers who took part in the study was analysed using Thematic Analysis (Braun & Clarke, 2006a). Codes were generated for the emerging themes which were grouped, refined and named. The final themes were then linked to the theoretical domains framework and applied to behaviours from the Capability, Opportunity Motivation-Behaviour (COM-B) model (Cane et al., 2012; Michie et al., 2005; Michie et al., 2011). The qualitative analysis is reported in chapter four.

Chapter 3

Results

3.1 Introduction

This chapter addresses the quantitative research questions pertaining to the efficacy of BEST. Section 3.2 reports recruitment and retention data about the study participants. Section 3.4 presents the results for the main study outcomes and addresses the use of sign. Finally, section 3.7 and 3.8 presents additional analyses of sensitivity and dosage.

The following research questions are addressed in this chapter:

1A) Is an intervention underpinned by usage-based theory (BEST, with and without sign) more effective than treatment as usual (TAU) for improving language and associated functional communication outcomes in pre-school children with language difficulties?

1B) Is signing a necessary 'active ingredient' of BEST, and does it differentially affect progress on different language goals (sentence structure and morphology)?

1C) What are the point and interval estimates of effect sizes across outcomes?

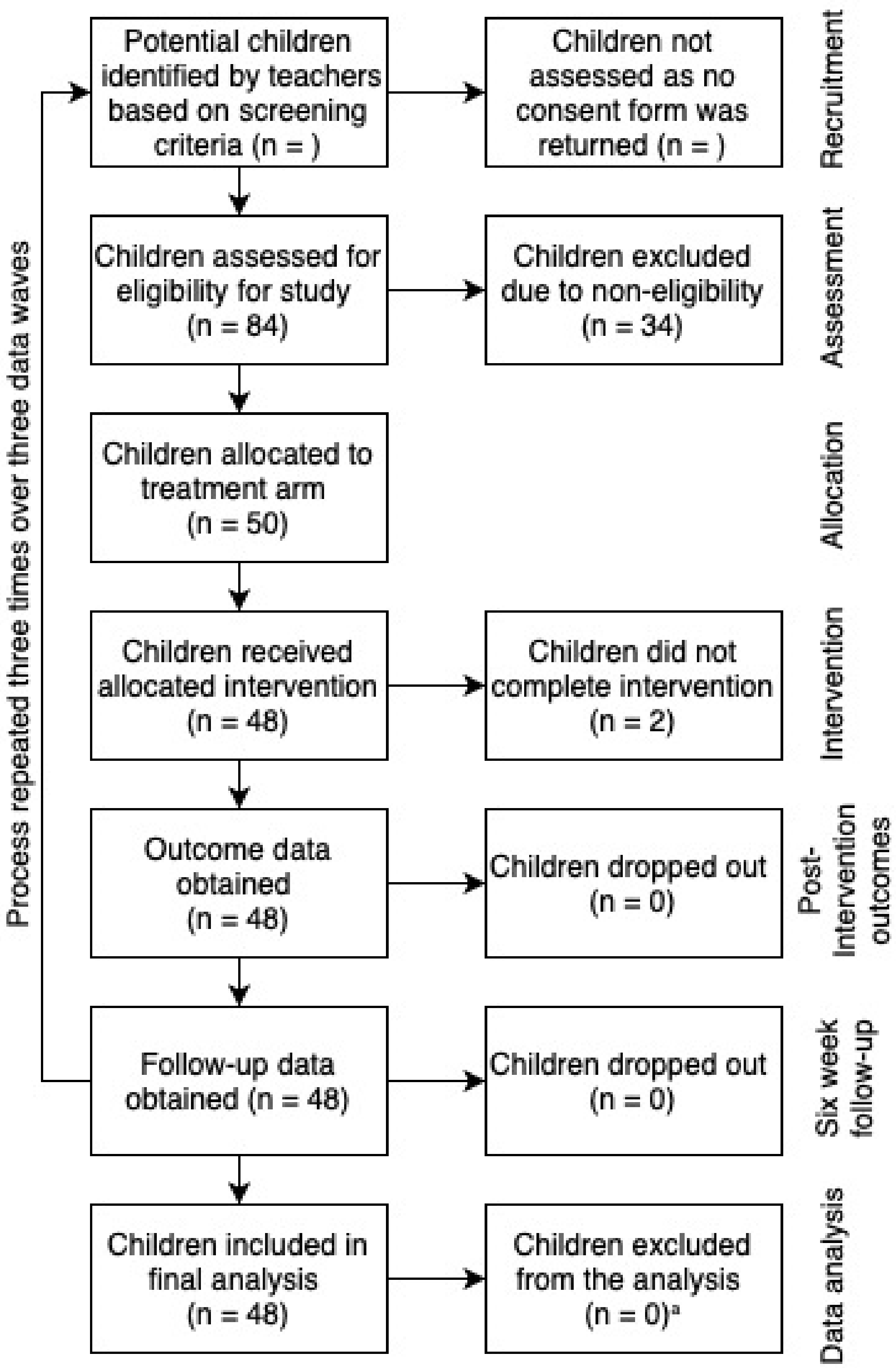
3.2 Recruitment, Retention and Participants

3.2.1 Recruitment

Figure 3.1 below shows the participant recruitment data for the study. Teachers distributed consent forms to children who they felt would benefit from additional language support (see 2.6.2.1 for full eligibility criteria and selection process). Ten forms were given to each school involved at each wave of the study (n=140 forms). However, it is unclear how many parents were initially approached; some teachers reported not sending all the forms home, whilst others reported photocopying and distributing more. It was therefore not possible to obtain exact numbers of children initially approached. Eighty-four consent forms were returned within the specified timeframe. Those 84 children were assessed, and 50 were eligible whilst 34 were found to be ineligible to take part. Children were deemed ineligible due to either: their age being outside of the 3;5-4;5 range; scoring above the 16th percentile on the receptive and expressive subscales of the NRDLS; having a clinical diagnosis; or parental report of English as the child's second, non-dominant language. Figure 3.1 displays recruitment and retention figures. The numbers reported are the total number of participants across the three data waves.

Figure 3. 1

Participant flow chart showing the participant journey through the study



[a] *Data was missing for some participants for certain measures. Missing data is discussed in detail below.*

3.2.2 Retention

The retention rate was 96%. Two children of the initial 50 did not complete the study. One child was revealed to be receiving intensive Speech and Language Therapy and it was decided that they would be too tired if also taking part in the BEST study. The other left the study part way through as the school no longer wished to take part in the study (only one child from that school was taking part). Data were destroyed for these participants.

3.2.3 Completion of Assessments

Participants were assessed at baseline, outcome and follow-up. Where applicable participants also completed the BEST Assessment progress checks (see table 2.6).

Information on completed assessments at each timepoint is summarised in table 3.1 below.

NRDLS assessment completion rate was 99%. The BEST Assessments were 95% complete, with missing data due to child absence and a child moving school. Despite best efforts, the FOCUS teacher questionnaires were 36% completed, and parental FOCUS responses were 28% complete.

Table 3. 1

Percentage of completed assessments at each timepoint

Outcome Measure	Number (%) of completed assessments (n)[a]
Baseline NRDLS (Comprehension and Production)	48 (100%)
Baseline Targeted BEST Assessment (Content and Morphology)	47 (98%)
Baseline Generalised BEST Assessment (Content and Morphology)	46 (96%)
Baseline Alternative Picture Version[b]	24 (50%)
Baseline Teacher SDQ	25 (52%)
Baseline Parent SDQ	25 (52%)
Baseline Teacher FOCUS	26 (54%)
Baseline Parent FOCUS	24 (50%)
Outcome NRDLS (Comprehension and Production)	48 (100%)
Outcome Targeted BEST Assessment (Content and Morphology)	47 (98%)
Outcome Generalised BEST Assessment (Content and Morphology)	48 (100%)
Outcome Alternative Picture Version[b]	25 (52%)
Outcome Teacher FOCUS	17 (35%)
Outcome Parent FOCUS	10 (21%)
Follow-Up NRDLS (Comprehension and Production)	47 (98%)
Follow-Up Targeted BEST Assessment (Content and Morphology)	43 (90%)
Follow-Up Generalised BEST Assessment (Content and Morphology)	43 (90%)
Follow-Up Alternative Picture Version[b]	43 (90%)
Follow-Up Teacher FOCUS	9 (19%)
Follow-Up Parent FOCUS	7 (15%)

Note. Number of children completing individual measures at each assessment point (% of completed assessments)

[a] Number of assessments out of a total possible 48. Data were destroyed for children who left the study during intervention delivery (n = 2).

[b] The Alternative Picture Version (APV) assessment was not analysed further due to 50% missing data.

3.3 Baseline Characteristics

3.3.1 Descriptive statistics at Baseline

Table 3.2 displays descriptive statistics for the sample across the three intervention arms for variables used to match the schools (e.g. CSCOT and IDACI score) and other demographic factors (e.g. sex, language profile and age). Intervention arms were created by matching demographic data to create three matched groups of schools that were approximately equivalent (see chapter two for the matching procedure). It was therefore anticipated that the groups would be equivalent at baseline on these variables.

Comparisons of these variables and additional demographic factors across the three data waves and across the 13 schools are also presented in Appendix 20.

3.3.1.1 Checking Normality of the data

Prior to testing whether treatment arms differed significantly on these baseline characteristics they were tested for normality using Shapiro-Wilk tests of normality to identify the correct statistical test. Income Deprivation Affecting Children Index (IDACI) and Communication Supporting Classrooms Observational Tool (CSCOT) score were found not to be normally distributed (included in Appendix 20) whilst other variables were normally distributed. ANOVAS were therefore conducted for all variables and Kruskal-Wallis Chi-Square tests of difference were conducted for IDACI and CSCOT scores. These confirmed the ANOVA test findings. ANOVA results, which are robust to many violations are therefore reported in table 3.2 below.

Table 3. 2

Descriptive statistics at baseline across Full Sample and Individual Treatment Arms (BEST with sign, BEST without sign and Treatment as Usual (TAU))

	Full Sample M(SD)		Treatment Arm						Between Groups ANOVA[e]/X^2
	N	M(SD)	BEST with sign		BEST without sign		TAU		
			N	M(SD)	N	M(SD)	N	M(SD)	
N	48	-	17	-	17	-	14	-	-
Sex									[F(2,45) = 3.52, p<.04]*
Male	28	-	14	-	7	-	7	-	
Female	20	-	3	-	10	-	7	-	
Language Profile[a]									[F(2,45) = 0.68, p>.51]
Expressive	15	-	6	-	6	-	3	-	
Receptive	5	-	1	-	3	-	1	-	
Mixed	28	-	10	-	8	-	10	-	
Age in months at entry	-	46.96 (6.13)	-	46.65(6.99)	-	47.71(5.67)	-	46.43(5.92)	[F(2,45) = 0.19, p>.82]
Level of Deprivation	-	0.26 (0.17)	-	0.31(0.14)	-	0.19(0.16)	-	0.29(0.19)	[F(2,45) = 2.38, p>.11]

Affecting Children[b]									
Classroom Oral Language Environment[c]	-	0.49 (0.16)	-	0.52(0.18)	-	0.54(0.12)	-	0.40(0.14)	[$F(2,45) = 3.51$, $p<.04$]*
Dosage[d]	-	9.52 (6.47)	-	13.65(2.47)	-	13.24(2.22)	-	-[f]	[$X^2(1) = 0.01$, $p=.937$]*

[a] Mixed language profile defined as scoring on or below the 16th centile of the production and comprehension subscales of the NRDLS. Expressive-only language difficulties defined as scoring on or below the 16th centile of the production subscale only. Receptive-only difficulties defined as scoring on or below the 16th centile of the comprehension subscale only.

[b] Income Deprivation Affecting Children Index (IDACI) Score. IDACI score was calculated using the school postcode at: https://lginform.local.gov.uk/reports/lgastandard?mod-metric=3910&mod-area=E92000001&mod-group=AllRegions_England&mod-type=namedComparisonGroup

A higher IDACI score indicates a postcode area of higher deprivation (see the measures section of chapter 2 for more information) (Department of Communities and Local Government, 2015).

[c] Communication Supporting Classrooms Observation Tool (CSCOT) Proportion Total Score (Dockrell et l., 2012). Total proportion score calculated from a one-hour classroom observation.

[d] Dosage refers to the average number of BEST sessions received out of a total possible 16. Children in the TAU arm do not have a dosage score as they did not receive the intervention.

*[e] Significant difference between groups (BEST with sign, BEST without sign, TAU) indicated by *. Calculated using ANOVA tests of difference, reported below.*

[f] No dosage scores are presented for TAU as they did not receive intervention during the study. Chi square Significance testing was conducted on the two intervention arms only.

ANOVAS were conducted to establish whether there were any significant differences between number, sex, age, classroom oral language environment and dosage for children across the intervention arms at baseline (see right hand column of table 3.2).

There was a significant difference between sex at baseline across the three treatment arms [$F(2,45) = 3.52$, $p = <.04$]. Tukey's HSD post-hoc tests revealed a significant difference between BEST with sign and BEST without sign ($p <.04$) such that BEST with Sign had a higher proportion of boys than BEST without Sign. There were no significant differences between BEST with sign and TAU ($p = .15$) or BEST without sign and TAU ($p = .86$).

There was no significant difference between language profile at baseline across the three treatment arms [$F(2,45) = 0.68$, $p>.51$].

There was no significant difference between age in months at entry at baseline across the three treatment arms [$F(2,45) = 0.19$, $p>.82$].

There was no significant difference between school IDACI scores across the three treatment arms [$F(2,45) = 2.38$, $p>.11$].

There was a significant difference between CSCOT Proportion total scores at baseline across the three treatment arms [$F(2,45) = 3.513$, $p = 0.0382$]. Tukey's HSD post-hoc tests revealed a significant difference between the TAU and BEST without sign arms at baseline ($p = <.05$, where BEST without sign CSCOT total proportion score was higher by around 0.12 points) suggesting that the TAU schools have fewer communication supporting environments, interactions and opportunities than the BEST without sign treatment arm schools. There were no significant differences between TAU and BEST with sign ($p = >.09$), or BEST with sign and BEST without sign ($p = >.94$).

There was no significant difference between the two BEST intervention arms for dosage [$X^2(1) = 0.01$, $p=.937$]. TAU was excluded as the dosage was 0 in this arm.

3.4 Results

3.4.1 Descriptive Statistics

Table 3. 3

Descriptive statistics for baseline, outcome and follow-up scores for production, comprehension, content and morphology scores

Outcome Measure	Baseline M(SD)			Outcome M(SD)			Follow-up M(SD)		
	BEST with sign	BEST without sign	Treatment as Usual	BEST with sign	BEST without sign	Treatment as Usual	BEST with sign	BEST without sign	Treatment as Usual
NRDLS[a] Production Standard Score	76.94 (6.49)	79.24 (10.65)	76.36 (7.91)	87.06 (14.05)	84.35 (9.67)	76.79 (9.02)	89.75 (13.64)	83.94 (10.59)	80.07 (11.38)
NRDLS[a] Comprehension Standard Score	82.41 (7.49)	81.00 (6.84)	81.21 (7.97)	91.59 (12.42)	87.41 (10.39)	83.79 (9.67)	93.44 (12.91)	86.88 (9.49)	87.71 (13.80)
Content (Targeted BEST Assessment[b])	44.96 (14.08)	43.49 (11.51)	48.98 (15.82)	77.86 (16.54)	78.43 (18.37)	49.85 (14.07)	78.68 (15.29)	74.20 (11.09)	58.63 (17.62)

Percentage Scores									
Morphology (Targeted BEST Assessment[b]) Percentage Scores	37.57 (19.32)	34.42 (19.54)	36.54 (21.11)	74.53 (22.06)	69.97 (24.13)	48.83 (22.46)	77.23 (18.54)	71.98 (22.11)	55.82 (24.11)
Content (Generalised BEST Assessment[c]) Percentage Scores	36.00 (13.76)	32.22 (14.06)	40.48 (16.30)	45.58 (14.51)	40.68 (10.80)	41.22 (14.89)	54.04 (18.19)	49.78 (14.31)	40.33 (17.59)
Morphology (Generalised BEST Assessment[c]) Percentage Scores	33.79 (16.31)	32.14 (19.23)	33.37 (22.26)	57.59 (25.35)	53.81 (23.03)	46.80 (22.65)	68.07 (20.78)	63.33 (18.90)	51.07 (26.01)

[a]*New Reynell Developmental Language Scales (Edwards et al., 2011)*

[b]*Targeted BEST Assessment (McKean et al., 2013)*

[c]*Generalised BEST Assessment designed by the Author (see Appendix 11)*

Table 3.3 displays descriptive statistics for outcome measures across baseline, outcome and follow-up time points. Between baseline and outcome there are considerable mean increases for both BEST arms on most measures, and a smaller increase for TAU. Maintenance with slight mean increases are seen for all intervention arms, but the TAU arm scores remain lower than the BEST arms on average.

3.4.2 Longitudinal Multi-Level Models

A longitudinal analysis was used to account for repeated measurements conducted with the same children at multiple timepoints. The data was first transformed into person-period (long-format) structure so that time invariant factors were duplicated for each participant, while time variant assessment sores were stacked using the stack() command in base R (Finch et al., 2019; R Core Team, 2019).

A data frame containing the variables and outcomes relevant to the research questions was constructed containing NRDLS Comprehension and Production standard scores, Targeted BEST Assessment Content Correct and Morphology Correct percentage scores and Generalised BEST Assessment Content Correct and Morphology Correct percentage scores. Also included were participant number, treatment arm, school, age in months, language profile, IDACI score, dosage (number of sessions received), time point and further variables necessary for the analysis (see Appendix 18 for descriptions).

Although this study had aimed to include SDQ and FOCUS score in the analysis, the poor questionnaire response rate meant that this was not possible (see table 3.1). In addition, the Alternative Picture BEST Assessment was excluded from analyses because only 50% of children completed it. Finally, the BEST Assessment Language Probes taken at week 4, 8 and 12 were not analysed in this study as trends were clear from the present analysis, but will be analysed in future. These variables were therefore excluded from the analysis.

3.4.2.1 The Maximal Theoretical Multilevel Model

In line with recommendations from Barr et al. (2013) the maximal theoretical longitudinal models were stipulated for each outcome measure. The theoretical maximal model is as follows:

lme(Score ~ Time * Intervention Arm + Age + Dosage + IDACI Score +

Language Profile ,

random = list(~ Time + Age + Dosage + Language Profile | School,

~ Time + Language Profile | Ppt No)

In these models, score is predicted by the following fixed effects: time point (baseline, outcome and follow-up), allocated intervention arm, the number of sessions received, language profile (expressive, receptive or mixed difficulties) determined by scores on the NRDLS, age in months at entry into the study, and interaction between intervention arm and time point.

Random intercepts for school are predicted by the following random slopes: timepoint, age in months at entry, number of sessions received, and language profile.

Random intercepts for participant are predicted by the following random slopes: timepoint and language profile.

Further rationale for the theoretical maximal model is included in Appendix 19.

3.4.2.2 Maximal Converging Model

The maximal theoretical model did not converge. The process outlined in Table 2.8 was therefore used to sequentially remove variables until the maximal parsimonious converging model was reached. This is also outlined in Table 3.4 below. Only two variables needed to be removed to reach model convergence, and the rationale for this is presented in table 2.8.

Table 3. 4

Order in which variables were removed from the maximal model to reach convergence

Order of Removal	Variable to be Removed
1	Random slopes for school: Language Profile
2	Random slopes for participant: Language Profile

Note. Following the procedure outlined in Table 2.8, the variables were removed in order, beginning with the least important, until the model reached convergence

The maximal converging model was therefore:

```
lme(Score ~ Time * Intervention Arm + Age + Dosage + IDACI Score +
Language Profile,
random = list(~ Time + Dosage + Age| School,
~ Time | Ppt No)
```

In this model, the score is predicted by the following fixed effects: time point (baseline, outcome and follow-up), allocated intervention arm, interaction between timepoint and intervention arm, age, the number of sessions received, IDACI score and language profile (expressive, receptive or mixed difficulties) determined by scores on the NRDLS.

Random intercepts for school are predicted by the following random slopes: timepoint, number of sessions received, and age in months at entry. Random intercepts for participant are predicted by random slopes for timepoint.

Table 3.5 below displays results of the multilevel models for the six main outcome variables: production, comprehension, targeted sentence content, targeted sentence morphology, generalised sentence content and generalised sentence morphology.

The parameters of interest are the interaction terms between time and intervention arm for each outcome, as these results demonstrate the efficacy of BEST. These are displayed in **bold text**. The results for each outcome are subsequently discussed in more detail. The intercept represents Treatment as Usual with a score of 0 on all covariate measures and is used to interpret the parameters of interest.

3.4.3 Multi-Level Model Results

Table 3. 5

Multilevel Results for Production, Comprehension, Content and Morphology, Generalised Content and Generalised Morphology Outcomes (Beta [95% CIs])

	Multilevel Model					
	Production Maximal Model[a]	Comprehension Maximal Model[b]	Targeted Content Maximal Model[c]	Targeted Morphology Maximal Model[c]	Generalised Content Maximal Model[d]	Generalised Morphology Maximal Model[d]
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	74.478***	77.396***	-12.549***	-19.444***	-2.987	-14.011***
	[71.59, 77.37]	[73.01, 81.78]	[-18.89, -6.21]	[-28.59, -10.30]	[-10.83, 4.85]	[-23.46, -4.56]
Outcome	0.452	4.881	0.495	12.294**	1.422	12.817**
	[-4.29, 5.19]	[-2.50, 12.26]	[-8.79, 9.78]	[1.46, 23.13]	[-5.71, 8.55]	[2.12, 23.51]
Follow-up	3.755	8.634**	8.845*	19.046***	-0.142	16.041**
	[-2.41, 9.92]	[0.91, 16.36]	[-0.74, 18.43]	[6.86, 31.24]	[-9.01, 8.73]	[3.68, 28.40]

BEST without sign at Baseline	-1.724	-2.915	-13.050	-3.995	-17.313	-11.820
	[-11.71, 8.26]	[-13.32, 7.49]	[-34.80, 8.70]	[-35.85, 27.86]	[-40.91, 6.29]	[-44.11, 20.47]
BEST with sign at Baseline	-0.498	-1.470	-8.953	6.601	-10.583	-2.781
	[-10.73, 9.73]	[-11.92, 8.98]	[-31.10, 13.20]	[-25.74, 38.94]	[-34.36, 13.20]	[-35.32, 29.76]
Age (centred)	-0.210*	0.130	0.789**	0.282	0.778*	0.406
	[-0.45, 0.03]	[-0.22, 0.48]	[0.20, 1.37]	[-0.85, 1.41]	[0.02, 1.53]	[-0.70, 1.52]
Dosage	0.154	0.189	0.291	-0.176	0.613	0.452
	[-0.56, 0.86]	[-0.50, 0.87]	[-1.24, 1.82]	[-2.47, 2.12]	[-1.02, 2.24]	[-1.84, 2.74]
IDACI (centred)	-5.936	-19.277***	-34.037***	-51.255***	-22.246*	-48.830***
	[-14.36, 2.48]	[-29.87, -8.68]	[-52.46, -15.61]	[-79.62, -22.89]	[-41.91, -2.58]	[-76.51, -21.15]
Expressive Profile at Baseline	0.494	10.396***	1.655	-10.248*	-1.287	-10.340*
	[-2.54, 3.53]	[7.85, 12.95]	[-4.81, 8.12]	[-20.39, -0.11]	[-8.35, 5.77]	[-20.59, -0.09]
Receptive Profile at Baseline	21.814***	-1.107	0.517	9.940	-0.763	5.543
	[17.09, 26.53]	[-5.42, 3.21]	[-9.71, 10.74]	[-5.92, 25.80]	[-12.04, 10.51]	[-10.30, 21.39]
BEST without sign at Outcome[e]	**5.322**	**2.745**	**34.722*****	**22.418*****	**6.076**	**8.513**
	[-1.12, 11.77]	**[-7.37, 12.86]**	**[22.12, 47.32]**	**[7.67, 37.17]**	**[-3.72, 15.87]**	**[-6.10, 23.13]**
BEST without sign at Follow-up[e]	**2.122**	**-1.769**	**23.269*****	**18.851****	**17.538*****	**15.926***

	[-6.31, 10.56]	**[-12.40, 8.86]**	**[10.19, 36.35]**	**[2.04, 35.66]**	**[5.26, 29.81]**	**[-0.94, 32.79]**
BEST with sign at Outcome[e]	**10.139*****	**6.671**	**32.089*****	**23.817*****	**7.585**	**11.206**
	[3.68, 16.60]	**[-3.58, 16.92]**	**[19.41, 44.77]**	**[9.02, 38.61]**	**[-1.98, 17.15]**	**[-3.18, 25.59]**
BEST with sign at Follow-up[e]	**9.435****	**4.123**	**21.953*****	**16.574***	**15.517****	**13.401**
	[0.93, 17.94]	**[-6.66, 14.91]**	**[8.58, 35.33]**	**[-0.69, 33.84]**	**[2.93, 28.11]**	**[-3.82, 30.62]**
Observations	143	143	137	137	137	137
Log Likelihood	-472.041	-480.081	-537.225	-581.596	-535.747	-585.683
Akaike Inf. Crit.	1,016.082	1,032.161	1,146.451	1,235.193	1,143.495	1,243.365
Bayesian Inf. Crit.	1,122.744	1,138.823	1,251.570	1,340.312	1,248.614	1,348.484

*Note: *p<0.1; **p<0.05; ***p<0.01*

[a] *New Reynell Developmental Language Scales (Edwards et al., 2011) Production Subscale*

[b] *New Reynell Developmental Language Scales (Edwards et al., 2011) Comprehension Subscale*

[c] *Targeted BEST Assessment (McKean et al., 2013)*

[d] *Generalised BEST Assessment (Designed by the Author)*

[e] *Emboldened text displays results for the parameters of interest (interaction terms for intervention arm and timepoint)*

The results are presented below. Scatter plots were created for intervention arm, time point, and both interaction and timepoint using ggplot2 (Wickham, 2016). Violin plots, a geometry from the ggplot2 package in R were added to the scatter plots to represent density trends in the data, whereby the kernel density (the horizontal width of the plot at any given point) demonstrates the density of the data. The widths of the plots are greater around the means, for example, as more data is situated there than at the top of the plots, which are narrower. The data points were jittered to allow clearer demonstration of where the points lie.

Violin plots for the pooled data and then separately for each intervention arm at baseline, outcome and follow-up are presented. Following this, children's individual trajectories with group means are illustrated. Relevant analyses from the multi-level models and effect size calculations are presented.

These results are presented in turn for NRDLS production standard scores, NRDLS comprehension standard scores, Targeted BEST Assessment content, Targeted BEST Assessment morphology, Generalised BEST Assessment content and Generalised BEST Assessment morphology correct percentage scores.

Missing data were not included in the plots, and the numbers of included and missing children at each assessment point can be found in Table 3.1 above. Missing numbers are also reported in the note below each graph.

3.4.3.1 Effect Sizes

All outcomes were standardised (mean = 0, SD = 1) using the scale() and centre() functions in base R (R Core Team, 2019), and compared to Treatment as Usual which represents the reference category in Table 3.5 above. Effect sizes were interpreted in line with Cohen (1988), whereby 0.2 represents a small effect size, 0.5 represents a medium effect size and an effect size of 0.8 or above is considered large.

3.4.3.2 Sensitivity Analysis

Following the main results, a sensitivity analysis is presented. This post-hoc analysis was introduced ensure that the results were robust to the scores of children who did not have demonstrate a particular impairment on the relevant measure at baseline.

Children with a receptive-only profile necessarily scored above the 16th percentile on production at baseline and children with an expressive-only profile necessarily scored above the 16th percentile on comprehension at baseline (see Table 2.4)

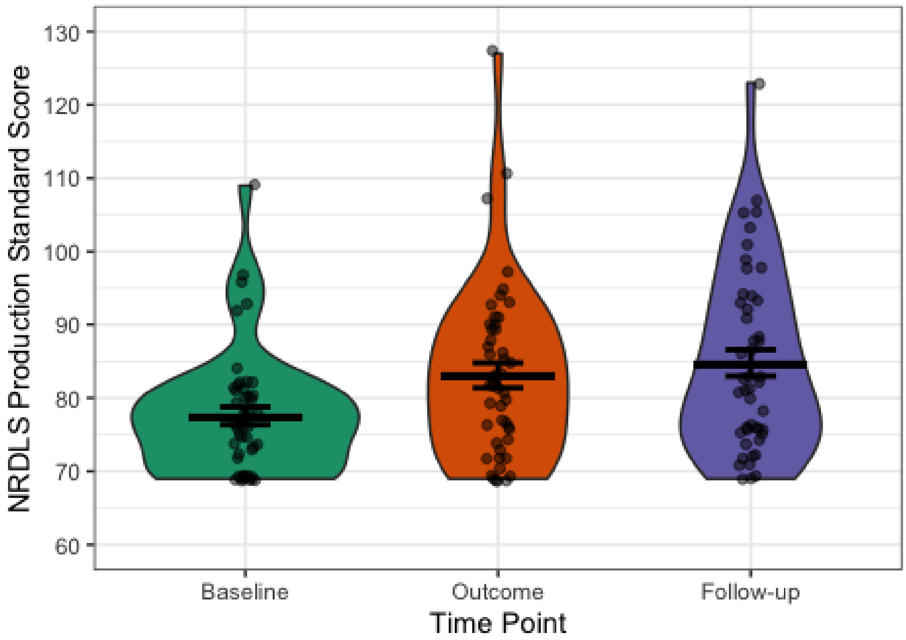
As a result, there are a subset of children included in each main analysis who did not demonstrate an impairment on or below the 16th percentile for the measure in question at baseline. Improving performance on this measure was not the primary focus of the intervention for these children.

As a sensitivity analysis, children with no impairment in expressive areas (the receptive-only language profiles) were excluded from measures of expressive outcomes. The analysis of expressive outcomes was conducted again with only the sample showing initial expressive difficulties (mixed and expressive-only language profiles). The same procedure was then repeated for receptive measures, removing the children with expressive-only difficulties. This analysis therefore demonstrates progress made on each measure only by children with an existing impairment in that domain.

3.4.4 Production (NRDLS Production Scale Standard Scores)

Figure 3. 2

NRDLS Production Standard Scores over three time points (baseline, outcome and follow-up), collapsed across intervention arms.

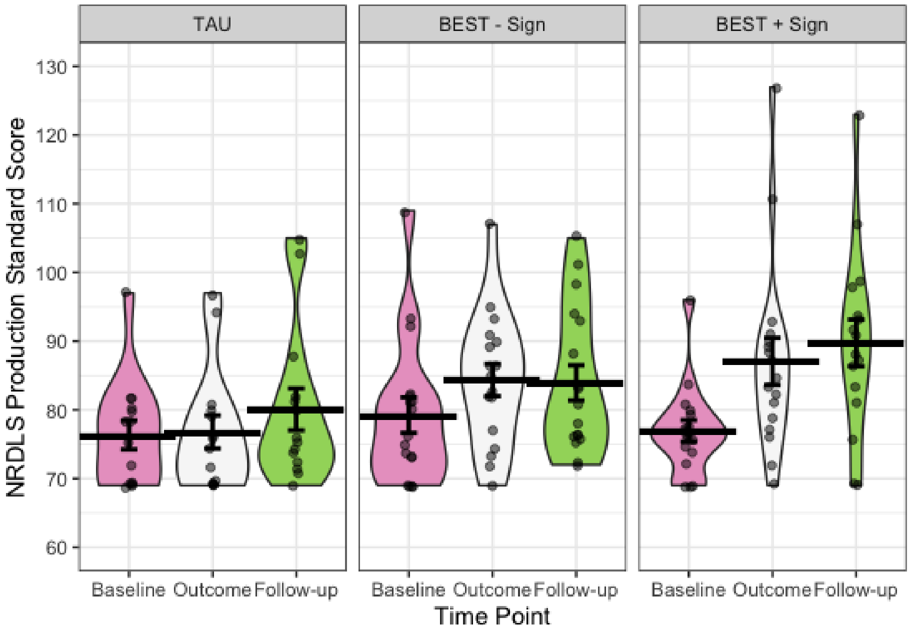


Note: At baseline n = 48; no missing data. At outcome n = 48; no missing data. At follow-up n = 47; 1 missing.

Figure 3.2 suggests that there is an overall small positive trend across the time points for production scores, whereby the mean scores increase over time. A larger increase is seen between baseline and outcome, and a minor increase is seen between outcome and follow-up. A flat trajectory was expected for standard scores, however the change in kernel densities suggest that some children have accelerated beyond their expected trajectory, whilst others remain seemingly unchanged.

Figure 3. 3

Production standard scores (NRDLS) by intervention arm and time point



Note: At baseline n = 48; no missing data. At outcome n = 48; no missing data. At follow-up n = 47; 1 missing.

Figure 3.3 shows the baseline, outcome, and follow-up scores (indicated by colour) across the three intervention arms (separated by facets). (N = 48; no missing data)

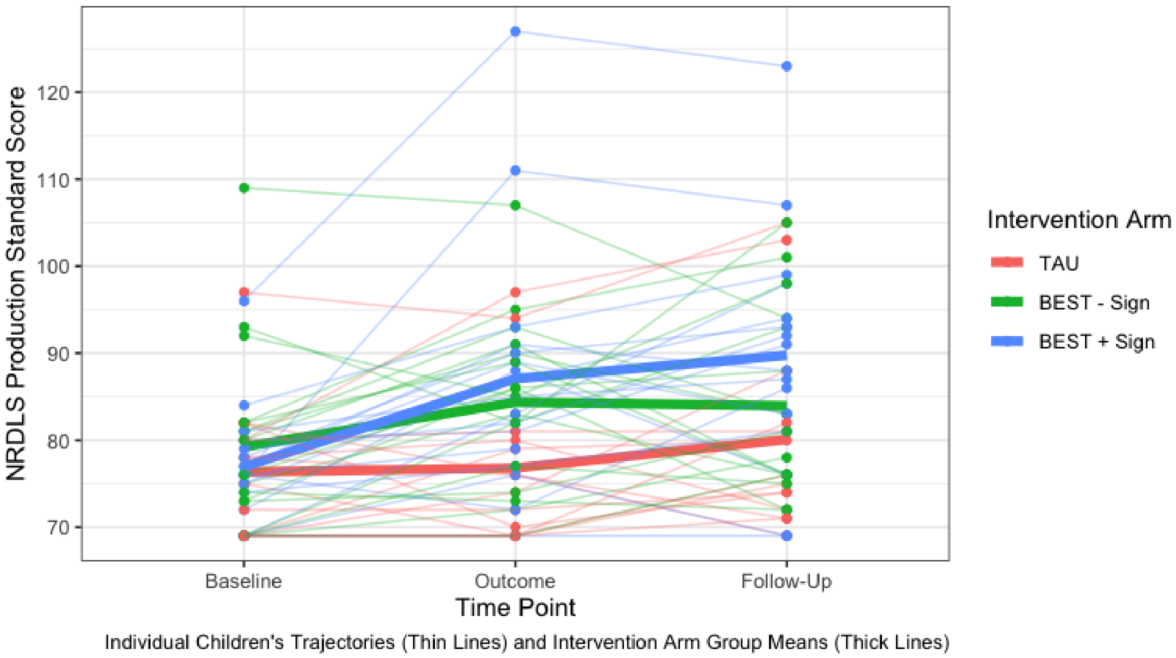
The baseline plots for the three treatment arms are relatively similar, although the mean score is marginally lower for BEST with sign. At outcome means increase for BEST without sign and BEST with sign, but not for TAU. At follow-up means further increase for BEST with sign, and children in TAU also show mean increases. For BEST without sign the mean scores remain the same at follow-up.

There is a positive trend over time for each intervention arm, with the most pronounced increases for BEST with sign, followed by BEST without sign. For both BEST with sign and BEST without sign the confidence intervals do not overlap between baseline and outcome, suggesting overall progress is made. For TAU the confidence intervals overlap for all three timepoints.

There is clear variability in the data with some high scores in the BEST arms contributing to the higher overall means. A flat trajectory was expected for standard scores, however the change in kernel densities suggest that some children have accelerated beyond their expected trajectory, whilst others remain seemingly unchanged. The sequential trend supports the hypothesis that there is a positive effect of BEST on standard production scores which is maintained after 6 weeks, which is most pronounced for BEST with sign.

Figure 3. 4

Children's individual production standard scores by intervention arm across timepoint



Note: At baseline n = 48; no missing data. At outcome n = 48; no missing data. At follow-up n = 47; 1 missing.

Figure 3.4 shows children's individual trajectories over the three time points. Red data represents each child receiving Treatment as Usual, green represents BEST without additional sign, and blue represents the BEST with additional sign condition. The denser lines represent the mean values for each treatment arm. There is clearly high variability within and between each treatment arm, although there are greater mean score increases for the BEST arms overall at outcome and follow-up.

3.4.5 Production (NRDLS Production Scale Standard Scores) Results

3.4.5.1 Parameters of Interest

There was a non-significant effect of BEST without sign at outcome (β = 5.32, 95% CI [-1.12, 11.77], *p* = .11). There was also a non-significant effect of BEST without Sign at follow-up (β = 2.12, 95% CI [-6.31, 10.56], *p* = .62).

These results suggest that an average child receiving BEST without Sign would gain around 5 points more than a TAU child between baseline and outcome. A child receiving BEST without Sign would gain around 2 more points than a TAU child between baseline and follow-up; whilst there may have been some effect of BEST without Sign at outcome, this appears reduced by follow-up and has a non-significant effect overall.

There was a significant effect of BEST with Sign at outcome (β = 10.14, 95% CI [3.68, 16.60], *p* < .01). The effect of BEST with Sign at follow-up was also significant (β = 9.44, 95% CI [0.93, 17.94], *p* < .05).

3.4.5.2 Covariables

There was a non-significant effect of age as expected due to the narrow range in the sample (β = -0.21, 95% CI [-0.45, 0.03], *p* = .11). The effect of dosage was non-significant (β = 0.15, 95% CI [-0.56, 0.86], *p* = .67), as discussed further in the Dosage Analysis section. The effect of IDACI score was also non-significant (β = -5.936, 95% CI [-14.36, 2.48], *p* = .20) suggesting that a unit increase of school deprivation score may have a negative effect on scores (in reality the change would be 0.1, not 1, so the effect is negligible).

The effect of time was non-significant at outcome and follow-up. At outcome a child receiving TAU gained around 0.5 more points than at baseline (β = 0.45, 95% CI [-4.29, 5.19], *p* = .85). By follow-up a child receiving TAU gained around 3.75 more points than at baseline (β = 3.76, 95% CI [-2.41, 9.92], *p* = .24).

The effectiveness of the matching process outlined in chapter 2 was confirmed as BEST without sign was not different to TAU at baseline in their NRDLS production language scores (β = -1.72, 95% CI [-11.71, 8.26], *p* = .74), and neither was BEST with Sign (β = -0.50, 95% CI [-10.73, 9.73], *p* = .93). There was a non-significant effect of expressive language profile at baseline (β = 0.49, 95% CI [-2.54, 3.53], *p* = .75). There was a significant effect of receptive language profile at baseline (β = 21.81, 95% CI [17.09, 26.53], *p* < .001). This is logical as by definition a child with a receptive language profile cannot be below the 16^{th} centile on the expressive subscale, so they would be expected to score significantly higher than children with expressive or mixed language profiles at baseline.

3.4.5.3 Effect Sizes

3.4.5.3.1 BEST without sign

At outcome the effect size for BEST without Sign was large (*d* = 1.08). At follow-up there was a negligible effect size (*d* = .18).

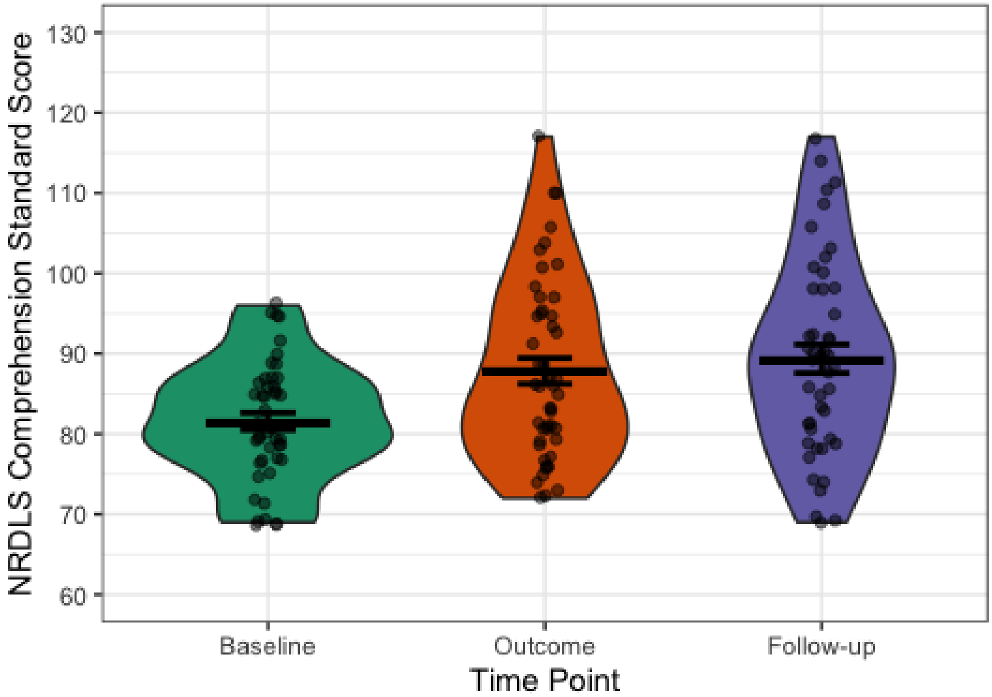
3.4.5.3.2 BEST with sign

At outcome and follow-up large effect sizes were detected (*d* = 1.11) and (*d* = 1.45) respectively.

3.4.6 Comprehension (NRDLS Comprehension Scale Standard Scores)

Figure 3. 5

NRDLS Comprehension Standard Scores over three time points (baseline, outcome and follow-up), collapsed across intervention arms

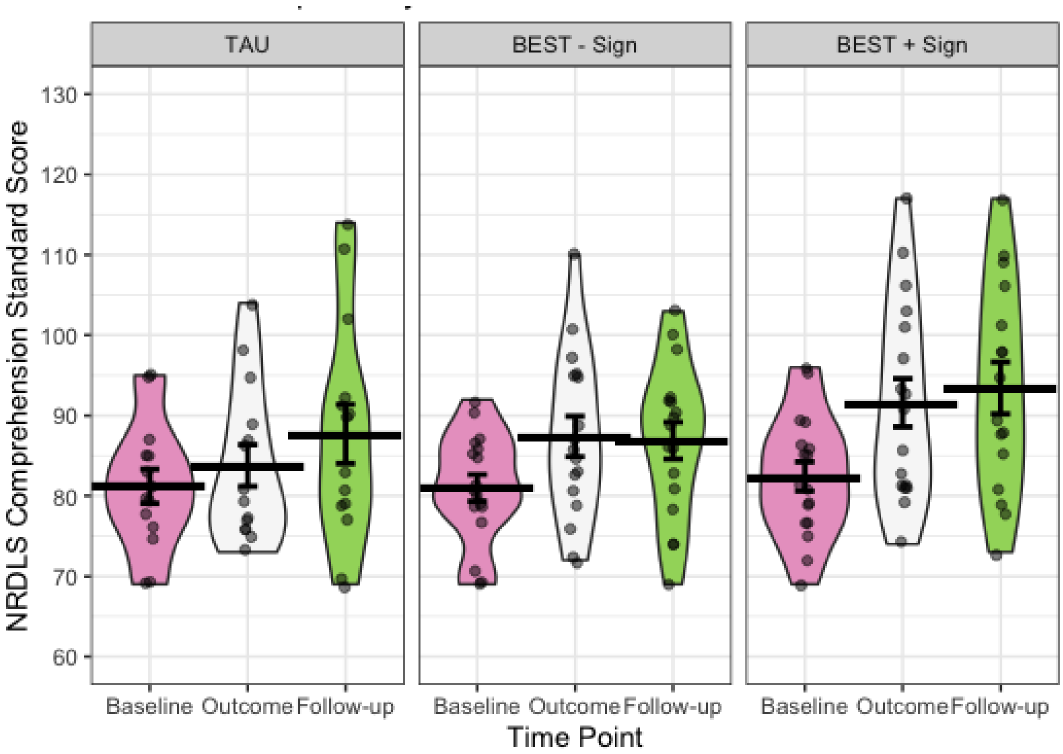


Note: At baseline n = 48; no missing data. At outcome n = 48; no missing data. At follow-up n = 47; 1 missing.

Figure 3.5 shows the scores for all children at each of the three timepoints. The average scores demonstrate an upwards trend over time, suggesting improvement since standard scores are age adjusted and therefore the expected trajectory is flat. The kernel densities support the suggestion that at least some of the improvement is due to some children accelerating, since the plots elongate upwards at the outcome and follow-up timepoints. Some scores remain at the bottom of the violin plot, however. The trend overall is similar to the one seen in figure 1.4 for production scores, with a bigger improvement between baseline and outcome, and a very marginal increase from outcome to follow-up.

Figure 3. 6

Comprehension Standard Scores across time points and intervention arms



Note: At baseline n = 48; no missing data. At outcome n = 48; no missing data. At follow-up n = 47; 1 missing.

Figure 3.6 shows the baseline, outcome, and follow-up scores (indicated by each colour) across the three intervention arms (separated by facets).

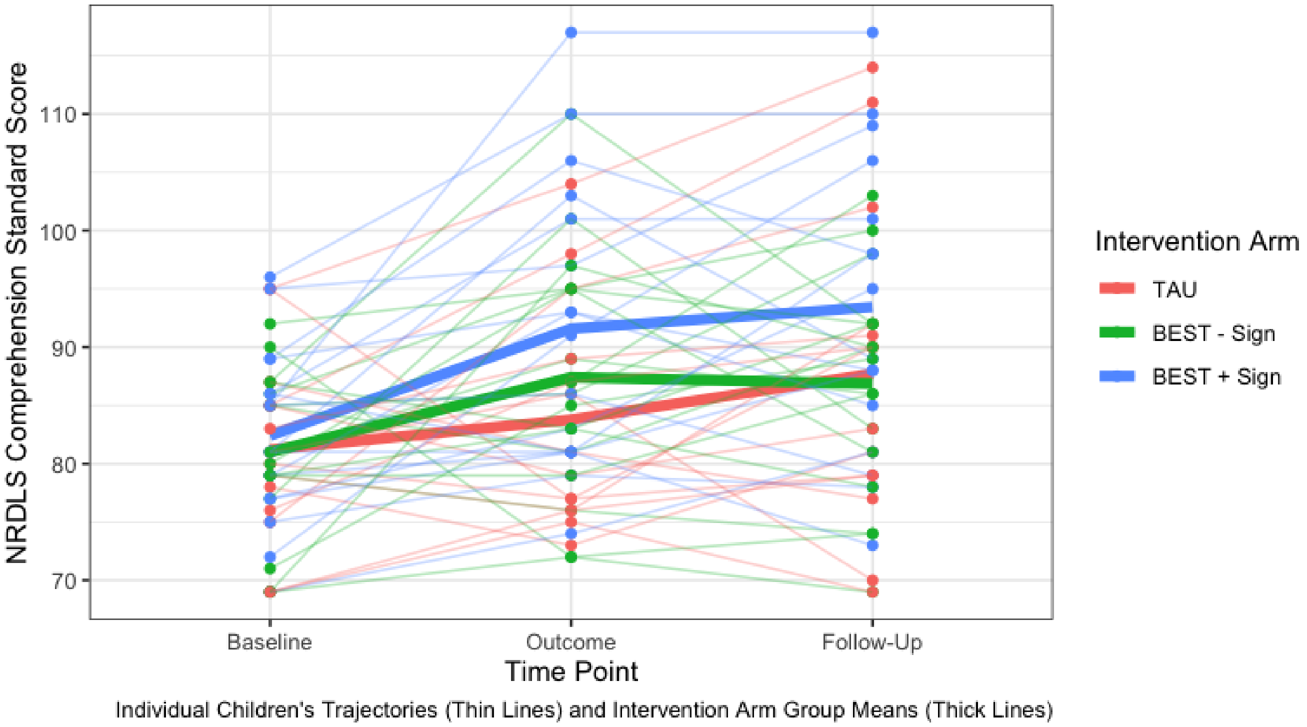
The baseline plots for the three treatment arms appear very similar. At outcome means increase for all three intervention arms, with most increase seen for BEST with sign, followed by BEST without sign and a small increase for TAU. At follow-up means further increase for BEST with sign, and children TAU. For BEST without sign the mean scores remain the same at follow-up.

There is a positive trend over time for each intervention arm, with the most pronounced increases for BEST with sign where increases are seen even post intervention. For BEST without sign there is increase between baseline and outcome which is maintained at follow-up. TAU shows a somewhat surprising increase from baseline to outcome to follow-up, considering that flat trajectories would be expected for children on standardised scores. Confidence intervals overlap for TAU, whilst they do not appear to overlap between baseline and outcome for the BEST arms.

There is variability in the data, with a wider range of scores for the BEST arms. A flat trajectory was expected for standard scores; however, all treatment arms demonstrate increased mean scores.

Figure 3. 7

Children's individual and mean comprehension standard scores by intervention arm across each time point



Note: At baseline n = 48; no missing data. At outcome n = 48; no missing data. At follow-up n = 47; 1 missing.

Figure 3.7 shows individual children's and intervention arm group mean trajectories over the three time points. Red data represents each child receiving Treatment as Usual, green represents BEST without additional sign, and blue represents the BEST with additional sign condition.

3.4.7 Comprehension (NRDLS Comprehension Scale Standard Scores) Results

3.4.7.1 Parameters of Interest

All parameters of interest were non-significant for comprehension scores.

The effect of BEST without Sign at outcome was non-significant (β = 2.75, 95% CI [-7.37, 12.86], *p* = .60).

There was a non-significant effect of BEST without Sign at follow-up (β = -1.77, 95% CI [-12.40, 8.86], *p* = .75).

There was a non-significant effect of BEST with Sign at outcome (β= 6.67, 95% CI [-3.58, 16.92], *p* = .21). The effect of BEST with Sign at follow-up was also non-significant (β = 4.12, 95% CI [-6.66, 14.91], *p* = .46).

3.4.7.2 Covariables

There was a non-significant effect of age (β = -0.13, 95% CI [-0.22, 0.48], *p* = .47). The effect of dosage was non-significant (β = 0.19, 95% CI [-0.50, 0.87], *p* = .59). There was a significant effect of IDACI score (β = -19.28, 95% CI [-29.87, -8.68], *p* < .01) suggesting that a unit increase of school deprivation score negatively impacted scores.

The effect of time was non-significant at outcome compared to baseline (β = 4.88, 95% CI [-2.50, 12.26], *p* = .20). The effect of follow-up was significant with an improvement of around 8.5 points compared to baseline (β = 8.63, 95% CI [0.91, 16.36], *p* = .03).

The BEST without sign group was not statistically significantly different to TAU at baseline (β = -2.92, 95% CI [-13.32, 7.49], *p* = .60), and neither was BEST with Sign (β = -1.47, 95% CI [-11.92, 8.98], *p* = .79). There was a significant effect of expressive language profile at baseline (β = 10.40, 95% CI [7.85, 12.95], *p* < .001). There was a non-significant effect of receptive language profile at baseline (β = -1.11, 95% CI [-5.42, 3.21], *p* = .62). Since by definition a child with an expressive language profile cannot be below the 16th centile for comprehension it is expected that they would score significantly higher than children with receptive or mixed language profiles at baseline.

3.4.7.3 Effect Sizes

3.4.7.3.1 BEST without sign

At outcome the effect size for BEST without Sign was small (*d* = .35). At follow-up there was negligible negative effect (*d* = -0.12).

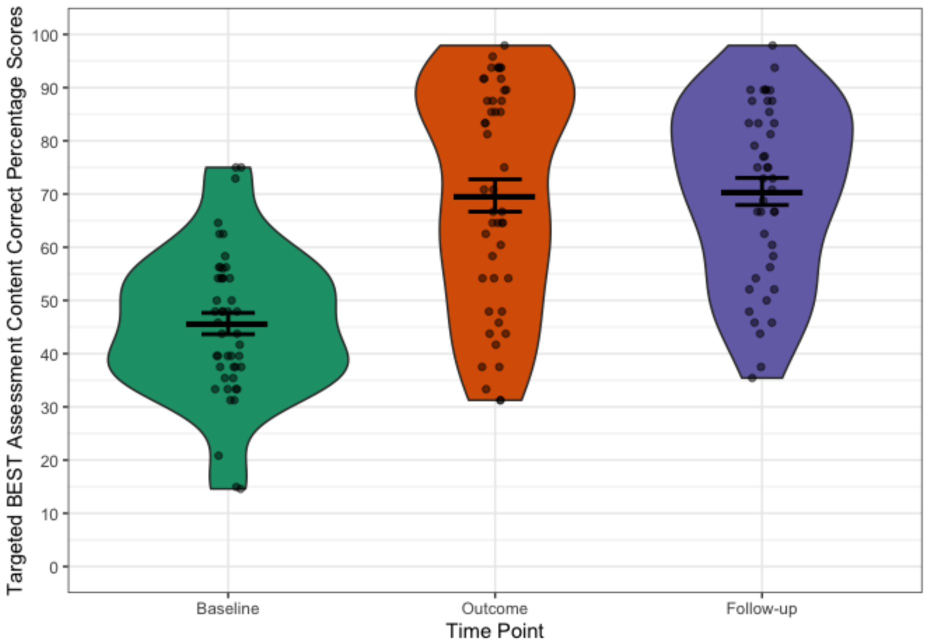
3.4.7.3.2 BEST with sign

At outcome and follow-up medium effect sizes were detected (*d* = .50) and (*d* = .50) respectively. There was wide variability in outcomes since results for comprehension outcomes were non-significant.

3.4.8 Targeted BEST Sentence Content (Targeted BEST Assessment Content Correct)

Figure 3. 8

Targeted BEST Assessment Content Correct Percentage Scores over three time points (baseline, outcome and follow-up), collapsed across intervention arms

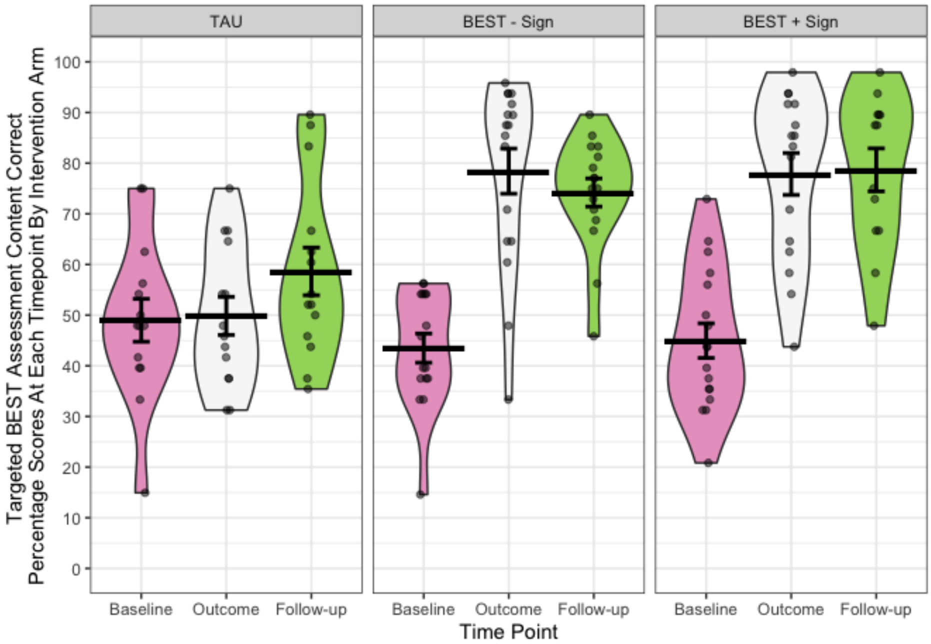


Note: At baseline n = 47; 1 missing. At outcome n = 47; 1 missing. At follow-up n = 43; 5 missing.

Figure 3.8 displays the progress made by all participants over the course of the study. The mean of the baseline scores sit a little below chance, with a large spread between scores of 15 and 75 percent. The mean score increased to 70 percent at outcome, with little change at follow-up. There is still a large spread at outcome and follow-up, from around 30 percent to just below ceiling.

Figure 3. 9

Targeted BEST Assessment Content Correct Percentage Scores across time points and intervention arms



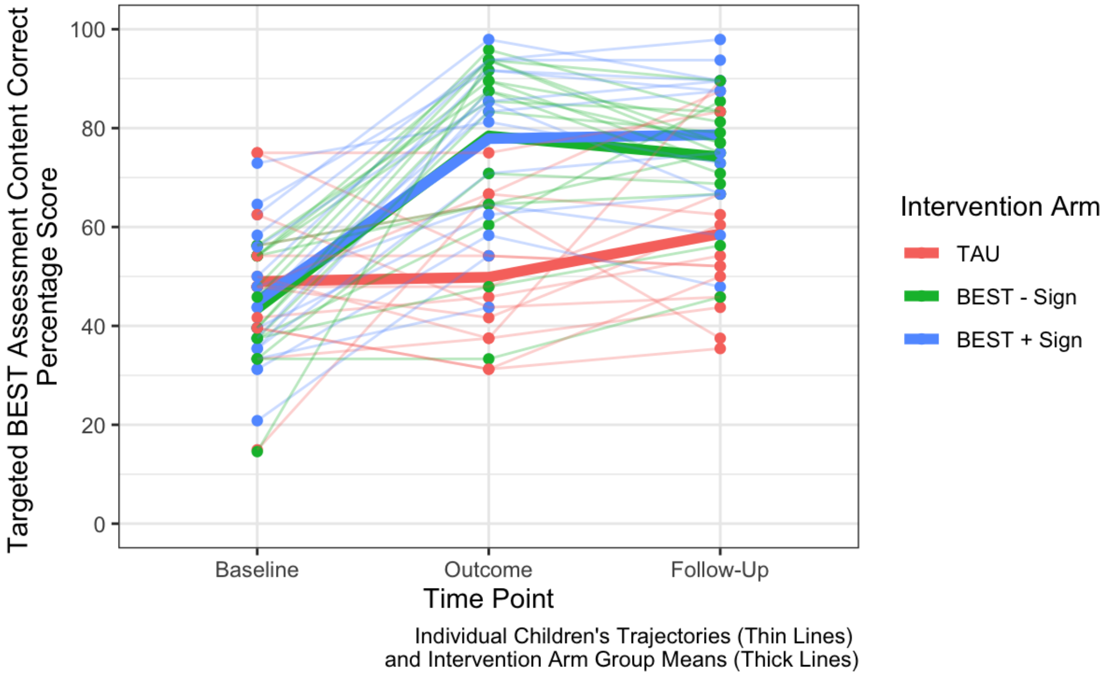
Note: At baseline n = 47; 1 missing. At outcome n = 47; 1 missing. At follow-up n = 43; 5 missing.

Figure 3.9 shows the baseline, outcome, and follow-up scores (indicated by each colour) across the three intervention arms (separated by facets). At baseline the mean scores are similar, although slightly higher for TAU. There is a fair amount of variability in the kernel densities suggesting differing spread of the data. At outcome there are large increases for BEST without sign and BEST with sign. At follow-up the largest increase is seen for TAU. The BEST with sign remains constant and the mean for BEST without sign drops by around 5 points. Due to high outcome scores for children in the BEST arms ceiling effects may affect the mean scores at follow-up, and children may be making similar amounts of progress across the three arms. Confidence intervals clearly do not overlap for baseline and outcome for the BEST arms, whilst they do for TAU. The scores support the hypothesis that BEST supports children with targeted content, but a difference between the BEST arms is not visible.

It must be noted that there is one missing data point at baseline and outcome, and five missing data points at follow-up due to missing assessment data.

Figure 3. 10

Children's individual and mean sentence content correct percentage scores by intervention arm across each time point



Note: At baseline n = 47; 1 missing. At outcome n = 47; 1 missing. At follow-up n = 43; 5 missing.

Figure 3.10 shows individual children's and intervention arm group mean trajectories over the three time points. Red data represents each child receiving Treatment as Usual, green represents BEST without additional sign, and blue represents the BEST with additional sign condition.

3.4.9 Targeted BEST Sentence Content (Targeted BEST Assessment Content Correct) Results

3.4.9.1 Parameters of Interest

All parameters of interest were significant for sentence content scores.

At outcome the effect of BEST without Sign was significant (β = 34.72, 95% CI [22.12, 47.32], p < .001).

There was also a significant effect of BEST without Sign at follow-up (β = 23.27, 95% CI [10.19, 36.35], p < .001).

There was a significant effect of BEST with Sign at outcome (β= 32.09, 95% CI [19.41, 44.77], p < .001). The effect of BEST with Sign at follow-up was also significant (β = 21.95, 95% CI [8.58, 35.33], p < .01).

3.4.9.2 Covariables

There was a significant effect of age (β = 0.79, 95% CI [0.20, 1.37], p < .05). The effect of dosage was non-significant (β = 0.29, 95% CI [-1.24, 1.82], p = .71). There was a significant effect of IDACI score (β = -34.04, 95% CI [-52.46, -15.61], p < .01) suggesting that a unit increase of school deprivation score negatively impacted scores.

The effect of time was non-significant at outcome compared to baseline (β = 0.50, 95% CI [-8.79, 9.78], p = .92). The effect of follow-up was also non-significant when compared to baseline (β = 8.85, 95% CI [-0.74, 18.43], p = .07).

The BEST without sign group was not statistically significantly different to TAU at baseline (β = -13.05, 95% CI [-34.80, 8.70], p = .27), and neither was BEST with Sign (β = -8.95, 95% CI [-31.10, 13.20], p = .45). The effect of expressive language profile at baseline was non-significant (β = 1.66, 95% CI [-4.81, 8.12], p < .62). The effect of receptive language profile at baseline was also non-significant (β = 0.52, 95% CI [-9.71, 10.74], p = .92).

3.4.9.3 Effect Sizes

All outcomes were standardised (mean = 0, SD = 1) using the scale() and centre() functions (R Core Team, 2019), and compared to the intercept group (TAU) at on sentence content scores at outcome. Cohen's d effect sizes of BEST on sentence content were interpreted in line with Cohen (1988).

3.4.9.3.1 BEST without sign

At outcome the effect size for BEST without Sign was large (*d* = 3.60). At follow-up there was also a large effect (*d* = 1.25).

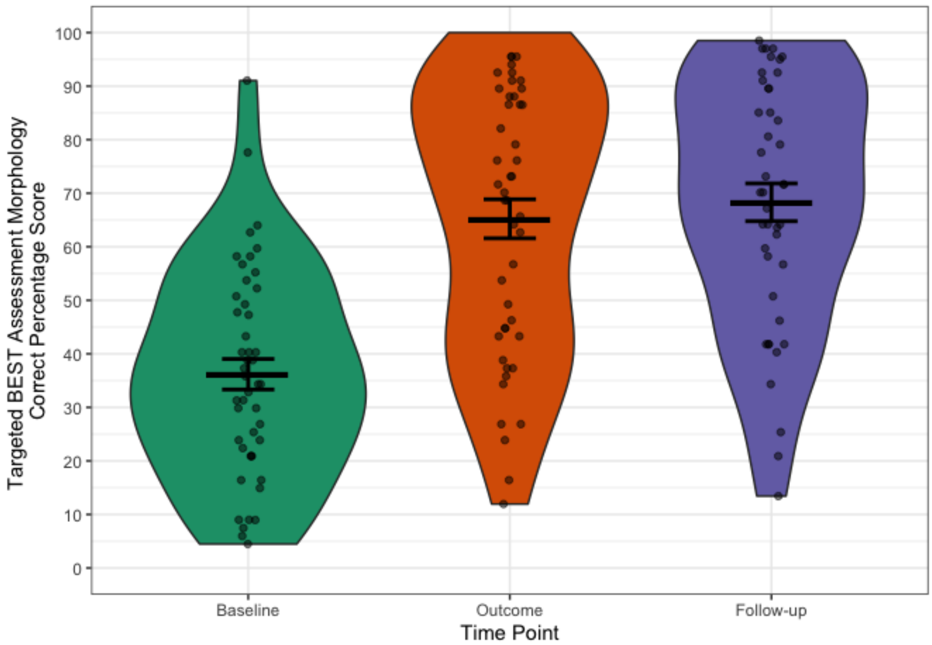
3.4.9.3.2 BEST with sign

Large effect sizes were detected at outcome and follow-up for BEST with sign: (*d* = 1.78) and (*d* = 2.14) respectively.

3.4.10 Targeted BEST Sentence Morphology (Targeted BEST Assessment Morphology Correct)

Figure 3. 11

Targeted BEST Assessment Morphology Correct Percentage Scores over three time points (baseline, outcome and follow-up), collapsed across intervention arms.

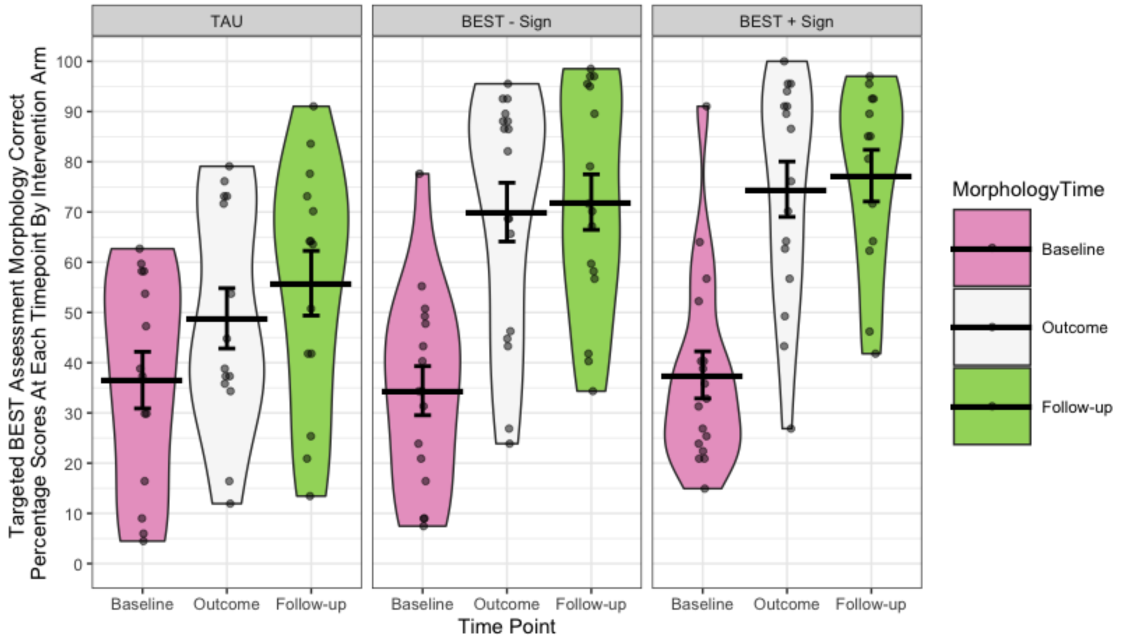


Note: At baseline n = 47; 1 missing. At outcome n = 47; 1 missing. At follow-up n = 43; 5 missing.

Figure 3.11 shows the pooled results for all treatment arms over the three time periods for morphology correct percentage scores. Overall the mean scores increased from around 37% to 65% at outcome and 68% at follow-up. Observation of the previous graphs indicates that the increases can be mainly attributed to the BEST conditions, although there was also some progress seen for the TAU group. Some progress is likely to be due to natural progression over the 14-week study period.

Figure 3. 12

Targeted BEST Assessment Morphology Correct Percentage Scores across time points and intervention arms



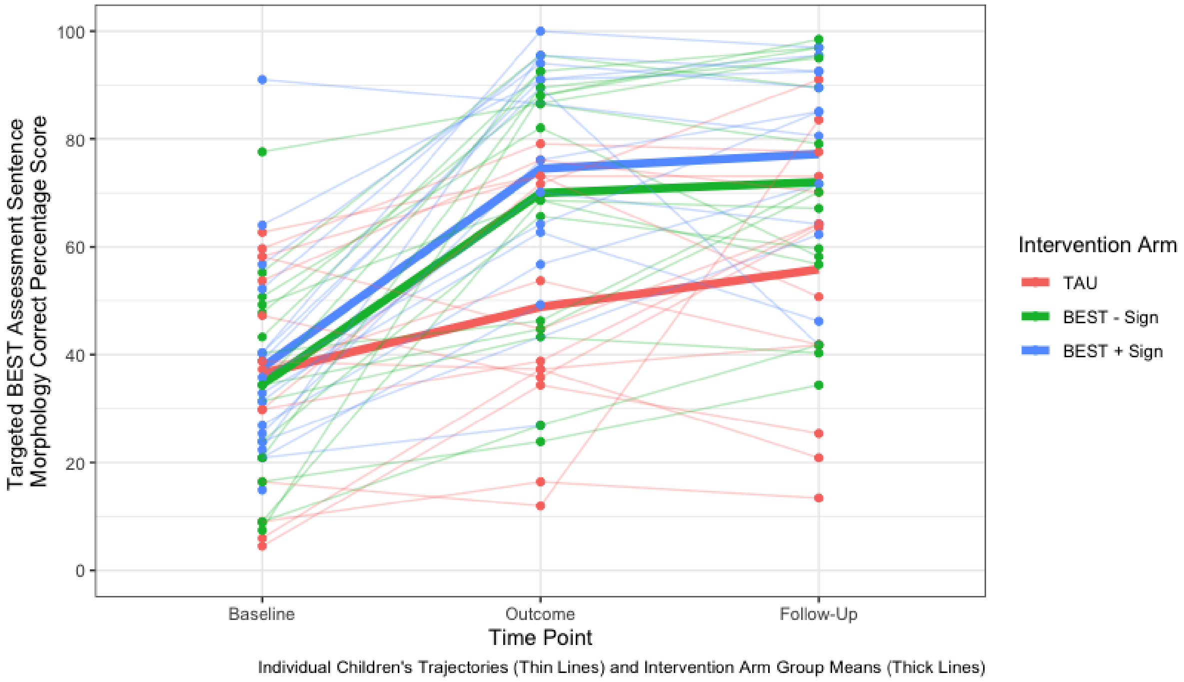
Note: At baseline n = 47; 1 missing. At outcome n = 47; 1 missing. At follow-up n = 43; 5 missing.

Figure 3.12 shows the baseline, outcome, and follow-up scores (indicated by colour) across the three intervention arms (across the three facets). At baseline scores are similar with higher upper scores for the BEST arms. At outcome the BEST arm means increase considerably and TAU also sees an increase. Means increase very slightly in the BEST arms and a little more in TAU at follow-up, however ceiling effects could affect mean follow-up outcome scores as many children in the BEST arms are already near ceiling at outcome. The plots support the hypothesis that BEST supports performance on the targeted morphology compared to TAU. There is a slightly more pronounced increase for BEST with sign than BEST without sign.

The mean scores, confidence intervals and kernel densities are similar across the BEST arms. There is one missing data point at baseline and outcome, and five missing data points at follow-up due to missing assessment data.

Figure 3. 13

Children's individual and mean sentence morphology correct percentage scores by intervention arm across each time point



Note: At baseline n = 47; 1 missing. At outcome n = 47; 1 missing. At follow-up n = 43; 5 missing.

Figure 3.13 shows individual children's and intervention arm group mean trajectories over the three time points. Red data represents each child receiving Treatment as Usual, green represents BEST without additional sign, and blue represents the BEST with additional sign condition.

3.4.11 Targeted BEST Sentence Morphology (Targeted BEST Assessment Morphology Correct) Results

3.4.11.1 Parameters of Interest

At outcome the effect of BEST without Sign was significant (β = 22.42, 95% CI [7.67, 37.17], *p* < .01).

There was also a significant effect of BEST without Sign at follow-up (β = 18.85, 95% CI [2.04, 36.66], *p* < .05).

There was a significant effect of BEST with Sign at outcome (β= 23.82, 95% CI [9.02, 38.61], *p* < .01). The effect of BEST with Sign at follow-up was non-significant at the p < .05 level (β = 16.57, 95% CI [-0.69, 33.84], *p* = .06).

3.4.11.2 Covariables

The effect of age was non-significant (β = 0.28, 95% CI [-0.85, 1.41], *p* < .63). The effect of dosage was non-significant (β = -0.18, 95% CI [-2.47, 2.12], *p* = .88). There was a significant effect of IDACI score (β = -51.26, 95% CI [-79.62, -22.89], *p* < .01).

The effect of time was significant at outcome compared to baseline (β = 12.29, 95% CI [1.46, 23.13], *p* = .03). The effect of follow-up was also significant when compared to baseline (β = 19.05, 95% CI [6.86, 31.24], *p* < .01).

The BEST without sign group was not statistically significantly different to TAU at baseline (β = -4.00, 95% CI [-35.85, 27.86], *p* = .81), and neither was BEST with Sign (β = 6.60, 95% CI [-25.74, 38.94], *p* = .70). The effect of expressive language profile at baseline was non-significant (β = -10.25, 95% CI [-20.39, -0.11], *p* = .06). The effect of receptive language profile at baseline was also non-significant (β = 9.94, 95% CI [-5.92, 25.80], *p* = .23).

3.4.11.3 Effect Sizes

All outcomes were standardised (mean = 0, SD = 1) using the scale() and centre() functions (R Core Team, 2019), and compared to the intercept group (TAU) at on sentence morphology scores at outcome. Cohen's d effect sizes of BEST on sentence morphology were interpreted in line with Cohen (1988).

3.4.11.3.1 BEST without sign

At outcome the effect size for BEST without Sign was large (*d* = 1.99). At follow-up there was also a large effect (*d* = 0.79).

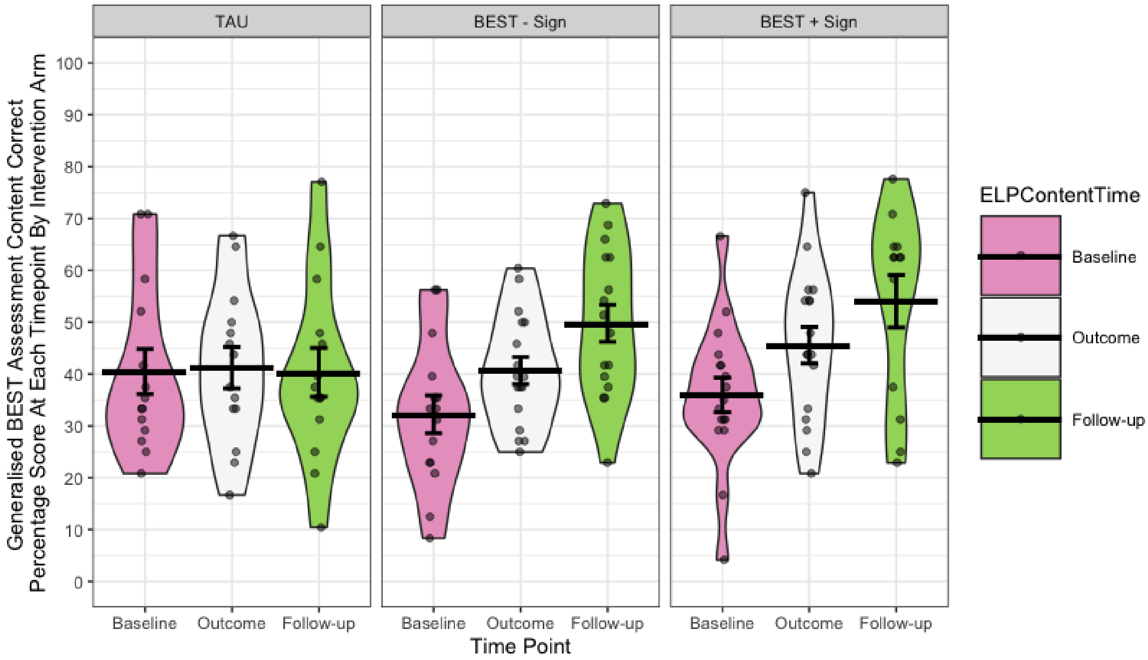
3.4.11.3.2 BEST with sign

Large effect sizes were detected at outcome and follow-up for BEST with sign: (d = 1.13) and (d = 1.25) respectively.

3.4.12 Generalised BEST Sentence Content (Generalised BEST Assessment Content Correct)

Figure 3. 14

Generalised BEST Assessment Content Correct Percentage Scores across time points and intervention arms

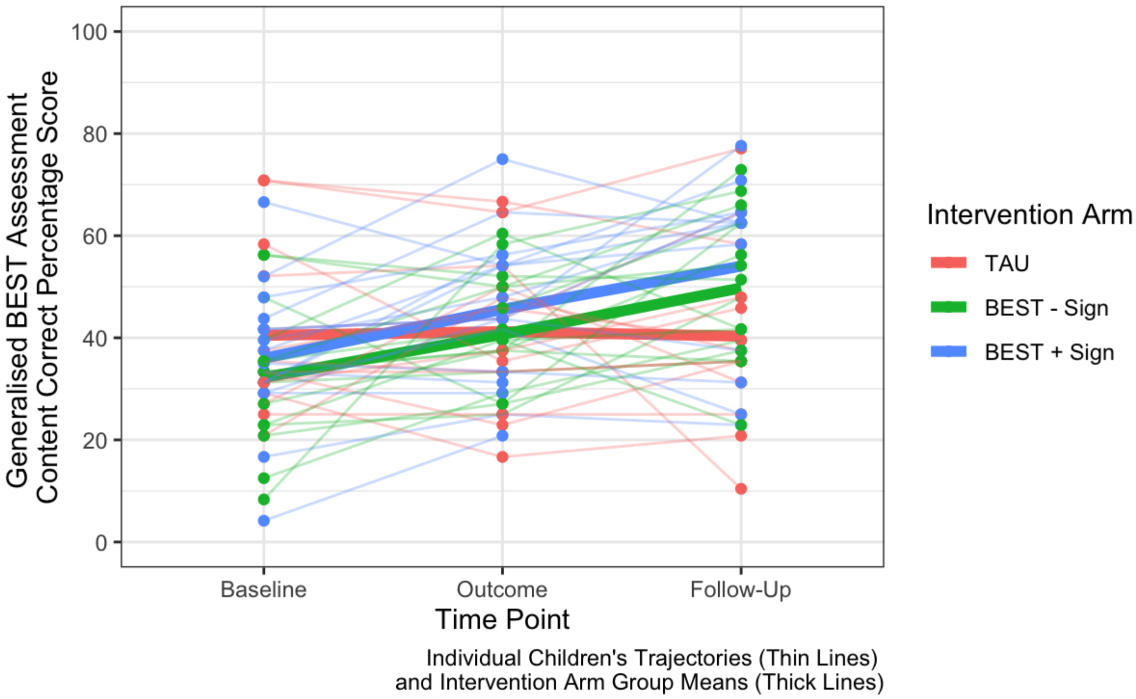


Note: At baseline n = 46; 2 missing. At outcome n = 48; no missing data. At follow-up n = 43; 5 missing.

Figure 3.14 shows the baseline, outcome, and follow-up scores for generalised content measured using the Generalised BEST Assessment (indicated by colour) across the three intervention arms (across the three facets). At baseline there is a small amount of variability across the scores, with the mean TAU score higher than the BEST arms. At outcome progress was made for the BEST arms, with the highest mean increase for BEST with sign. No progress was made for the TAU group. At follow-up progress was once again seen for the BEST arms with greater mean increase in the BEST with sign group. Again, there was no increase for TAU. Overall this graph supports the hypothesis that children receiving BEST show greater mean generalisation scores increases for content, and this is slightly more pronounced for the BEST with sign group.

Figure 3. 15

Individual and mean Generalised BEST Assessment content correct percentage scores by intervention arm and time point



Note: At baseline n = 46; 2 missing. At outcome n = 48; no missing data. At follow-up n = 43; 5 missing.

Figure 3.15 shows individual children's and intervention arm group mean trajectories over the three time points. Red data represents each child receiving Treatment as Usual, green represents BEST without additional sign, and blue represents the BEST with additional sign condition.

3.4.13 Generalised Sentence Content Results

3.4.13.1 Parameters of Interest

At outcome the effect of BEST without Sign was non-significant (β = 6.08, 95% CI [-3.72, 15.8;7], *p* = .23).

There was a significant effect of BEST without Sign at follow-up (β = 17.54, 95% CI [5.26, 29.81], *p* < .01).

At outcome the effect of BEST with Sign was non-significant (β= 7.59, 95% CI [-1.98, 17.15], $p < .12$).

There was a significant effect of BEST with Sign at follow-up (β = 15.52, 95% CI [2.93, 28.11], $p < .05$).

3.4.13.2 Covariables

There was a non-significant effect of age (β = 0.78, 95% CI [0.02 1.53], $p = .05$). The effect of dosage was non-significant (β = 0.61, 95% CI [-1.02, 2.24], $p = .47$). There was a non-significant effect of IDACI score at the < .05 level (β = -22.25, 95% CI [-41.91, -2.58], $p = .05$).

The effect of time was non-significant at outcome compared to baseline (β = 1.42, 95% CI [-5.71, 8.55], $p = .70$). The effect of time at follow-up was also non-significant when compared to baseline (β = -0.14, 95% CI [-9.01 8.73], $p = .98$).

The BEST without sign group was not statistically significantly different to TAU at baseline (β = -17.31, 95% CI [-40.91, 6.29], $p = .18$), and neither was BEST with Sign (β = -10.58, 95% CI [-34.36, 13.20], $p = .41$). The effect of expressive language profile at baseline was non-significant (β = -1.29, 95% CI [-8.35, 5.77], $p < .72$). The effect of receptive language profile at baseline was also non-significant (β = -0.76, 95% CI [-12.04, 10.51], $p = .90$).

3.4.13.3 Effect Sizes

All outcomes were standardised (mean = 0, SD = 1) using the scale() and centre() functions (R Core Team, 2019), and compared to the intercept group (TAU) on Generalised BEST content scores at outcome. Cohen's d effect sizes of BEST on the Generalised BEST sentence content scores were interpreted in line with Cohen (1988).

3.4.13.3.1 BEST without sign

At outcome the effect size for BEST without Sign was large (*d* = .81). At follow-up there was also a large effect (*d* = 1.01).

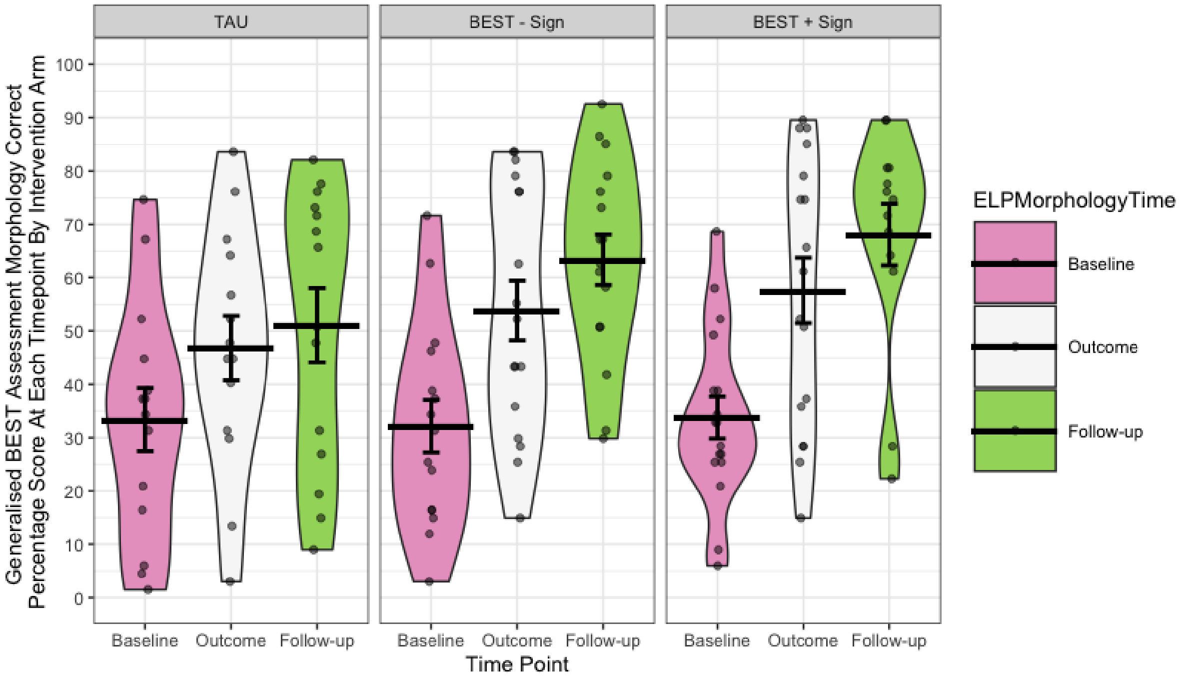
3.4.13.3.2 BEST with sign

A medium effect size was detected at outcome (*d* = .56), and a large effect size was detected at follow-up (*d* = 1.61).

3.4.14 Generalised BEST Sentence Morphology (Generalised BEST Assessment Morphology Correct)

Figure 3. 16

Generalised BEST Assessment Morphology Correct Percentage Scores across time points and intervention arms

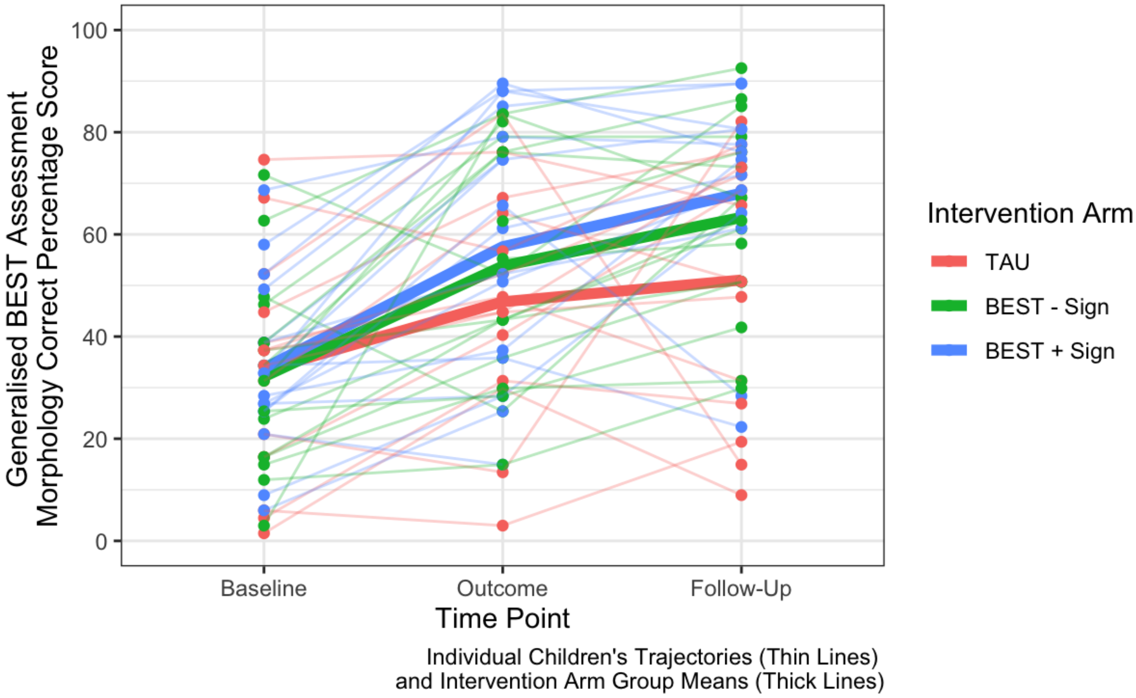


Note: At baseline n = 46; 2 missing. At outcome n = 48; no missing data. At follow-up n = 43; 5 missing.

Figure 3.16 shows the baseline, outcome, and follow-up scores for generalised morphology (indicated by colour) measured using the Generalised BEST Assessment across the three intervention arms (shown across the three facets). At baseline scores and ranges are very similar. At outcome mean increases were seen for all arms, and were slightly greater for BEST, with the highest mean increase for BEST with sign. At follow-up a small increase was seen for the TAU group, and slightly greater increases for BEST without sign and BEST with sign. The TAU increase is perhaps surprising given the lack of change for the Generalised BEST content scores; however, the Generalised BEST results are in line with the Targeted BEST Assessment figures (see Figures 3.14 and 3.15). The trend is in line with the other figures presented in this chapter and suggests increased progress made by children receiving BEST, which is slightly more pronounced for the BEST with sign group. It must be noted that these analyses are with a smaller n than the main study and are therefore susceptible to power issues.

Figure 3. 17

Individual and mean Generalised BEST Assessment morphology correct percentage scores by intervention arm and timepoint



Note: At baseline n = 46; 2 missing. At outcome n = 48; no missing data. At follow-up n = 43; 5 missing.

Figure 3.16 shows individual children's and intervention arm group mean trajectories over the three time points. Red data represents each child receiving Treatment as Usual, green represents BEST without additional sign, and blue represents the BEST with additional sign condition. Most improvement is seen on average for BEST with sign, followed by BEST without sign. Least improvement is seen for Treatment as Usual.

3.4.15 Generalised Sentence Morphology Results

3.4.15.1 Parameters of Interest

All parameters of interest were non-significant

At outcome the effect of BEST without Sign was non-significant (β = 8.51, 95% CI [-6.10, 23.13], *p* = .26).

There was also a non-significant effect of BEST without Sign at follow-up (β = 15.93, 95% CI [-0.94, 32.79], *p* = .07).

The effect of BEST with Sign at outcome was non-significant (β= 11.21, 95% CI [-3.18, 25.59], *p* = .13). The effect of BEST with Sign at follow-up was non-significant (β = 13.40, 95% CI [-3.82, 30.62], *p* = .13).

3.4.15.2 Covariables

The effect of age was non-significant (β = 0.41, 95% CI [-0.70, 1.52], *p* = .48). The effect of dosage was non-significant (β = 0.45, 95% CI [-1.84, 2.74], *p* = .70). There was a significant effect of IDACI score (β = -48.83, 95% CI [-76.51, -21.15], *p* < .01).

The effect of time was significant at outcome compared to baseline (β = 12.82, 95% CI [2.12, 23.51], *p* < .05). The effect of follow-up was also significant when compared to baseline (β = 16.04, 95% CI [3.68, 28.40], *p* < .05).

The BEST without sign group was not statistically significantly different to TAU at baseline (β = -11.82, 95% CI [-44.11, 20.47], *p* = .49), and neither was BEST with Sign (β = -2.78, 95% CI [-35.32, 29.76], *p* = .87). The effect of expressive language profile at baseline was non-significant (β = -10.34, 95% CI [-20.59, -0.09], *p* = .06). The effect of receptive language profile at baseline was also non-significant (β = 5.54, 95% CI [-10.30, 21.39], *p* = .50).

3.4.15.3 Effect Sizes

All outcomes were standardised (mean = 0, SD = 1) using the scale() and centre() functions (R Core Team, 2019), and compared to the intercept group (TAU) at on Generalised BEST Assessment sentence morphology scores at outcome. Cohen's d effect sizes of BEST on the Generalised BEST Assessment morphology scores were interpreted in line with Cohen (1988).

3.4.15.3.1 BEST without sign

At outcome the effect size for BEST without Sign was medium (d = .79). At follow-up there was also a medium effect (d = .66).

3.4.15.3.2 BEST with sign

A medium effect size was detected at outcome (d = .55), and a large effect size was detected at follow-up (d = 1.02).

3.4.15.4 Sensitivity Analysis

By virtue of the eligibility criteria, children presented with either expressive, receptive or mixed (expressive and receptive) language profiles, indicating their primary difficulties. The children who had only one type of difficulty necessarily scored above the 16^{th} percentile on the other measure. To test whether this influenced the results presented above, a sensitivity analysis was conducted, removing children who scored highly at baseline for the relevant outcome.

Parameters of interest are reported in the respective sections below if they are different from the original results.

These results are tentative since removing participants results in reduced power, however it is valuable to ascertain whether the effects are due to children who were unimpaired on specific domains in the first place.

A full table of the multi-level model results, excluding children performing above the 16th centile at baseline, can be found in Appendix 22.

3.4.15.4.1 Production

Children scoring above the 16th percentile on the NRDLS Production subscale at baseline (n = 5) were removed from the sample and the sensitivity analysis production multilevel model was carried out for the remaining 43 participants. The results were mainly unchanged when high scoring children were removed, indicating that the findings outlined above are not due to children who did not have particular difficulties in production from the outset. This increases robustness of the analysis.

Of the parameters of interest, BEST without sign at outcome became significant where it was not in the original analysis (β = 6.73, 95% CI [0.71, 12.75], p = .03). There were no significant changes in any other parameters of interest.

3.4.15.4.2 Comprehension

As above, children scoring on or above the 16th percentile for comprehension at baseline were removed from the data (n = 14) and the sensitivity analysis comprehension multilevel model was conducted. The outcomes were unchanged by the removal of high-scoring children in this sensitivity analysis.

There were no significant changes in any of the parameters of interest.

3.4.15.4.3 Sentence Content

Sentence content is an expressive language measure and was therefore treated in the same way as production above. Children scoring above the 16th percentile on the NRDLS Production subscale at baseline (n = 5) were removed from the sample and the sensitivity analysis multilevel model for sentence content was carried out.

There were no significant changes in any of the parameters of interest.

3.4.15.4.4 Sentence Morphology

Sentence morphology is a measure of expressive language. Children scoring above the 16th percentile on the NRDLS Production subscale at baseline (n = 5) were removed from the sample and the sensitivity analysis multilevel model for sentence content was carried out.

Of the parameters of interest, BEST without sign at follow-up became non-significant, whereas it was significant in the original analysis (β = 6.73, 95% CI [-4.31, 28.97], p = .15). There were no significant changes in any other parameters of interest.

The results are robust to the sensitivity analysis and do not appear to be driven by children who score above the 16th centile on the measure in question.

3.4.15.5 Dosage Analysis

The BEST intervention consists of 16 therapy sessions. Despite close adherence to the protocol, some children missed sessions due to absence or scheduling difficulties. The average dosage received (out of a possible 16 BEST sessions) did not differ statistically across the two BEST arms. A dosage analysis was carried out to establish whether the number of sessions received affected language outcomes.

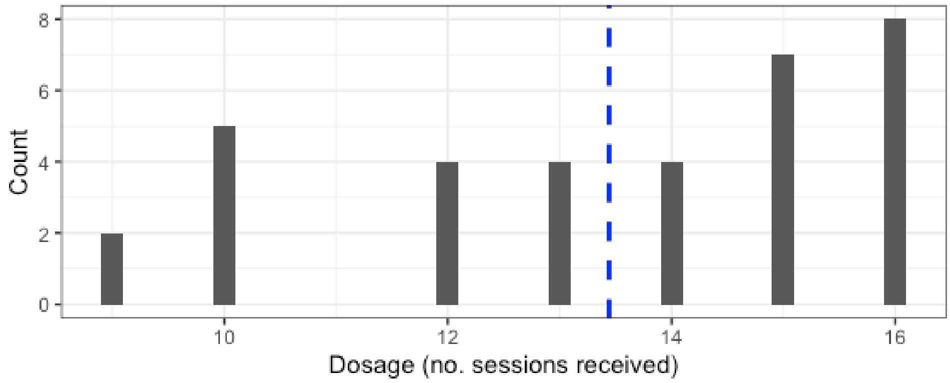
3.4.15.5.1 High vs Low Dosage Analysis

The BEST manual does not stipulate a minimum 'adequate dosage', so the following procedure was followed: children in the BEST arms who received over 75% of sessions (a minimum of 13 out of 16) were considered to have received the 'adequate dosage', based on a mean dosage of 13.44 sessions (see figure 3.18). The number of children who received hypothesised high and low dosage were calculated (see table 3.6).

An analysis was carried out to establish whether there was an effect of receiving the adequate dosage (high dosage) compared to receiving 12 sessions or fewer (low dosage). Change scores on comprehension and production scores were calculated for children in the high and low dosage groups.

Figure 3. 18

Dosage received by children across study



Note: Average number of BEST sessions (Dosage) received by children in the BEST intervention arms. The blue dashed line represents the mean dosage (13.44).

Table 3. 6

Distribution of hypothesised sufficient (>75%) and insufficient (<=75%) dosage received by children in the BEST with sign and BEST without sign arms

	BEST with sign		BEST without sign	
	High Dosage (>75%)	Low Dosage (<=75%)	High Dosage (>75%)	Low Dosage (<=75%)
n	12	5	11	6
Total n	17		17	

Between groups ANOVAs were conducted. No significant differences were identified between the high and low dosage group change scores between baseline and outcome for production [F (1,32) = 0.85, p = .36], comprehension [F (1,32) = 1.449, p = .24), sentence content [F (1,30) = 0.079, p = .78] and sentence morphology [F (1,30) = 0.108, p = .75].

Between baseline and Follow-Up there was a significant difference between low and high dosage children's change scores for production [F (1,31) = 4.244, p = <.05]. Unexpectedly the low dosage group showed larger change scores than the high dosage group (p = <.05). Differences between change scores between low and high dosage groups were non-significant for comprehension [F (1,31) = 0.014, p = .91), sentence content [F (1,26) = 0.003, p = .96) or sentence morphology [F (1,26) = 0.108, p = .75].

3.4.15.5.2 Correlation between Dosage and Growth

Pearson product moment correlations were applied out to establish whether there was a gradient relationship between number of sessions received (dosage) and change scores on each of the outcome measures.

Figure 3. 19

Relationship between dosage and production change scores for BEST with sign and BEST without sign

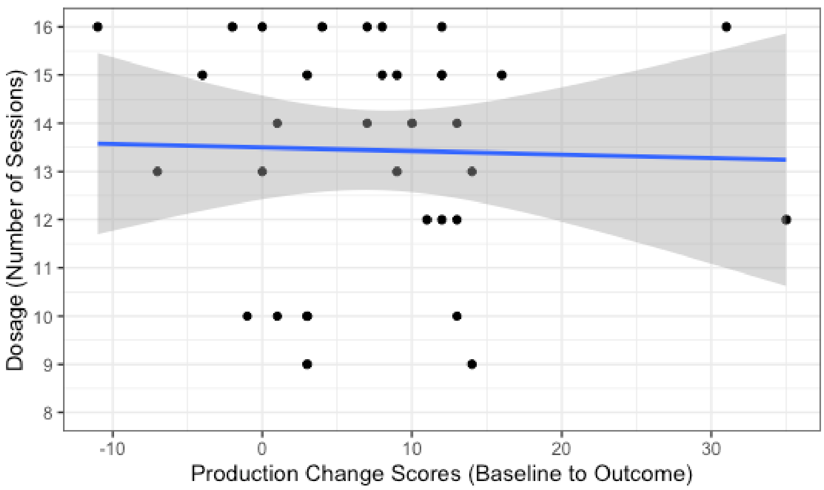


Figure 3.19 displays the correlation between dosage and NRDLS change scores for production. No significant correlation was found between dosage and change scores [r (32) = -0.03, p = 0.87]

Correlations were also non-significant between dosage received and comprehension change scores [r (32) = 0.15, P = .40], dosage and content correct percentage change scores [r (30) = -0.14, p = .44] and dosage and morphology correct percentage change scores [r (30) = -0.08, p = 0.68].

Graphs displaying the Pearson product momentum correlations for comprehension, content correct and morphology correct can be found in Appendix 23.

This analysis suggests that in the present sample there is no statistically significant advantage of receiving the hypothesised sufficient dosage over the insufficient dosage.

3.4.15.6 Functional Communication

This study was designed to include functional communication as an outcome measure using the FOCUS questionnaire. Due to the poor response rate for the FOCUS questionnaires however, analysis of functional communication across timepoints was not possible (see Table 3.1 above for questionnaire response rates). Alternative methods of gathering this data to improve response rates in future research are discussed in chapters 4 and 5.

3.5 Summary of Main Results

The key findings of the main analysis and confirmatory 2sensitivity analysis are summarised below and are then discussed in chapter 5.

Research question 1A asks whether an intervention underpinned by usage-based theory (BEST, with and without additional sign) is more effective than treatment as usual (TAU) for improving language and associated functional communication outcomes in pre-school children with language difficulties.

The results presented in this chapter indicate that BEST is more effective than TAU in improving children's language outcomes. The areas that specifically benefited from BEST were production, targeted content, targeted morphology and generalised content. These results are summarised in figures 3.20 and 3.21 below (see summary of Research Question 1C).

Research question 1B: 'Is signing a necessary 'active ingredient' of BEST, and does it differentially affect progress on different language goals (sentence structure and morphology)?' can be answered by comparing the results of BEST with and without sign on each outcome measure.

There was specific evidence for the interaction of BEST with sign for increasing production scores.

There was no effect of sign on outcome scores for comprehension.

For Targeted BEST Assessment sentence content there was no specific effect of sign over and above BEST without sign $t(34) = 1.29$, $p = 0.21$. For Targeted BEST Assessment sentence morphology there was a specific effect of BEST with sign over and above BEST without sign $t(34) = 2.37$, $p = 0.02$.

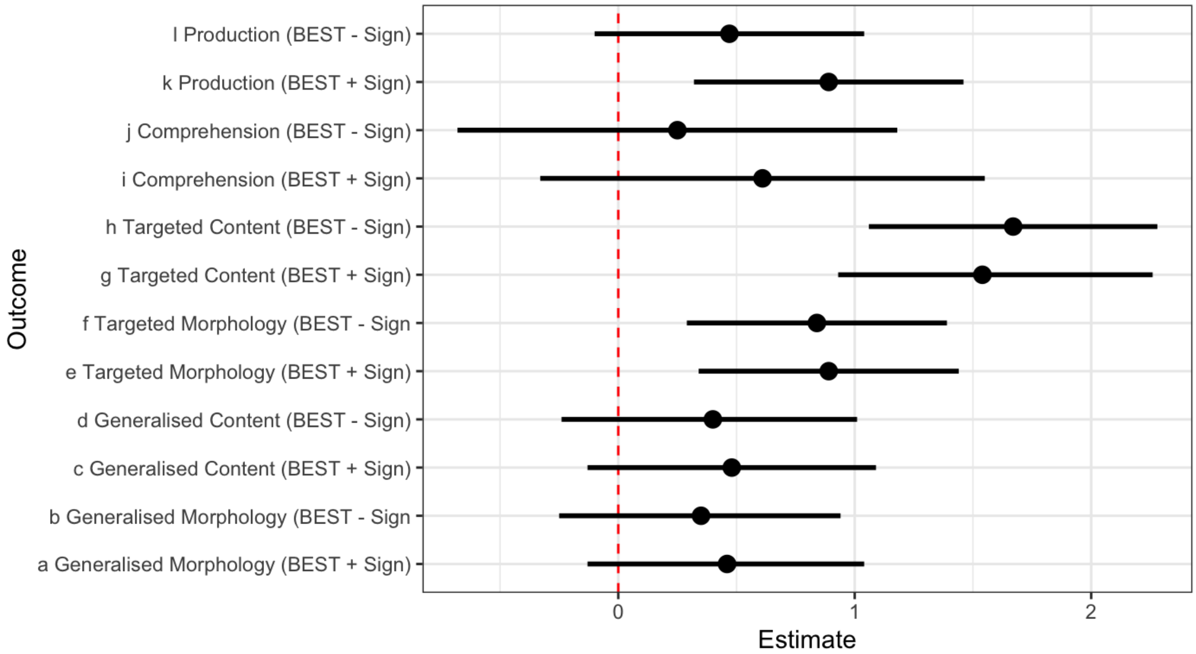
For the Generalised BEST Assessments of content and morphology there was no significant evidence for BEST with sign over and above that of BEST without sign.

Although BEST with sign was not statistically significantly better than BEST without sign for any outcomes, the visualisations presented in this chapter suggests that sign as an adjunct to BEST appears to support generalised production development beyond the BEST target structures (measured on the NRDLS) and also supports the development of targeted production and morphology. These findings suggest that role of sign is important for production outcomes, but does not outweigh the benefits of BEST for content and morphology outcomes. Results are discussed further in chapter five.

To effectively answer question 1C: ‘What are the estimated effect sizes across outcomes for BEST with and without Sign, relative to Treatment as Usual?’, Figures were created to illustrate all measures at outcome (figure 3.20) and follow-up (figure 3.21). Outcomes were standardised to be comparable between all outcome measures. The black dots represent the betas for each outcome and the black lines depict the spread of the confidence intervals. The red dashed line represents the reference TAU with no scores on covariables. Any betas whose confidence intervals cross the red line are non-significant, whilst those that do not cross the red line are significant findings at the 95% confidence level.

Figure 3. 20

Standardised effects by outcome measure and intervention arm at outcome (T1)



Note. Scores are standardised and compared with Treatment as Usual at outcome. Red dashed line depicts zero change.

[a] *Generalised morphology correct percentage scores*

[b] *Generalised morphology correct percentage scores*

[c] *Generalised content correct percentage scores*

[d] *Generalised content correct percentage scores*

[e] *Targeted morphology correct percentage scores*

[f] *Targeted morphology correct percentage scores*

[g] *Targeted content correct percentage scores*

[h] *Targeted content correct percentage scores*

[i] *Comprehension standard scores*

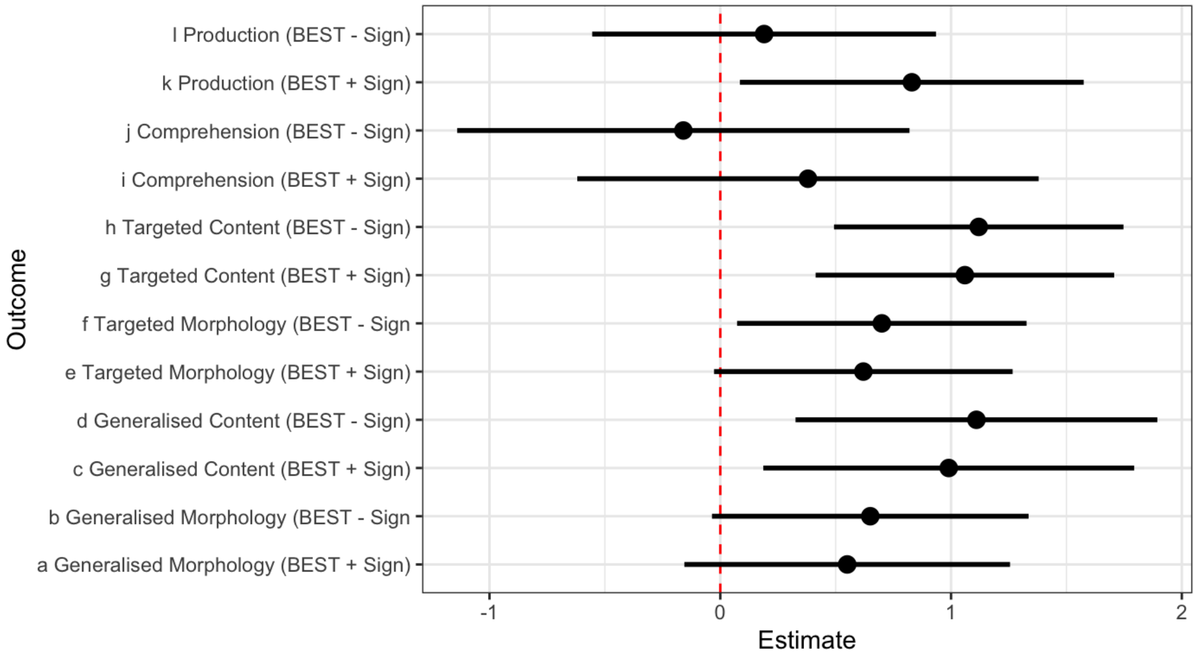
[j] *Comprehension standard scores*

[k] *Production standard scores*

[l] *Production standard scores*

Figure 3. 21

Standardised effects by outcome measure and intervention arm at follow-up (T2)



Note. Scores are standardised and compared with Treatment as Usual at follow-up. Red dashed line depicts zero change.

[a] *Generalised morphology correct percentage scores*

[b] *Generalised morphology correct percentage scores*

[c] *Generalised content correct percentage scores*

[d] *Generalised content correct percentage scores*

[e] *Targeted morphology correct percentage scores*

[f] *Targeted morphology correct percentage scores*

[g] *Targeted content correct percentage scores*

[h] *Targeted content correct percentage scores*

[i] *Comprehension standard scores*

[j] *Comprehension standard scores*

[k] *Production standard scores*

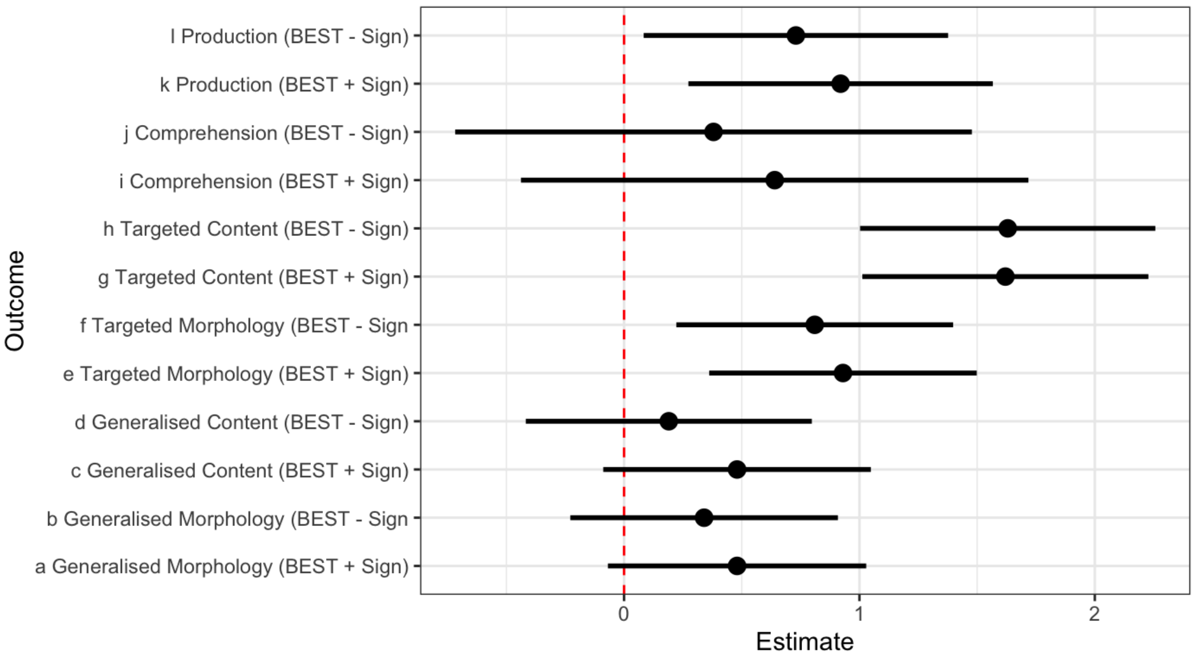
[l] *Production standard scores*

At outcome (figure 3.20), we see significant progress on production outcomes for BEST with sign only. No significant progress is observed for comprehension outcomes. Significant progress can be observed for targeted content and morphology for BEST with and without sign. No significant progress is made on generalised outcomes for either intervention arm compared with TAU.

At follow-up (figure 3.21), similar patterns are observed to those at outcome whereby significant progress is maintained on production scores for BEST with sign, but not for BEST without sign. Significant progress is not made at follow-up for any comprehension measures. Targeted content progress is maintained, whereas significant targeted morphology progress is not maintained for BEST with sign. Unlike at outcome, progress reaches significance for generalised content for both BEST with and without sign, whereas the same pattern of non-significant progress is maintained for all generalised morphology outcomes.

Figure 3. 22

Sensitivity analysis: Standardised effects by outcome measure and intervention arm at outcome (T1)



Note. Sensitivity Analysis scores are standardised and compared with Treatment as Usual at outcome. Red dashed line depicts zero change.

[a] *Generalised morphology correct percentage scores*

[b] *Generalised morphology correct percentage scores*

[c] *Generalised content correct percentage scores*

[d] *Generalised content correct percentage scores*

[e] *Targeted morphology correct percentage scores*

[f] *Targeted morphology correct percentage scores*

[g] *Targeted content correct percentage scores*

[h] *Targeted content correct percentage scores*

[i] *Comprehension standard scores*

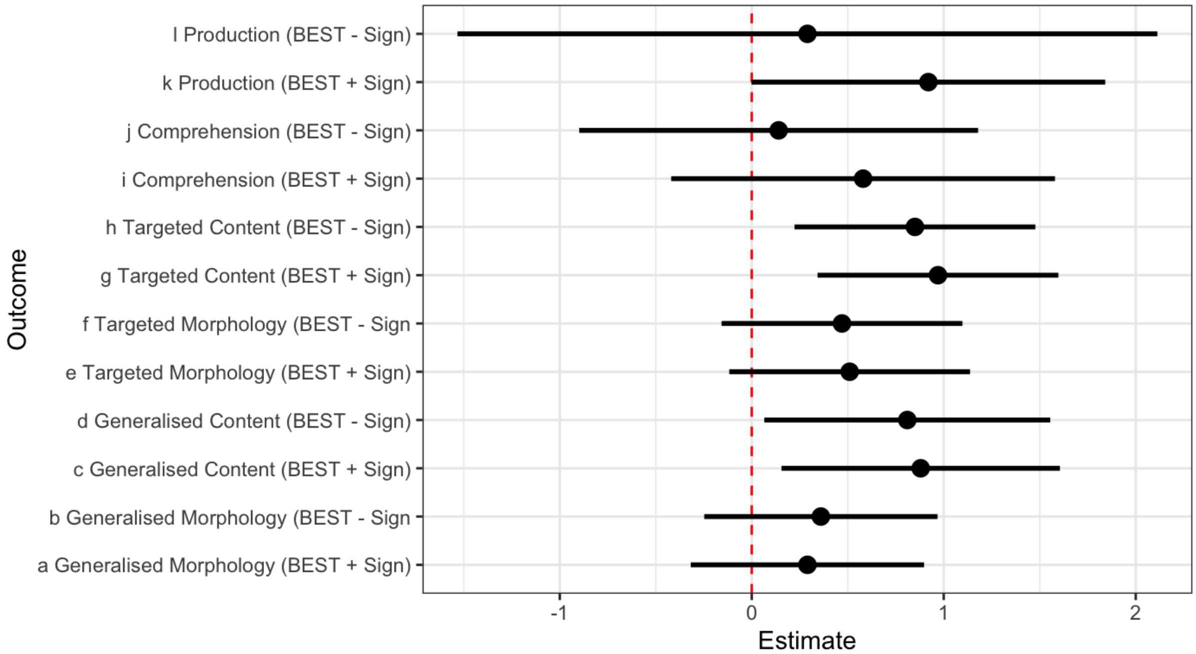
[j] *Comprehension standard scores*

[k] *Production standard scores*

[l] *Production standard scores*

Figure 3. 23

Sensitivity analysis: Standardised effects by outcome measure and intervention arm at follow-up (T2)



Note. Sensitivity Analysis scores are standardised and compared with Treatment as Usual at follow-up. Red dashed line depicts zero change.

[a] *Generalised morphology correct percentage scores*

[b] *Generalised morphology correct percentage scores*

[c] *Generalised content correct percentage scores*

[d] *Generalised content correct percentage scores*

[e] *Targeted morphology correct percentage scores*

[f] *Targeted morphology correct percentage scores*

[g] *Targeted content correct percentage scores*

[h] *Targeted content correct percentage scores*

[i] *Comprehension standard scores*

[j] *Comprehension standard scores*

[k] *Production standard scores*

[l] *Production standard scores*

Results of the sensitivity analysis (figure 3.22) are similar to those of the main analysis at outcome, with the exception of production outcomes which are significant for both intervention arms, as opposed to just BEST with sign in the main analysis. All other outcomes follow the same pattern as the main analysis.

At follow-up (figure 3.23), significant progress is no longer seen for production or targeted morphology outcomes for either intervention arm.

Significant progress is maintained for targeted content for BEST with and without sign, and significant progress is also observed at follow-up for generalised content, as in the main analysis.

3.6 Conclusion

The following questions have been addressed in this chapter.

1A) Is an intervention underpinned by usage-based theory (BEST, with and without sign) more effective than treatment as usual (TAU) for improving language and associated functional communication outcomes in pre-school children with language difficulties?

1B) Is signing a necessary 'active ingredient' of BEST, and does it differentially affect progress on different language goals (sentence structure and morphology)?

1C) What are the point and interval estimates of effect sizes across outcomes?

There is evidence for efficacy of BEST for improving language outcomes at this stage of the evaluation process, with significant positive effects for all outcomes excluding comprehension and Generalised morphology.

There is some evidence across outcomes for the use of sign as part of the BEST intervention, with positive trends in all outcomes excluding sentence content, although this does not manifest as a significant difference for the majority of outcomes. This is not unexpected for a study with a sample of this size due to wide confidence intervals and the fact that BEST would be expected to have a larger effect than sign alone.

Effect sizes are presented for each outcome measure and range between no effect and very large effect. There is a high degree of variability between individual children on measures and large effect sizes are still seen for non-significant outcomes. This suggests that some children respond very well to BEST whilst others do not at all.

The findings of this study appear robust to a sensitivity analysis, whereby children without marked impairments at baseline on certain measures do not drive the level of improvement seen across the intervention arms.

There is no evidence from the present study that the 'hypothesised sufficient dosage' is necessary for positive outcomes. Growth curve modelling could seek to address this in future research to establish the optimal number of sessions.

A larger-scale evaluation of BEST is warranted due to the positive outcomes of the present study (Craig et al., 2008; Eldridge et al., 2016). Further work must also take into account the acceptability and feasibility study reported in chapter 4 and seek to maintain acceptability to parents and schools, whilst seeking to improve return rates of language and behaviour questionnaires.

Chapter 4

Informing and Developing Future Research through Stakeholder Consultation: An Acceptability and Feasibility Study

4.1 Introduction

The research described in this thesis is situated within the piloting and feasibility stages of the evaluation frameworks outlined in chapter one (Craig et al., 2008; Eldridge et al., 2016). In line with these guidelines, it is important to transparently report barriers encountered during this study, and to make recommendations to avoid costly mistakes in subsequent evaluations. This chapter outlines an acceptability and feasibility study and presents process data concerning the present study to understand teacher's attitudes towards the BEST intervention and how the research was conducted. Stakeholder consultation is employed in this part of the research to gather teacher's views about the acceptability and feasibility of the BEST intervention and of how the study was conducted. This work is used to inform the design of a subsequent study of BEST. Post-hoc analyses and behaviour change theory are then applied to draw a series of theoretically motivated recommendations for the refinement of future research. Reflective process data consisting of field notes and a reflective log recorded by the Author are then presented, and further recommendations are made based on these observations.

4.2 Conducting Research within Schools

4.2.1 Advantages of working with schools

There are numerous advantages to collaborating with schools to conduct research. Schools provide access to numerous potential participants who have established relationships with teachers and senior leadership. Logistically, children are already at the school, and participation in school-based research does not usually place additional demands on parents/carers in terms of transportation to separate sites. Space to conduct the study is also available through the school, at no extra cost to the research team (Bartlett et al., 2017). School-based research can be mutually beneficial, as children are provided with an intervention that may otherwise be unavailable to them. Teachers and senior leadership also often receive additional training or monitoring as an incentive for taking part. This can take place at every stage of the research process, from development to evaluation, and includes the opportunity to take part in Stakeholder Consultation to inform research design.

Recruitment may be aided by already established relationships between teachers and parents (Bruzzese et al., 2009). Retention rates are also likely to be higher than those of community based studies due to mandatory school attendance (Bartlett et al., 2017). Evidently school-based research is worthwhile and has many benefits, however it must be carefully managed to maintain good relationships with staff.

4.2.2 Barriers to Effective Working in Schools

There are several potential barriers to working with schools: Obtaining ethical approval and fully informed consent for school-based research is often a lengthy process, with many levels of approval necessary from the NHS, the Local Authorities, the University ethics board and the schools themselves. This may be intensified when personal, audio, or video data are collected and sufficient time should be built in to the research plan to allow for this (Bartlett et al., 2017). Despite the benefits of working alongside teachers outlined above, lack of caregiver trust in the research team can still pose a barrier to recruitment, while relationship building can take significant time and effort. Additional strategies such as multiple channels of contact and providing incentives to participants and their families may therefore be required (Berry et al., 2013). Practical barriers to conducting research in schools include scheduling, space allocation, and staff availability (Pincus & Friedman, 2004). It is important that researchers understand these barriers and demonstrate understanding to teachers when approaching potential schools. These barriers need to be avoided to promote positive research outcomes.

4.2.3 Enhancing Effective Collaboration with Schools

When working with schools, a key contact who is committed to the research project and available to help with navigating the school system should be sought early on in the research process (White, 2012). Relationships with all school staff should be cultivated and maintained, since they can be instrumental to the success of the study, for example by staff encouraging the return of consent forms and facilitating the practical aspects of the study (Pincus & Friedman, 2004). The researcher should be mindful of their role as a guest in the school, and be flexible and accommodating around curricular and extra-curricular activities (Bartlett et al., 2017), since in some instances the research may be terminated if it is perceived to interrupt everyday activities (Pincus & Friedman, 2004). Offering CPD or other useful incentives can increase effective collaboration, as this can act as a motivating reason to remain in the study (Pincus & Friedman, 2004; Wagner et al., 2004).

4.3 Stakeholder Consultation

This part of the study involved a stakeholder consultation with teachers who had been involved in the BEST research. Stakeholders are individuals or groups who have a stake in the processes and outcomes of research or policy change (Deverka et al., 2012). Whilst this work was initially carried out to inform a future larger-scale trial of BEST, strong themes of acceptability and feasibility also became evident throughout the process which are explored to enhance contextualisation of the study.

This chapter triangulates the empirical data presented in chapter 3 by introducing a qualitative analysis of the stakeholder consultations with teachers and results of an acceptability and feasibility study. School staff acted as advisors to the Author in light of their experience working alongside the researchers during the trial (National Institute for Health Research, 2012).

The discussions with teachers aimed to identify common themes, which were used to generate theoretically motivated recommendations for future research in line with the MRC framework (Craig et al., 2008) and behaviour change theory (Michie et al., 2011), and informed by a framework of acceptability (see 4.6.1-4.6.3 below). The resulting recommendations can be applied to a larger-scale evaluation of BEST, as well as to other school-based research since many of them can be generalised beyond this study. All school staff gave informed consent to take part and for their anonymised comments to be employed in the present analysis'.

4.4 Research Question

The thematic analysis was inductive in nature, and therefore not driven by a-priori research questions. Despite this, the following research question was conceptualised overarching aim of this research was to examine:

2) What theoretically motivated recommendations can be made from applying Behaviour Change Theory and the Theoretical Framework of Acceptability to stakeholder consultation outcomes, to understand and increase the acceptability and feasibility of the research and therefore inform the next stage of BEST development?

4.5 Methods

4.5.1 Recruitment Procedure

Following completion of the empirical study reported in this thesis, the Local Authority Speech and Language Therapy Lead emailed teachers on behalf of the Author inviting them to take part in additional stakeholder consultation work. This approach was taken because the lead SLT was involved in supporting the recruitment for the BEST trial and had existing relationships with school staff. School staff were informed of the purposes of the consultation, and that if they chose to take part their responses would be fully anonymised and passed on to the Author for analysis.

An independent Research Associate with extensive qualitative research experience followed up affirmative email responses with phone calls and school staff provided verbal consent to participate in the consultation over the phone. The Author was heavily involved with the development of the stakeholder consultation questions, however consultations were performed by the Research Associate to avoid biases due to the relationships developed over the course of the study.

4.5.2 Participants

Six teachers agreed to take part in individual stakeholder consultations. These teachers and their respective schools remained anonymous to the Author. The consultation was completed via telephone by the independent Research Associate.

4.5.3 Anonymity

Consultation data were anonymised and summarised by the Research Associate. Any comments containing identifiable information were removed so the researcher could not identify which teachers had made which comments. Teachers taking part were made aware of this process so that they would feel as comfortable as possible discussing the study with the Research Associate.

4.6 Qualitative Methods

An inductive analysis was first conducted using thematic analysis. A deductive framework analysis categorising and interpreting the data using prespecified themes in the Capability, Opportunity Motivation-Behaviour (COM-B) model was then undertaken.

4.6.1 Thematic Analysis

Qualitative methods were used to analyse stakeholder consultations conducted with school staff and the Author's field notes to produce evidence-based recommendations to facilitate improvements for future research. Thematic analysis is a broad and flexible analytical method used to draw out important themes from qualitative data. It follows a process of familiarisation with the data, generating descriptive codes, generating themes to group codes by, and defining and naming the themes. Thematic analysis is applied across a variety of data, including interview and field notes, and is more flexible than other approaches to qualitative analysis such as grounded theory and conversation analysis (Braun & Clarke, 2006a).

4.6.2 Behaviour Change Theory

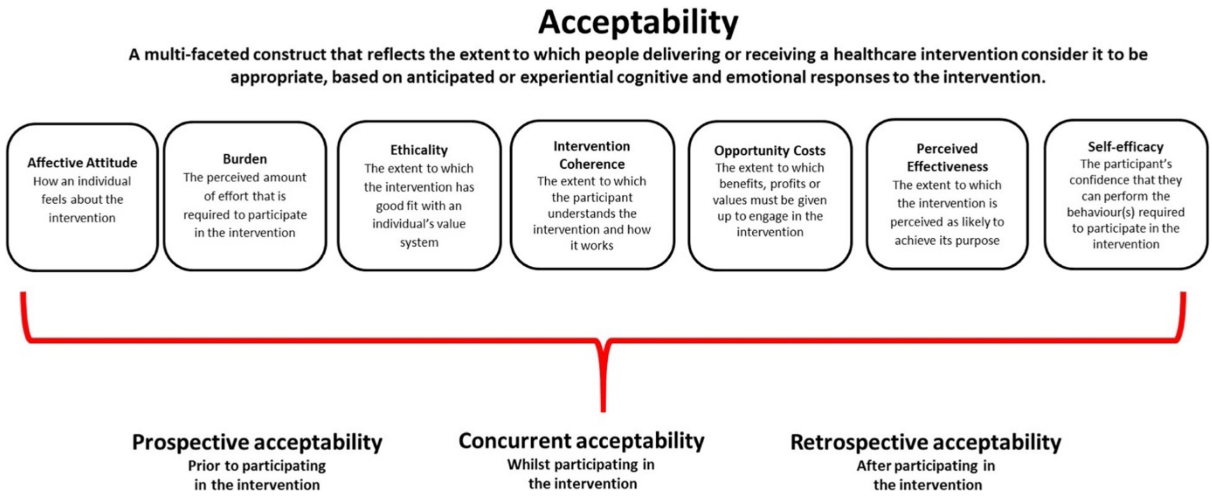
Behaviour change theory (BCT) is used to investigate barriers to certain behaviours and identify ways of promoting desired behaviours. BCT is frequently applied to public health interventions such as HIV preventative measures (Albarracín et al., 2005) and improving hand hygiene (Lhakhang et al., 2015), and improves likelihood of positive intervention outcomes. The Capability, Opportunity Motivation-Behaviour (COM-B) model was designed using behaviour change theory to characterise the facilitators of a target behaviour (physical capabilities, psychological capabilities, social opportunity, physical opportunity, automatic motivation, and reflective motivation). COM-B forms the behaviour system at the centre of the Behaviour Change Wheel (BCW) which can be used to link behaviour to appropriate intervention functions to facilitate behaviour change (Michie et al., 2011). The intervention functions that link to components of the COM-B model are education, persuasion, incentivisation, coercion, training, restriction, environmental restructuring, modelling, and enablement. Analysis of themes and their relation to behaviour change theory are used in this analysis to promote an evidence-based approach to making recommendations.

4.6.3 Acceptability

It is important to establish the acceptability of interventions to those who deliver and receive them (Craig et al., 2008). This information can be used to increase recruitment, retention and positive engagement with future studies, as well as negating any potential negative effects. Sekhon et al. (2017) define acceptability as a reflection of the degree to which recipients or those delivering an intervention find it appropriate based on cognitive and psychological responses.

Figure 4. 1

The Theoretical Framework of Acceptability (TFA) reproduced with permission from Sekhon et al., 2017 (pp.8)



Sekhon et al. introduced the Theoretical Framework of Acceptability (TFA) (see figure 4.1), consisting of the following constructs: prospective, concurrent, and retrospective acceptability: affective attitude, burden, ethicality, intervention coherence, opportunity costs, perceived effectiveness, and self-efficacy. The outcomes of stakeholder consultations discussed below are therefore also considered within the context of each TFA construct, to determine acceptability of the intervention to school staff.

4.6.4 Procedure

An independent Research Associate conducted telephone stakeholder consultations with school staff who took part in the BEST study and who agreed to take part additional acceptability and feasibility work. One school staff member communicated their responses via email due to their lack of availability for a telephone conversation. Conversations were not recorded verbatim but were instead summarised to protect the anonymity of the school staff to the Author. This method was not considered problematic for the specific aims of this research, which were to identify broad, high-level themes, unlike in many typical thematic analyses which may be more in-depth (Braun & Clarke, 2006a). The Research Associate asked open ended questions centred around the themes listed in table 4.1 below. The themes were agreed upon a-priori by the Author and Research Associate but were not scripted since the stakeholder consultation was not designed to follow a formal interview schedule. Responses were summarised and anonymised, and all responses that may have led the researcher to guess which teacher had made the comments were removed. The summarised responses were presented in a short report (see Appendix 24).

The analysis was completed in five phases. Themes were first identified via an inductive (bottom-up) thematic analysis at the semantic level (Braun & Clarke, 2006a). This phase followed guidelines set out by Braun & Clarke, whereby the data was 'read and re-read' for familiarisation, and preliminary ideas for themes were noted. Initial codes were generated for the semantic content of the data, and the themes were linked with similar themes, refined and named (see figure 4.2).

During phase two, the identified themes were linked to the Theoretical Domains Framework (TDF) (see table 4.2).

In the third phase the theoretical domains were then plotted onto the COM-B model to identify the components underpinning them (see table 4.2).

In the fourth phase the intervention functions associated with each of the behaviours from the COM-B model were identified (Cane et al., 2012; Michie et al., 2005; Michie et al., 2011).

In phase five, a series of theoretically motivated recommendations for future school-based studies based on behaviour change theory were generated. These are presented below.

Acceptability was assessed separately using the Theoretical Framework of Acceptability (TFA) (Sekhon et al., 2017), whereby the intervention was considered against the seven TFA constructs (burden, intervention coherence, opportunity costs, and perceived effectiveness) (see table 4.3).

Following the above stages, the Author used field notes taken during the study to consider whether any further barriers to study implementation could be addressed using the same method described above (see 4.10).

4.6.5 Key Areas of Discussion

The main purpose of conducting stakeholder consultation was to inform future research carried out in partnership with schools. Teachers were made aware that it would also be reported in this analysis. Telephone (n=5) and email (n= 1) conversations were carried out by a Research Associate who was not otherwise involved in the present study. Teachers' responses were anonymous and did not in any way affect the treatment that their schools received during or after the study. Open-ended questions were used to gather information broadly relating to topics outlined in Table 4.1.

Table 4. 1

Key areas of discussion in stakeholder consultation

General feeling around speech and language needs in the classroom
The research team (The Author and RAs)
Communication between the school and research team
Communication between the parents/carers and research team
Recruitment, assessment and eligibility
Forms (including consent forms, information sheets and SDQ and FOCUS questionnaires
Practical considerations and study timing
Teaching Assistants (TAs)
BEST sessions
Homework booklets
Teacher's overall impressions
Appealing CPD incentives

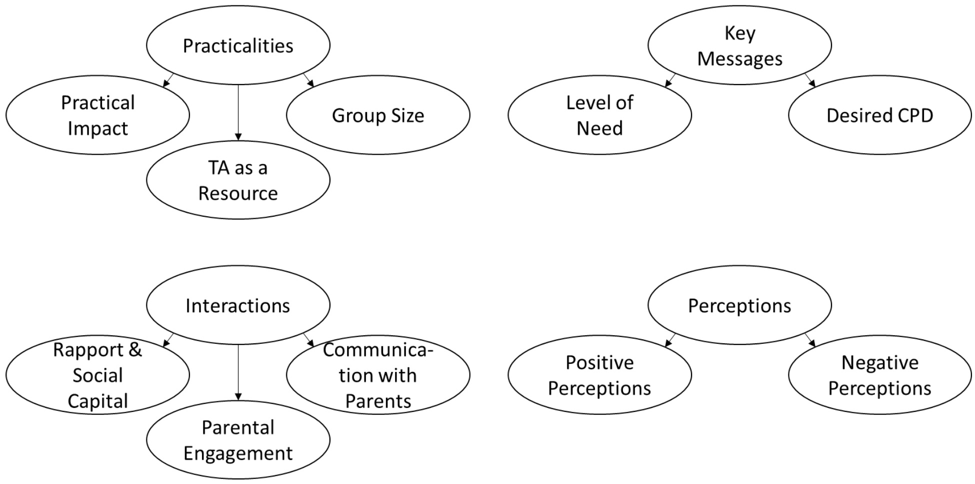
The feedback from the stakeholder consultations is presented in detail in table 4.2 below.

4.7 Results

The first stage of the analysis identified four top-level themes and 10 sub-themes which were drawn out of the data. These themes are displayed in the figure below. See Table 4.2 below for the stakeholder data.

Figure 4. 2

Themes and sub themes for the semantic content of the data



As shown in figure 4.2 the themes are: Practicalities (Impact; Teaching Assistants (TA) as a resource; Group size), Key messages (Level of need; Desired Continuing Professional Development (CPD)), Interactions (Rapport and social capital; Communication with parents; Parental engagement), and Perceptions (Positive; Negative).

These key themes were considered as barriers and enablers to successful language intervention research studies and are set out under phase 1 in the table below. Not all themes that arose from phase one of the analysis resulted in recommendations. The themes 'impact of sessions in the classroom', 'level of need', and 'positive perceptions' are therefore not included in the table.

The themes were linked to the components of the theoretical domains framework to establish the related behaviour change domains, and then to the COM-B model (Cane et al., 2012; Michie et al., 2005; Michie et al., 2011). This procedure was used to generate evidence-based recommendations for future work evaluating BEST, as well as for language intervention research more generally. These are displayed in table 4.2, and the full set of recommendations are reported below.

Table 4. 2

Phases of Generating Recommendations for Future Language Intervention Research

Phase 1: Key themes	**Phase 2: Application of the TDF Framework**	**Phase 3: Application of the COM-B Model**	**Phase 4: Linking COM-B components to intervention functions**	**Phase 5: Creating recommendations based on stakeholder consultation (and relation to intervention components)**
Barrier: TA as a resource	(3) Social/professional role and identity	Reflective motivation	Education, persuasion, incentivisation, coercion	Ensure that TAs are aware of their role and theoretical underpinnings (education) Consider whether TAs should be used, and how they could benefit more if so (incentivisation)
Barrier: impact of group size	(11) Environmental context and resources	Physical opportunity	Restriction, environmental restructuring, enablement	Consider limiting group size to 3-4 children (environmental restructuring)
Enabler: providing targeted CPD	(1) Knowledge	Psychological capability	Education, training, enablement	Offer tailored specific and broad CPD (education, training)

Enabler: rapport and social capital	(12) Social influences	Psychological capability	Education, training, enablement	Maintain working relationships and strong communication (enablement)
Barrier: Communication with parents	(12) Social influences	Reflective motivation	Education, persuasion, incentivisation, coercion	Written and verbal communication should be clear and easy to understand (education)
Barrier: engagement with parents	(8) Intentions	Reflective motivation	Education, persuasion, incentivisation, coercion	Consider electronic forms; make instructions clearer (education); explicitly remind and monitor homework uptake (persuasion)
Barrier: Negative perceptions	(6) Beliefs about consequences	Reflective motivation	Education, persuasion, incentivisation, coercion	Remind school staff of theoretical reasons behind key principles (education)

Table 4. 3

Consideration of stakeholder consultation data in terms of the Theoretical Framework of Acceptability; constructs and definitions from Sekhon et al, (2017)

Acceptability construct	**Definition**	**Relevance to stakeholder consultation data**
Affective Attitude	How an individual feels about the intervention	Affective attitude was not probed in detail during stakeholder consultation. The face-to-face meetings conducted with schools during the recruitment phase (see chapter 2) were generally very positive, and all schools apart from one initially signed up to take part in the intervention. This implies that the study generally appealed and was acceptable to Head Teachers, since they were happy to take part. Notably this construct of the TFA applies to perceived acceptability of the intervention before it took place, which the stakeholder consultation did not specifically investigate.
Burden	The perceived amount of effort that is required to participate in the intervention	Burden examines perceived amount of effort required to take part in the intervention before participating. Stakeholder consultation suggested that teachers perceived the intervention as being a burden on resources, especially on TA time away from the classroom. This aligns closely with the responses to stakeholder consultation hence the recommendation being made to consider whether the TA should be included in future school-based research.
Ethicality	The extent to which the intervention has good fit with an individual's value system	As with affective attitude, ethicality wasn't explored explicitly since in order to take part in the study every Head Teacher and Parent/Carer provided fully informed consent, and queries were addressed at this point. The main issue raised concerning ethicality was the detailed nature of the consent forms which has been addressed in the recommendations section above. No other issues concerning ethicality were raised during stakeholder consultation.
Intervention Coherence	The extent to which the participant understands the intervention and how it works	Some teachers reported understanding the underlying principles of BEST, including how the intervention built upon itself, and that it did not matter if children were not always paying attention because they were still benefiting from the input presented during sessions. Conversely, some commented that the intervention was repetitive and boring, and that it wasn't beneficial because children already had experience of using visual prompts. This feedback suggested a lack of intervention coherence for some staff and contributed to recommendation number 4, which involves placing greater emphasis on conveying the key messages in accessible ways and checking that they are understood and remembered.

Opportunity Costs	The extent to which benefits, profits or values must be given up to engage in the intervention	Teachers reflected on the perceived loss of resources due to TAs time being spent in BEST sessions. This is perhaps due in part to the lack of flexibility due to working with the external research team, and while the loss of TA time at specific times was clearly a burden, it is possible that the opportunity cost would be lower if school staff were able to deliver the intervention in their own time. Future research could endeavour to monitor implementation of BEST by schools, and whether opportunity costs are reduced when the school is able to act autonomously.
Perceived Effectiveness	The extent to which the intervention is perceived as likely to achieve its purpose	Teachers mentioned the intervention taking some children forward and discussed the specific benefits for some children taking part, while others spoke of the intervention taking children backwards. Staff reported that there were no perceived benefits to taking part as a TAU school. The TFA definition also suggests that the purpose needs to be made clearer to school staff. The purpose of the present study was to test efficacy, however if found to be effective the purpose would undergo a shift towards explicitly improving language outcomes. Overall perceived effectiveness is mixed and does not necessarily reflect the empirical evidence presented in the previous chapters.
Self-Efficacy	The participant's confidence that they can perform the behaviour(s) required to participate in the intervention	Many school staff mentioned the simplicity of the intervention, suggesting that they would find its implementation manageable and that they have strong self-efficacy. Other staff mentioned wanting to receive training and to watch sessions again before implementing BEST themselves, indicating less perceived self-efficacy and the need for more support. TAU school staff in particular indicated that they would like to receive more training and observe more sessions before implementing them in their schools.

4.8 Barriers and Enablers

The feedback gathered from stakeholder consultation with school staff was extremely varied and highlighted strengths as well as areas for improvement in future trials. This work is therefore instrumental for establishing what can be learned from the present study to impact the implementation of future research.

The thematic analysis generated themes that can be seen as either enablers or barriers to successful trial implementation (phase 1). These themes were then linked to the Theoretical Domains Framework (phase 2), and the theoretical domains were plotted against the COM-B model (phase 3). The relevant intervention functions were identified in the COM-B model (phase 4), and these were used to generate evidence-based recommendations (phase 5). Themes are discussed in this section in the order that they appear in the table above. Some themes were not identified as barriers or enablers and were therefore not considered further in this analysis.

4.8.1 Barrier: Teaching Assistants as a Resource

The feedback identified the inclusion of TAs in the BEST sessions as a barrier to school participation. TA involvement resulted in compromised staff to pupil ratios in the classroom and also caused delay for the researchers when cover had to be found for TAs working 1:1 with children. TAs reported feeling underused which also links to the negative perceptions barrier described below (see 4.8.7). Reasons for the repetitive nature of BEST were explained to teachers and teaching assistants, however this analysis revealed that knowledge needs to be shared more effectively, since key messages were not always well communicated. The project was seen as a drain on resources in some classrooms due to its inclusion of TAs, as well as the length of sessions (see 4.8.2 below).

4.8.2 Barrier: Impact of Group Size

Sessions involving higher numbers of children were perceived negatively, and staff mentioned that some sessions lasted for too long and that children's behaviour and attention became challenging due to the session length. The project was perceived as a drain on classroom resources in one school because of the duration of the sessions as well as the use of TAs, as discussed in 4.8.1 above.

4.8.3 Enabler: Providing Targeted CPD

During the study the CPD offering played an important role in aiding recruitment, especially for schools in the TAU arm that did not directly benefit from therapy delivery during the study. The stakeholder consultation identified that schools would like to receive CPD on promoting language in the classroom as well as specific training on areas such as how to implement BEST. This is also reflected by the level of need described by school staff, since some schools have a high proportion of children with language difficulties.

4.8.4 Enabler: Rapport and Social Capital

Feedback about the rapport established between the Author and schools was very positive, highlighting the crucial role that interpersonal relationships play in creating and maintaining strong working relationships in this type of research. Communication with schools, researcher availability and flexibility were all positively perceived. It is likely that good communication led to retention and good relationships with schools, possibly suggesting communication was more effective, or better working relationships were established, within the schools that carried on participating into the second wave.

4.8.5 Barrier: Communication with Parents

Particular issues were identified with the content and complexity of information in the consent forms and information sheets. Teachers identified that the communication style was not always appropriate for the language and literacy levels of the parents. Despite this, the stickers given to children after assessment and therapy sessions were well received, and teachers pointed out that they signalled to parents that their child had been working on the project which seemed to be a positive reminder and allowed parents to feel proud of their children.

4.8.6 Barrier: Engagement with Parents

There is a lack of consensus on whether face-to-face meetings are beneficial to parents. During the study all schools were offered face-to-face meetings, however the uptake was low. When meetings did take place, teachers felt they were not beneficial or well received by parents. On the other hand, when meetings did not take place it was suggested by teachers that they would have been beneficial. Lack of these meetings may also have impacted on successful communication with parents which is discussed above.

4.8.7 Barrier: Negative Perceptions

BEST was sometimes perceived as repetitive and boring, and some feedback suggested a lack of understanding of the structure of BEST (e.g. repeated exposure to the input). This highlights the difficulty of conveying information about the study to schools, especially as the knowledge needs to be retained over the duration of the study or therapy implementation. Although careful efforts were made to convey this information over the course of the study, the results of this analysis suggest that more needs to be done to ensure understanding of the design and implementation of BEST.

4.8.8 Acceptability

The stakeholder consultation feedback was considered against the theoretical framework of acceptability (TFA) (Sekhon et al., 2017). The teacher responses suggested that they had mixed perceptions of acceptability. The questions asked during the consultation did not probe all constructs in the TFA, however teachers specifically reflected on four of the constructs. Table 4.3 above summarises the findings from the stakeholder consultation and researcher's observations in terms of acceptability.

4.9 BCT and Acceptability Recommendations

Overall, perception of acceptability was varied and depended on factors including the treatment arm the school was included in. Acceptability was generally higher for schools in the BEST arms. Taking part in the study as a TAU school resulted in low perceived effectiveness and lower self-efficacy. Other aspects of acceptability could be improved through practical adaptations, such as shortening the length of sessions and considering the role of TAs. These suggestions are reflected in the recommendations.

The empirical results of the trial reported in this thesis provide evidence for the efficacy of BEST and suggest that a future larger-scale study is warranted. The consultation data was therefore used to generate the evidence-based, theoretically motivated recommendations below which aim to tackle the barriers discussed in this chapter. These recommendations apply not only apply to future research evaluating BEST, but also to other school-based intervention studies.

1. The Author and RAs should maintain good working relationships with schools by prioritising flexibility and strong communication, showing awareness of the issues facing teachers and parents, and answering questions and addressing concerns adequately. They should be available to answer questions through a variety of methods.
2. All written and verbal communication with parents/carers, including consent forms and information sheets, should be clear and tailored to their language and literacy needs. Face-

to-face meetings and QR codes linked to explanatory videos should accompany written communication to ensure it is easily understood. Information presented to parents should be reassuring regarding their children's identified difficulties.

3. An electronic method of completing questionnaires should be considered.
4. Greater emphasis should be placed on communicating the key messages and theoretical rationale of BEST and ensuring that they are understood and remembered. Teachers should be presented with posters displaying BEST principles, explicitly highlighting repetition and turn taking as theoretically motivated elements of the therapy.
5. Sessions should be made more manageable by ensuring sufficient training of the persons delivering BEST and implementing smaller groups e.g. 3-4 children.
6. The advantages and disadvantages of including TAs in the research should be seriously considered in light of the feedback from schools about the disruption it causes and the Author's reflections that sessions were manageable without including TAs. If TAs are included, they should receive clear explanations of BEST. Ways in which they could benefit further from taking part in the sessions should be identified.
7. Future research should explicitly monitor uptake of the BEST homework, possibly using a diary format and/or a text messaging reminder service. Videos should be created to demonstrate how to complete the homework. Effective systems should be put in place to remind teachers to send booklets home and encourage parents to complete homework with their children.
8. CPD training and materials offered to schools in return for taking part in the study should focus not only on BEST but also on general strategies for improving SLCN in the classroom. There should be a focus on issues identified by teachers in the summary table above.

4.10 Researcher Perspectives

The qualitative methods above were also applied to the Author's field notes which were made throughout the course of the study. The following discussion summarises the main themes encountered within these notes and further recommendations are drawn.

4.10.1 Barriers to Effective Implementation

The stakeholder consultation analysis identified issues from the perspective of school staff who were involved in the trial. The pilot and feasibility phase of intervention evaluation ensures that the intervention can be delivered as intended (Craig et al., 2008), and that any barriers to effective implementation have been adequately addressed before further research takes place.

The researcher made field notes and kept a reflective log, recording barriers encountered throughout the study. The barriers fell under 4 main categories: 1) schools; 2) families; 3) methodological barriers; and 4) ethical barriers. These are discussed below, and prospective solutions are suggested. Some barriers were raised in the consultation evaluation above, whilst others are novel to the Author's reflections.

4.10.2 Schools

Due to the nature of the research, the participating schools and specific members of school staff played a crucial role in the successful implementation of the research. School-based research is inherently challenging due to the conflicting demands placed upon teachers (Bartlett et al., 2017). Despite this, several barriers were identified through the Author's experiences that could be ameliorated in future research.

4.10.2.1 Training school staff

Staff received training before being involved in BEST sessions (see Appendix 12 for TA manual). Training involved a demonstration of the resources and structure of the sessions and discussion of the BEST key principles. Questions were answered and school staff were provided with the BEST manual and a manual written specifically for TAs for future reference (see McKean et at., 2013 for the BEST manual).

The training approach was perceived as effective at the time of delivery by the Author, however some of the staff feedback recorded in the consultation summary above suggests that core messages were not well conveyed. A video should be used to provide clear, standardised training in future. Posters conveying the BEST key messages should also be developed and displayed in classrooms where BEST sessions take place. This approach would involve recruiting children whose parents/carers were happy for them to appear in the video for training purposes, and would therefore prove difficult, but would be worthwhile for training school staff.

It was not possible to deliver the training at the same time in relation to the first BEST session in each school due to scheduling difficulties. Some staff therefore received the training session immediately before delivering the first session, while others had a gap of up to a week before delivering the therapy. Although it is unlikely that this affected delivery, a training video could be sent in advance of the first session to allow training to be delivered in a consistent manner.

4.10.2.2 Variation in school staff and staff absence/unavailability

The study was designed whereby the researcher would fill the ‘first adult’ role and TAs would fill the ‘second adult’ role (see Table 2.1). In practice however, some schools were unable to consistently provide a TA to support all sessions. In these scenarios either another member of staff filled the ‘second adult’ role, or the researcher delivered the session alone (see 2.15.4). Across the whole study the staff members who filled the role of ‘adult 2’ therefore included TAs, class teachers and SENCOs and many sessions were delivered without the support of a second adult.

It is likely that the varying staff roles impacted on therapy delivery to an extent, due to the differing levels of training and the impact of a school staff member who knew the children. In particular, the Author's observations suggest that those with greater levels of training seemed to find it harder to implement the BEST key principles (i.e. not giving praise during the sessions and not providing the child with forced choice alternatives or correcting the child's utterances), likely because these may be considered counterintuitive when compared to 'standard' teaching practice (Keller et al., 2005). When issues regarding adherence to BEST principles occurred, the researcher was able to rectify them through discussions with staff, but it is worth acknowledging the effect that this could have on sessions that were are as closely monitored.

Sessions may have also been affected when school staff were unavailable to support BEST sessions. TAs were initially selected to support BEST sessions because they have close relationships with the children and children were accustomed to following their instruction, for example regarding behaviour. Behavioural differences may therefore have occurred across sessions depending on the presence of a school staff member. It would be beneficial to control this more closely in future research.

4.10.2.3 Organisational and Practical Barriers

Schools had difficulty providing suitable spaces for therapy sessions for the duration of the study. BEST sessions were therefore carried out in a number of settings (e.g. classroom, corridor, library), and with different physical formats (e.g. sitting around a table; on the floor etc). Some sessions were conducted in rooms that had a lot of visual distraction (e.g. that looked out onto the playground or into corridors) which the Author felt negatively affected children's attention and increased background noise levels. The Author and RAs employed techniques such as ensuring children faced away from distractions where possible, but ultimately this is a caveat of conducting research in schools and is difficult to control.

Working around the classroom timetable and school staff's availability to complete the BEST sessions twice weekly was also a barrier to effective implementation. It was challenging to accommodate regular sessions in all schools around school trips, assemblies and other timetabled activities. Many schools also struggled in terms of staff to pupil ratio, and on some occasions a TA would have to delay the session in order to find cover staff, or leave a class understaffed to complete the BEST sessions.

The timetabling of nursery and reception classes posed difficulty for arranging sessions since each had a unique setup (e.g. some ran morning and afternoon sessions, some ran Monday morning to Wednesday lunchtime and Wednesday lunchtime to Friday afternoon sessions, and some ran both simultaneously). It was therefore necessary to visit some schools up to four times per week rather than the standard two sessions per week to deliver sessions to all children if they were part of separate cohorts and in school at different times. Future research protocols should therefore build in contingency time to accommodate multiple BEST sessions at each site.

4.10.2.4 Session Length

During the first BEST sessions of the first wave the researcher and school staff immediately felt that sessions were too long in duration. Children's concentration was negatively impacted and teachers were unhappy that TAs were absent from the classroom for such long periods. As a result, the Author and supervisory team devised an adaptation to the BEST procedure whereby during phase 2 of each set of sentences (the output phase), the researcher or TA would manipulate the toys while all children's responses to "what's happening?" were elicited. Once all the children had given responses to a sentence one child would be given the toys to act out the sentence. The next child would be given the toys to act out after all children had responded to the next sentence, and this process continued for all sentences so that all children had turns to act out sentences using the toys. To compensate for the children's reduced interaction with the toys during the structured therapy sessions, at the end of each session all toys were placed in the middle of the table, and children were able to play with them for at least 5 minutes. In line with the principles of BEST this was not compulsory, but children enjoyed this and generally played with the toys enthusiastically. It was decided that the researcher would not comment on what the children made the toys do during this free play period, so that children in different BEST groups were not exposed to different inputs as a result of this change.

Recording using a dictaphone and transcribing sessions immediately afterwards was also implemented early during the first wave which saved time and reduced the burden on the researcher. The primary outcome of this amendment to the original BEST manual was that sessions were reduced in length. A second anecdotal outcome was that observed behaviour and concentration were improved, although this was not measured empirically.

TAs and teachers reported that sessions were more manageable when reduced in length, and this was particularly the case for the larger groups of four to six children. The Author felt that overall quality and theoretical effectiveness of the intervention were not affected by this change. The fundamental principles of BEST remained intact at all times, and it is possible that children were better able to access the intervention due to increased attention and improved behaviour during the sessions. It is also possible that the original BEST manual specifying that children should manipulate the toys for every output sentence added more distraction than benefit, but further research is required to draw conclusions on this.

4.10.2.5 Child Recruitment

Despite the strict recruitment protocol in place (see chapter two), in some schools teachers remained unsure of which children should be put forward for assessments; in some cases forms were sent to children who were clearly ineligible (e.g. they had strong language abilities for their age or were outside of the eligible age range). This indicates that information needs to be presented more clearly and that teachers should be reminded of the crucial elements more frequently throughout the recruitment phase of the research.

4.10.2.6 Attrition

The school attrition rate was higher than anticipated prior to commencement of the study. During the first data wave two out of eight schools (25%) dropped out before the beginning of assessments. One of these was due to schools' concerns about the introduction of General Data Protection Regulations (GDPR) causing delay that surpassed the deadline for baseline assessments to be completed. The second drop-out was likely due to a lack of communication between the Author and the school, since despite multiple meetings and communication, the school reported feeling unsure about which children to put forward and their concerns could not be addressed despite every effort. One school dropped out after the study commenced, which was attributed to mismatched expectations between the Author and the school staff. Although situations such as these are unlikely to be avoided in all cases, these examples highlight the need for clear communication of key messages and expectations of the study.

4.10.2.7 School Questionnaire Return Rate

Teacher's questionnaire response rate was lower than expected. Questionnaire response rate is notoriously poor in research of this nature (Sturgis et al. 2006), however future research should anticipate this and build in designated time for teachers to complete forms. A more systematic means of completing questionnaires for both teachers and parents/carers would be beneficial in future research, and should include automatic reminders, for example via text messaging, to encourage timely completion of questionnaires.

4.10.3 Families

Parental engagement varied between individuals and throughout the study. This section describes the difficulties encountered when working with families and suggests possible solutions.

4.10.3.1 Return of Consent Forms

Parental consent form return rates varied between schools. This may have been influenced by the teacher's engagement with the project and subsequently the degree to which they encouraged families to participate. Indeed, some teachers reported spending time talking with parents about the study, whilst others sent the forms home then did not follow-up with parents. The researcher offered to run face-to-face meetings with families as discussed in the consultation summary above, but uptake was low.

4.10.3.2 Parental Questionnaire Return Rate

As with consent forms, return rate for parental questionnaires was generally low, and decreased at each time point. Teachers were asked to encourage parents to complete and return forms, but this did not greatly improve return rates. Ultimately, poor questionnaire response rate is a common barrier in research of this nature (Sturgis et al., 2006), although methods such as sending questionnaires electronically and reminding parents via text message could be piloted in future studies. If time and resources permitted, meeting with parents face-to-face to complete questionnaires may also yield a greater response rate.

4.10.3.3 Homework compliance and monitoring

Parents were asked to complete the homework booklets that were sent home following each therapy session. Written instructions were provided to parents at the beginning of the study and are summarised in chapter two. The present study did not directly measure parental homework uptake, although based on teacher report it is likely that uptake was highly variable; some parents self-reported completing homework to teachers, whereas others did not. This potential difference in input that children were receiving at home, based on how often homework was completed with parents, is a confound that needs to be addressed in future studies. As suggested above for questionnaires, parents could be reminded to complete homework via text message. Future research should explicitly monitor uptake via a homework diary or similar method.

4.10.4 Study Methodology

4.10.4.1 Blinding

As discussed in chapter two, blinding is essential for conducting robust intervention research (Eldridge et al., 2016; Schulz et al., 2010). Research assistants were blinded to the treatment arm of the schools they were working within to avoid subconscious biases in assessments.

A number of measures were put in place to maintain blinding. Firstly, email signatures were used by Research Assistants to remind participating schools not to disclose their treatment arm during communication. Where applicable, Research Assistants also reminded teachers in person that they were blinded to intervention arm, and therefore any questions pertaining to the treatment arm or BEST should be referred to the Author. Treatment arm was coded on consent forms so that the Research Assistants could collect them from schools without blinding being compromised. Within the University working environment, the Author and the Research Assistants stored data in separate locked filing cabinets, so that the treatment arm assigned to each school was not included on any documents that were seen by the RAs. Conversations and meetings were also carefully considered within the shared office space so that no information was inadvertently revealed regarding treatment arm.

4.10.4.2 Signing

The treatment arms were implemented in order to test the effects of sign, however two issues were encountered during the study regarding blinding.

First, as is further discussed in chapter five, there were observed limitations to how much of the signing was observed by children due to difficulties with maintaining children's attention throughout the study (see 5.4.1). Similar difficulties have been reported by Vogt & Kauschke (2017).

Secondly, as discussed in the important changes to the protocol section of chapter two, the method of delivery had to be altered slightly when the second adult was not available to support the session. While neither of the changes above are likely to have affected overall outcomes, it is important that future research anticipates such changes to the methodology so that proposed adaptations can be made in advance, rather than post-hoc. Both issues are discussed further in chapter five.

4.10.5 Ethical Issues

4.10.5.1 GDPR

Compliance with General Data Protection Regulations (GDPR) became a legal obligation on 25/5/2018, coinciding with the first data wave, when initial consent forms had already been distributed. A further consent form explaining the GDPR implications was therefore sent to parents/carers, to explain exactly how their and their children's data would be held and analysed, and the process by which they could withdraw from the study should they wish to. Face-to-face meetings were offered to all schools and parents/carers to discuss the implications of GDPR. One school arranged a face-to-face meeting and parents from that school did not express concerns regarding GDPR. The consent forms were merged for the second and third data wave to avoid the need for a second GDPR form, but content remained the same.

4.10.5.2 Ethical Approval

It was not immediately clear whether the present evaluation research required ethical approval through the NHS Research Ethics Committee (REC) via the Integrated Research Application System (IRAS), or whether ethical approval should be sought through Newcastle University. This study highlighted a grey area in the appropriate route of obtaining ethical approval for similar research, since children were recruited through schools rather than through the NHS, but it was also possible that they were on clinical caseloads for NHS Speech and Language Therapists. Guidance from the IRAS REC eventually stipulated NHS ethical approval was unnecessary, since participants were not recruited *by virtue* of being on NHS caseloads.

4.10.6 Additional Recommendations from Author's Field Notes

The above discussions highlight several barriers encountered in the present study that were not identified in the thematic analysis of stakeholder consultation with school staff. This is likely because they affected the Author and RAs, rather than the schools. Four further recommendations were therefore drawn from the Author's experiences:

9. Develop a training video for Teaching Assistants to ensure training can be delivered at uniform times and in a standardised manner across each school in the study.
10. Monitor the effect of TAs' presence on therapy outcomes, children's behaviour and TA and researcher perceptions to further determine whether the inclusion of TAs is justified.
11. Build in contingency time for when multiple groups run simultaneously in the same school due to timetabling difficulties.
12. Ensure the protocol considers decisions about what adaptations to therapy delivery will be made a-priori when variations occur, e.g. when supporting staff are not available.

4.10.7 Limitations of the Thematic Analysis

There are a number of limitations to this phase of the research. Firstly, the thematic analysis applied to the data has not traditionally been applied to stakeholder consultation summary data. This method was chosen for the present study as it was appropriate for summarising the main themes emerging from the data without compromising participants' anonymity. The analysis is therefore unlikely to generalise to other settings or studies, as it is a summary of the opinions of school staff from this specific project. The effectiveness of the thematic analysis technique being applied to this data is also limited as the interviews were not transcribed verbatim. Despite this, the analysis did identify themes that were anticipated by the Author.

The consultation was not conducted to probe The Framework of Acceptability (TFA), and the data were applied post-hoc to the framework. Therefore, while the responses identified some views on acceptability, it is possible that other views may have been gathered if questions had been tailored to probe the TFA framework. The data doesn't allow for clear differentiation between acceptability of the BEST intervention and acceptability of participation in the research trial as a whole, which should also be the focus of future research.

4.11 Conclusion

By following the recommendations set out in this chapter, future research of a similar nature can build on the current evidence base and improve efficiency by avoiding ineffective or time-consuming methods when a better alternative has been identified. Issues raised in this chapter are discussed further in chapter 5.

Chapter 5

Discussion

5.1 Introduction

This thesis reports the findings of a non-randomised pilot study evaluating the efficacy of Building Early Sentences Therapy (BEST), the role of sign as an adjunct to BEST, and the acceptability and feasibility of the intervention to stakeholders. This chapter begins with a summary of the content and findings of the previous 4 chapters. The research questions are then addressed in turn and are interpreted in the context of literature relating to intervention research and the potential of usage-based therapies to support children with language difficulties. Strengths and limitations of this study are then presented, followed by the theoretical and clinical implications of this work. Future avenues for research are then highlighted before the Author's concluding remarks.

In light of the previous literature reviewed in chapter 1, the following a-priori hypotheses were made:

1A) BEST will be effective for supporting generalised and targeted language structures in children with language difficulties compared to Treatment as Usual.

1B) Limited available evidence also suggests that the inclusion of a signing system in BEST will provide additional support for generalised and targeted language structures compared to BEST without sign.

1C) There are no a-priori hypothesised effect sizes based on limited previous evidence for interventions of this nature.

2) There will be greater acceptability to parents and teachers due to the quasi-experimental design employed in the present research, compared to the student pilot and feasibility RCT (see chapter two).

5.2 Thesis Summary

Chapter 1 identified the benefits of early intervention for targeting children with language difficulties prior to the establishment of more fixed trajectories (McKean et al., 2017), however few interventions have undergone the rigorous evaluation process defined by the Medical Research Council (Craig et al., 2008). Evidence for existing interventions is therefore limited, with implications for clinical decision making and service delivery. Further, few interventions are underpinned by explicit theories of language development.

Usage-based theory has potential for informing the development of language interventions since its principles readily translate into effective dose form components (Frizelle et al., submitted). One such usage-based intervention is Building Early Sentences Therapy (BEST), designed to support children with expressive and mixed (expressive and receptive) language difficulties. BEST aims to harness cognitive mechanisms underpinning typical language development to facilitate mapping and analogy across items and structures, and to support cognitive processes known to be challenging to children with DLD. Sign is a hypothesised 'active ingredient' of BEST, providing additional visual cues and facilitating bootstrapping.

Parents and teachers involved in the previous student-led randomised controlled pilot study found the randomisation process to be unacceptable. The present study therefore employed a quasi-experimental design and sought to further explore themes of feasibility and acceptability. The study sits within the piloting and feasibility stage of the MRC guidelines, measuring intervention efficacy under ideal conditions.

Children made significant progress on all measures excluding the comprehension subscale of the NRDLS. Sign was particularly effective for promoting progress on standardised production outcomes.

A stakeholder consultation was carried out and found the procedures to be generally acceptable to teachers. Evidence-based recommendations for future research are presented.

5.2.1 Research Questions Addressed in this Thesis

The following questions were addressed by the present research.

1A) Is an intervention underpinned by usage-based theory (BEST, with and without additional sign) more effective than treatment as usual (TAU) for improving language and associated functional communication outcomes in pre-school children with language difficulties?

1B) Is signing a necessary ‘active ingredient’ of BEST, and does it differentially affect progress on different language goals (sentence structure and morphology)?

1C) What are the point and interval estimates of effect sizes across outcomes?

2) What theoretically motivated recommendations can be made from applying Behaviour Change Theory to stakeholder consultation outcomes, to increase acceptability of the research and therefore inform the next stage of BEST development?

The research questions are addressed in order, followed by a discussion of the study strengths and limitations and implications.

5.3 Research Question 1A: Is an intervention underpinned by usage-based theory (BEST, with and without additional sign) more effective than treatment as usual (TAU) for improving language outcomes in pre-school children with language difficulties?

To address this question, longitudinal multi-level models accounting for fixed and random effects were implemented for the following outcome measures: standardised production and standardised comprehension (New Reynell Developmental Language Scales (NRDLS) (Edwards et al., 2011)), targeted sentence content, targeted sentence morphology (Targeted BEST Assessment (McKean et al., 2013)), generalised sentence content and generalised sentence morphology (Generalised BEST

Assessment, designed by the Author for this study). The resulting outcomes were used to establish whether BEST is more effective than TAU for improving language outcomes in 3;5-4;5-year-olds.

The child's score on each outcome measure was predicted by the timepoint, intervention arm, child's age, child's received dosage, school IDACI score and child's language profile, and the interaction between time point and intervention arm. Random intercepts for *school* were predicted by the following random slopes: timepoint, dosage received and child's age in months at entry. Random intercepts for *participant* were predicted by random slopes for timepoint.

5.3.1 Standardised Production

Standardised production was measured using the norm-referenced expressive subscale of the NRDLS. On average children in the BEST arms made more progress than children receiving TAU. This difference was significant for children receiving BEST with sign only, however closer inspection of the coefficient plots 3.20 and 3.21 show that at outcome the lower confidence interval was close to zero for BEST without sign at outcome, indicating a positive trend for BEST without sign. BEST was beneficial for the children's generalised production scores in this study. Children maintained significant progress at follow-up compared to TAU.

The confidence intervals appear relatively wide, as is often the case with smaller sample sizes. Despite this, children receiving BEST performed better than TAU across all production outcomes. This finding was robust to the sensitivity analysis (plots 3.22 and 3.23).

This outcome was anticipated since BEST was designed to support children's expressive language. Nevertheless, the amount of progress, particularly by children receiving BEST with sign, is large (J. Cohen, 1988). As the NRDLS scores are age adjusted, one might predict minimal change over the study duration, and these changes therefore represent progress. There is evidence of efficacy of BEST which warrants further testing to establish whether these effects are maintained in an effectiveness study.

5.3.2 Standardised Comprehension

Next we move to comprehension outcomes, which were measured using the receptive subscale of the NRDLS. Children receiving BEST did not make significant progress on comprehension across the study. Indeed, children receiving BEST without sign at follow-up performed worse on average than TAU which did not occur for any other outcome or timepoint during the study. The sample had made progress at follow-up, although this was not limited to children in the BEST arms, more likely reflecting general progress made over time and due to the variety of interventions received by all children.

It is possible that children may have continued to make progress with comprehension after the six-week maintenance period. Camarata et al. (2009) showed that children make incidental receptive gains during expressive intervention involving modelling, recasting and imitation, although this intervention was carried out 24 sessions over 12 weeks and was therefore one third longer than BEST in duration.

The timepoint variable was significant at follow-up; this was not specific to the BEST arms and is therefore more likely to be a reflection of general progress made over time and due to the interventions received by all groups as part of TAU.

5.3.3 Targeted sentence content

Targeted sentence content was assessed using the Targeted BEST assessment: a specific, non-standardised measure that was designed to measure targeted content and morphology as part of the BEST intervention. Children receiving BEST both with and without sign made significant progress on targeted content at both outcome and follow-up. Targeted content showed the largest gains of any assessments at outcome and the effects were maintained at follow-up. It was unsurprising that BEST improved children's ability with the targeted content compared to children receiving TAU given that it assessed content they were exposed to many times over the course of the intervention. The maintenance of the effects six weeks post intervention suggest that these changes did not reverse and that progress with targeted content was robust.

5.3.4 Targeted Sentence Morphology

Targeted sentence morphology was measured using the same Targeted BEST Assessment as targeted sentence content. Children receiving BEST made progress on targeted morphology. Whilst this progress ceased to be significant at follow-up for BEST with sign, inspection of plot 3.21 shows that the lower confidence interval marginally crossed the TAU reference line, suggesting that children still maintained progress, although the effects were reduced by follow-up.

5.3.5 Generalised Sentence Content

The Generalised BEST Assessment probe was designed by the Author for this study to measure whether any change detected for the Targeted BEST content extended beyond the target vocabulary. To achieve this the assessment presented different content to that of the Targeted BEST Assessment, which was matched for age of acquisition and using the same predicate argument structures.

Children receiving BEST showed some evidence of progress compared to TAU at outcome, but this change appeared genuinely larger at follow-up, when the results became significant for both BEST Arms. Although not conclusive at this stage, these findings suggest that BEST (with and without sign) may have a longer-term impact on children’s ability to generalise their knowledge to novel sentence content.

5.3.6 Generalised Sentence Morphology

Generalised sentence morphology was measured using the same Generalised BEST Assessment as generalised sentence content. Children receiving BEST made some progress on generalised morphology although this progress was non-significant compared to TAU. Unlike with the targeted morphology, children made additional progress between outcome and follow-up, rather than their average scores reducing. These results suggest that the positive effects on targeted morphology may have had some effect on generalised morphology.

5.3.7 Sensitivity Analysis

All children were evaluated on all measures detailed in section 2.11, excluding those who missed assessments (see Table 3.1). As discussed in section 3.7, some children did not begin the study with clinically significant language difficulties on one or other of the NRDLS subscales (see Table 2.4 for an explanation of language profiles and Table 3.2 for descriptive statistics). To ensure that overall progress was not driven by children who did not have a difficulty below the 16th centile to begin with, these children were removed from the sample and the analysis was conducted again. This was also important since children with these profiles would be less likely to receive targeted or specialist intervention in a 'real-world' setting.

The sensitivity analysis in chapter 3 demonstrated that children who scored above the 16th centile on a particular measure of the NRDLS did not drive the overall results, since removing them did not change the overall observed rate of progress for most outcomes. Encouragingly, the finding relating to BEST without sign at outcome became more robust for NRDLS production scores, although overall scores for BEST without sign at follow-up for morphology were slightly reduced. The sensitivity analysis suggests that one can have confidence in the initial results due to the limited impact of children with higher baseline scored (see plots 3.20 and 3.21).

5.3.8 Question 1A Conclusions

The results of this study provide support for the efficacy of BEST with respect to production, targeted content, targeted morphology and generalised content outcomes. This suggests that BEST is efficacious in improving generalised and targeted expressive language for children with language difficulties. The effects are generally maintained six weeks post intervention and extend to the standardised NRDLS expressive subscale suggesting that BEST positively impacts underlying representations of predicate argument structure and the results are not driven by practice effects alone.

Interestingly the results for generalised content and morphology do not follow the same trend as targeted content and morphology, as children make more progress at follow-up for generalised structures but less with targeted. This could reflect some practice effects that were present for targeted content and morphology that diminished by follow-up, but the confidence intervals are wide for all outcomes and suggest that further research with a larger sample would be required to draw robust conclusions.

The results for research question 1A must be interpreted in light of both the stage of the evaluation process and high level of involvement of the Author. The study evaluates intervention efficacy which is measured under ideal conditions (Robey & Schultz, 1998). This arguably results in a high level of treatment fidelity, reliability and control due to the Author's close involvement with schools and delivery of all BEST sessions. If schools were to implement BEST outside of study conditions, treatment fidelity and dosage may be impacted due to the competing demands upon school staff time and variation in skills and experience, although this assumption requires further research (McCartney et al., 2011). SLTs/SLT Assistants may be better suited to delivering BEST if serious issues for intervention fidelity were identified in TA delivery, but the issue would then perhaps be whether they could deliver sessions with a high enough dosage given their roles across multiple areas or schools.

5.4 Research Question 1B: Is signing a necessary 'active ingredient' of BEST, and does it differentially affect progress on different language goals (sentence structure and morphology)?

The results of this study support the inclusion of sign in the intervention. NRDLS production is the only language outcome with a clear statistical advantage for the use of sign as an adjunct to BEST, however as a standardised language measure this suggests a robust change in children's mental representations as a result of receiving intervention.

Children did not demonstrate progress on standardised comprehension across either time point or intervention arm, although there is some visual evidence of increased progress for children receiving BEST with sign (see 3.4.6). This may be due to the hypothesised benefit of sign for mapping and highlighting the morphological frame within model sentences. It is also likely to represent the small number of children with receptive-only difficulties included in the trial (n=5).

Sign does not appear to impact the remaining non-standardised measures, however this is likely due in part to ceiling effects limiting the upward potential for change, considering how much progress was made, even with BEST without sign.

At follow-up the same pattern occurs as for outcome. There is clear evidence that robust progress with BEST with sign on NRDLS production scores is maintained at follow-up, demonstrating a clear advantage over BEST without sign. We also see maintenance of the progress made with BEST with sign on NRDLS comprehension scores. Interestingly the non-standardised measures all demonstrate slightly higher scores for BEST without sign at follow-up, but these may not represent a genuine effect due to the spread of the confidence intervals and variability within the data. Note that the sensitivity analysis also reverses these effects at follow-up (see Figure 3.21), but again this may not reflect a meaningful difference. Further research with higher statistical power would be required to confirm these findings.

5.4.1 The Impact of Implementation with and without TA Support

As discussed in section 2.15.4, school staff were only available to support around half of the total BEST sessions due to practical constraints. Through completing treatment fidelity ratings of the videos, Dr Sean Pert identified a trend whereby when BEST was delivered by two adults the child tended to focus more on the TA who was manipulating the toys, rather than looking between the TA and the Author presenting the signs (See section 2.9 and 2.15.4 for more detail).

It could therefore be argued that the effect of signing may have been greater had the circumstances allowed the child to attended more to the signs during the sessions. This hypothesis should be examined in subsequent research by ensuring that children are equally exposed to both the toys and the signing, possibly by removing the TA involvement (also see recommendation 6 in chapter 4 which is discussed further below).

Any discussion around implementing signing must account for the fact that it takes time, effort, and financial resources to deliver training for any signing system (see Appendix 14 for the signs needed to deliver BEST). Signing could arguably cause difficulties for comprehension if delivered without a high level of fidelity between and within practitioners, for example if two practitioners were to use a different sign for the same referent. Part of the question surrounding whether signs should be delivered as an adjunct to language intervention must therefore consider the financial and time costs involved in training practitioners to deliver signing reliably. The results of this study suggest that despite the potential costs of practitioners and school staff learning to implement the signs correctly, there is a benefit to this being undertaken, particularly for language production.

Further, as discussed in 1.6.6, there is some anecdotal evidence that parents have concerns surrounding the use of sign, namely that it will delay the onset of oral language. This study provides preliminary evidence for some additional benefits of sign and could help to alleviate parental concern around sign being implemented in BEST. It also suggests that sign does not have a detrimental effect on language learning under the present study conditions.

Through delivering the intervention the Author also reflected on the necessity for future BEST training packages to include sufficient additional training time to learn the signs correctly before intervention delivery commences.

5.4.2 Question 1B Conclusion

The results suggest an overall positive benefit of sign as an adjunct to BEST. Progress on all language outcome measures is greater for BEST with sign than BEST without sign at outcome, excluding targeted content for which significant progress was still demonstrated within the BEST with sign group compared to TAU. The findings suggest there is value in training practitioners to deliver BEST with sign which outweighs the additional time costs, although the results for the non-standardised measures are inconclusive due to ceiling effects. The advantages of BEST with sign are maintained at follow-up for the NRDLS outcomes, suggesting that children have undergone a change in their underlying representations, thus supporting analogy. The positive findings regarding the use of sign could be driven by the sign highlighting the morphological ‘frames’ within which content ‘slots’ are placed (Tomasello, 2003). This also supports previous work suggesting that grammaticality plays an important part in recall (Polišenská et al., 2015).

The evidence from this study suggests that signing increases children’s awareness of the sentence’s morphological frame, which in turn supports analogy and increases production skills. We can hypothesise that the increase in ability to analogy facilitates more abstract representations which also increases the child’s processing capacity and accelerates progress, as it has been argued that more abstract representations increase processing capacity thus enhancing learning potential (Bishop, 2014). Establishing a joint attentional frame between adult and child, with a sign referring to the spoken word at the centre of it, may also provide the child with two complementing exposures to the same input, thus harnessing mechanisms underpinning language acquisition.

This argument counters the findings of Ting et al. (2012) that signing may not be beneficial for children with an underlying language difficulty since oral and signed language may compete for limited processing resources available to the child. Ultimately, more research is needed to establish the specific mechanisms underpinning the role of sign.

The improvements for BEST with sign must be interpreted relative to the overall effect of BEST. There is evidence that sign has an additional benefit and the estimated effect of BEST with sign is larger than the estimate effect of BEST without sign on almost all outcomes. However, these differences were non-significant (see Figure 3.18 and 3.19). Capturing the statistical effects of sign would require a larger sample size since we would expect BEST to have a greater effect than sign due to its multiple components, which are shown to be beneficial elements of language intervention (Camarata et al., 2009; Frizelle et al., accepted, submitted).

The findings do not support anecdotal parental concerns that the use of sign may delay the development of oral language. Whilst they cannot be considered conclusive at this stage, the results instead indicate that there may be value in utilising sign in intervention for children with language difficulties, particularly for generalised production outcomes.

5.5 Research Question 1C: What are the point and interval estimates of effect sizes across outcomes?

Figures 3.20 and 3.21 demonstrate the standardised effect sizes of all language measures probed in this study at outcome and follow-up respectively. The effect sizes reported in chapter three indicate that BEST has substantial effects on language outcomes. There is also a wide level of variability between individual children's progress however, resulting in some outcomes demonstrating medium to large effect sizes without statistical significance across the whole intervention arm such as Targeted BEST sentence content. Results must be interpreted with caution considering this.

Although variable, the effect sizes for most measures found in this study are generally larger than those identified in previous studies (see section 1.3). The variability suggests that there is further work to be done regarding the refinement of the target population of BEST, given that some children made marked progress while others seemingly did not benefit as much. The children with receptive-only language profiles also seemed to begin the study with higher scores, so it is not clear whether the language profile had a differential effects and results are not conclusive from this study. Whilst this project aimed to examine the role of underlying baseline characteristics, poor questionnaire response rates meant that the underlying social, emotional and behavioural difficulties and language profiles could not be analysed further.

It is also important to consider the outcomes of an intervention are clinically meaningful. Clinically meaningful effects the impact of an intervention for children and their families. The increase seen on NRDLS production scores is impressive given that many studies do not demonstrate change on standardised measures, or rely instead on raw scores on measures created specifically for the study. This suggests that the effects of BEST extend beyond the study and have implications for wider language abilities. Future research could seek to establish the 'meaningfulness' of these changes through qualitative interviews with children and their families.

5.6 Research Question 2: What theoretically motivated recommendations can be made from applying Behaviour Change Theory to stakeholder consultation outcomes, to increase acceptability of the research and therefore inform the next stage of BEST development?

The empirical results reported in chapter 3 suggest that further evaluation of BEST is justified. The purpose of early pilot studies is not only to address the question of efficacy, but also to lay the groundwork for future research at the latter stages of the evaluation process (Craig et al., 2008; Eldridge et al., 2016).

In light of this, chapter four describes a qualitative acceptability and feasibility study based on a stakeholder consultation exploring the strengths and weaknesses of the research from the perspective of teachers involved. In line with Braun & Clarke (2006), interview data was coded into themes which were linked to the Theoretical Domains Framework (TDF) and the Capability, Opportunity Motivation-Behaviour (COM-B) model. Functions associated with behaviours from the COM-B model were then identified, allowing data-driven recommendations for future research to be made (Braun & Clarke, 2006; Cane et al., 2012; Michie et al., 2005; Michie et al., 2011). Acceptability of the intervention was also examined using the Theoretical Framework of Acceptability (TFA) (Sekhon et al., 2017).

Further recommendations for a future study were added based on the Author's recorded experiences of conducting the research which were carefully documented in a reflective log. Together the recommendations from consultation with stakeholders and the Author's personal reflections can inform future studies, and indeed are already being implemented in the current Language Intervention in the Early Years (LIVELY) study (McKean et al., 2020) comparing BEST with the Derbyshire Language Scheme (DLS) (Knowles & Masidlover, 1982). The recommendations are restated in Table 5.1 below.

Table 5. 1

Twelve recommendations emerging from Content Analysis of stakeholder consultation and Author experience of conducting the research

1.	The Author and RAs should maintain good working relationships with schools by prioritising flexibility and strong communication, showing awareness of the issues facing teachers and parents and answering questions and addressing concerns adequately. They should be available to answer questions through a variety of methods.
2.	All written and verbal communication with parents/carers, including consent forms and information sheets, should be clear and tailored to their language and literacy needs. Face-to-face meetings and QR codes linked to explanatory videos should accompany written communication to ensure it is easily understood. Information presented to parents should be reassuring regarding their children's identified difficulties.
3.	An electronic method of completing questionnaires should be considered, including the use of QR codes to link parents to an accessible video explaining the study.
4.	Greater emphasis should be placed on communicating the key messages and theoretical rationale of BEST and ensuring that they are understood and remembered (see Figure 5.1 below). Teachers should be presented with posters displaying BEST principles, explicitly highlighting repetition and turn taking as theoretically motivated elements of the therapy.
5.	Sessions should be made more manageable by ensuring sufficient training of the persons delivering BEST and implementing smaller groups e.g. 3-4 children.

6. The advantages and disadvantages of including TAs in the research should be seriously considered in light of the feedback from schools about the disruption it causes, and the Author's reflections that sessions were manageable without including TAs. If TAs are included, they should receive clear explanations of BEST. Ways in which they could benefit further from taking part in the sessions should be identified.
7. Future research should explicitly monitor uptake of the BEST homework, possibly using a diary format and/or a text messaging reminder service. Videos should be created to demonstrate how to complete the homework. Effective systems should be put in place to remind teachers to send booklets home and encourage parents to complete homework with their children.
8. CPD training and materials offered to schools in return for taking part in the study should focus not only on BEST but also on general strategies for improving SLCN in the classroom.
9. Develop a training video for Teaching Assistants to ensure training can be delivered at uniform times and in a standardised manner across each school in the study.
10. Monitor the effect of TAs' presence on therapy outcomes, children's behaviour and TA and researcher perceptions to further determine whether the inclusion of TAs is justified.
11. Build in contingency time for when multiple groups run simultaneously in the same school due to timetabling difficulties.
12. Ensure the protocol considers decisions about what adaptations to therapy delivery will be made a-priori when variations occur, e.g. when supporting staff are not available.

Perhaps most notable when examining the recommendations are the number relating specifically to strong communication and the building and maintaining of effective working relationships (recommendations 1, 2, 3, 4 and 7). The importance of social capital in relationship building with schools and recruitment of participants is rarely discussed in published studies, however these recommendations suggest that relationships with school staff and communication with families are fundamental to the success of a project. Not only should these recommendations be heeded in a future study involving BEST, but also in schools-based research in general, to maximise participation, engagement and retention of participants and schools.

Another common theme emerging from the above recommendations is ensuring that communication and materials are accessible to parents and carers. Although every effort was made during the study to create accessible messages to families regarding them providing consent and also around completing the BEST homework, the consultation data suggests that this could still be improved (recommendations 2, 3, 7).

In summary, 12 recommendations were produced using a qualitative methodology in chapter 4. The implications of these findings are discussed in section 5.8 after the consideration of the strengths and limitations of this research.

5.7 Strengths and Limitations of the Research

Before drawing conclusions about the implications of this research it is important to consider the relative strengths and weaknesses of the study. This is necessary to weigh the findings appropriately and consider their strength, rigour and generalisability. The following section therefore discusses the strengths and limitations of this work.

5.7.1 Strengths

5.7.1.1 Study Design

The design of the research reported in this thesis is one of its key strengths. Based on a previous randomised controlled pilot study, the pragmatic decision was taken to employ a non-randomised design, and this was shown to increase acceptability to parents and schools. Non-randomised designs are valuable in situations where randomisation is considered unethical or impractical (Craig et al., 2008), as was the case with this study. Matching at the school level also left the groups less susceptible to bias due to the sample size.

As well as being appropriate for the sample size and stage of the evaluation process, the non-randomised design also delivers the equivalent 'moderate' level of evidence, that is assigned to most other published studies, in line with Law et al's (2012) classification. With the addition of this thesis to the evidence base, BEST is now equally well evaluated as many alternative interventions for the target population and arguably more transparently reported than many.

The present work benefits from a solid grounding in the frameworks designed specifically for SLT interventions (Fey & Finestack, 2009; Robey & Schultz, 1998), supplemented with the gold-standard MRC guidelines (Craig et al., 2008). The situation of the research within the piloting and feasibility phase of the evaluation process ensured that the appropriate aims of work at this level of development were met. The study focused on testing procedures, estimating recruitment and retention rates and determining efficacy, as well as evaluating the acceptability and feasibility of the intervention through consultation with school staff. These vital elements of the evaluation process, along with the recommendations generated in chapter four, feed into a current larger-scale trial (McKean et al., 2020), reducing the risk of conducting research that wastes time and money (Craig et al., 2008).

To ensure that future research can easily extract the relevant information, this thesis followed gold standards of reporting in line with CONSORT guidelines (Eldridge et al., 2016) and any devitions from the guidelines are noted and justified.

A further methodological strength was the collaboration with NHS SLT gatekeepers. This undoubtably had advantages for school and participant recruitment in line with previous recommendations (Sturgis et al., 2006). At the school level, SLTs had existing professional relationships with Senior Leadership Teams and SENCOs which they utilised to make direct contact. This approach increased engagement with, and recruitment of, schools compared to a ‘cold-calling’ method.

Child recruitment relied heavily upon the support of teachers, since they selected the children they considered to be eligible and approached families. Trust and engagement established via the SLTs may have had a positive impact on teacher’s attitudes towards the study and increased recruitment as a result.

5.7.1.2 Methodological Rigour

As detailed throughout chapter two, the Author took numerous steps to ensure that the research was conducted as robustly as possible to maximise the confidence we can have when interpreting the results.

The Author delivered extensive training to RAs regarding conducting robust and reliable standardised assessments. Measures were piloted in nurseries before the study began and video recordings of role-play sessions were scrutinised to increase fidelity of both the assessments and the BEST sessions themselves. Reliability of delivery of language measures was consistently high throughout the study.

Inter-rater reliability checks and the utilisation of a scoring manual designed by the Author and updated with ambiguous responses ensured that measures are reliable and that we can have confidence when interpreting these findings.

Treatment fidelity was a priority during the study. A number of checks were built into the protocol to ensure that BEST was delivered in line with the manual and are reported transparently in section 2.9. The Author ensured adherence to the BEST manual and when necessary recorded appropriate changes to intervention delivery to ensure treatment fidelity remained high. The Author also addressed any issues of treatment fidelity with the TAs supporting sessions. Treatment fidelity represents an important strength of the present study, since outliers are not driven by children receiving differing dose form, increasing the confidence with which findings can be interpreted.

The nature of this efficacy study was such that the Author was able to quickly identify and correct problems with intervention implementation. This highlights the strength of the current design, however subsequent effectiveness trials and implementation in schools will not have this level of control. It is important that alternative ways are found to increase treatment fidelity for subsequent work where the Author may not be on hand to make adjustments. These suggestions including displaying posters promoting the key principles for BEST delivery in classrooms and are discussed in more detail below.

At the piloting and feasibility stage of intervention evaluation, it is particularly important to gain qualitative insights about barriers to participation so that methodological barriers can be addressed early (Craig et al., 2008). The empirical results of this study were triangulated by employing Behaviour Change Theory (BCT) to produce recommendations for future research that were driven by data collected from stakeholders. Few intervention studies take such a rigorous approach to making future recommendations and may therefore miss important themes emerging from consultation with stakeholders and participants. Through the unique perspective this method afforded insight was gained into the increased acceptability to teachers and families than the initial randomised controlled pilot study. This study also echoes findings that teachers and SENCOs are keen to implement interventions for language difficulties in the early years (White & Spencer, 2018).

5.7.1.3 Advanced Analytical Methods Suitable for Quasi-Experimental Design

This research employs longitudinal multi-level modelling to effectively cluster data at the level of the school and the individual participant. Multi-level modelling posed a significant advantage over methods traditionally employed in the field (e.g. Standard Ordinary Least Squares Regression (OLS) (Riccio & Jemison, 1998)) in this study, since it was able to account for the greater similarities between some participants (e.g. those that attend the same school) than others (e.g. those attending different schools). The multi-level model also controlled for effects which are not controlled for in other comparable methods such as regression, ANOVA and ANCOVA (Finch et al., 2019). The models employed in this study are therefore more robust than other similar methods and do not suffer from violations of the assumption that error terms for different children are un-correlated.

Multilevel modelling is an appropriate statistical method for addressing the research questions examined in this thesis for a number of reasons. In non-randomised study designs, bias occurs due to the lack of randomisation (Angrist & Pischke, 2008, 2014), which poses a threat to internal validity (Murray, 1998). This is compounded by the fact that children who attend the same school will be inherently more similar to each other than to children who attend a different school due to factors such as teaching style, socio-economic status and school ethos. The sum of these factors increases the likelihood of similarities between an individual and other members of their group, compared to individuals outside of the group and is known as intra-cluster correlation (Bray et al., 2009).

Additionally, the method accounted for the clustering within schools, (whereby children attending the same school were more similar to each other than to those from different schools due to their shared environmental characteristics) and for each timepoint which were clustered within each child. Multi-level models are specifically designed to deal with data which violate the assumption of independence within pooled models, allowing the clustering of error terms to reflect the similarities between and within individuals across schools and time.

Whilst the completion rate was very high for face-to-face assessments, some data were missing due to child absence (see table 3.1). Children who missed one or more assessment were able to be included in the analysis, maintaining the sample size which is not possible using traditional analyses such as Analysis of Variance and Analysis of Covariance (Finch et al., 2019).

5.7.2 Limitations

5.7.2.1 Minor Deviations from the Study Protocol

Several small deviations from the protocol occurred which should be acknowledged and may be useful for informing future trial methodology.

During the process of matching schools, an error resulted in the matching process not taking the differences in class size and the number of nursery and/or reception classes in each school into account. Despite this, the three treatment arms were found to be equally matched on the majority of criteria and it is unlikely that this would have had affected the results.

Unblinding occurred on one occasion during the study, when a teacher forwarded an email containing the school intervention arm to an RA. It was not felt that this compromised the results of the assessments, although the findings should be interpreted in light of this. This also highlighted a potential discussion point: Many intervention studies refer to blinding but do not detail the steps taken to ensure its integrity, especially when researchers are working in close proximity to one another. As discussed in chapter 2, blinding is a complex issue that requires planning and diligence to maintain and which may not always be transparently reported.

5.7.2.2 Sources of Possible Variability across Schools

The Author noticed that when two adults delivered BEST with sign, it was difficult for the child to attend to the toys and the signing since the they were delivered simultaneously. This observation was made by the Author while conducting the research and was not tested empirically. Children's attention may have affected progress differentially across the intervention arms, however this issue has more implications for signing. Across both BEST arms, children who did not watch the adults were still able to access the spoken input of the BEST sentences. Signing, on the other hand, is solely visual and would not have been accessible to children who were not attending to the adult providing the signs at the appropriate time. It is therefore possible that children in the signing arm would have been more disadvantaged by behavioural or concentration difficulties than children in the BEST without Sign arm. Indeed, the Author noticed a trend whereby children frequently attended to the toys (which were manipulated by adult 2 in the cases where a school staff member was involved) and therefore did not appear to attend consistently to the signs provided by the Author. Similar difficulties were reported by Vogt & Kauschke (2017).

As discussed in the 'Important Changes to the Protocol' section (see section 2.15), the method of delivery was altered slightly when the second adult was not available to support the session. These changes were approved by the Authors of BEST since they resulted in no difference to the exposure of model and target sentences. They were therefore unlikely to have affected overall outcomes. Nevertheless, it is important that future research plans any such changes a-priori rather than post-hoc. Although this is a methodological limitation of the present research, it could also indicate that the effects of signing may have been stronger had children consistently attended to the signing, which should be addressed in a subsequent study.

Another potential source of variability in the present study was the homework. Each BEST session has a corresponding homework booklet to be completed with parents or carers at home. In the present study, homework booklets were distributed following every session, although the study did not monitor uptake. Children received approximately 30 to 40 minutes of intervention per week. If children also completed the homework for 10 to 15 minutes per day as directed, their overall exposure to the BEST content could have increased by up to two thirds. It is therefore likely that some children received substantially more input than others as a result of completing the homework frequently with caregivers at home.

There were no systematic differences between the groups in terms of expected homework uptake, so this is unlikely to have biased the treatment arms. However, monitoring of this in future research may improve our understanding of the relative value of the homework sessions, and the quantity of exposure, for BEST as an intervention.

As is discussed in chapter 3, significant differences were observed between the CSCOT total proportion scores across the three treatment arms at baseline. Tukey's HSD post-hoc testing revealed the BEST without sign arm had significantly higher scores on the CSCOT than Treatment as Usual. It is not clear whether this represents a genuine difference between the TAU and BEST without sign arms. There are 3 possible reasons for this:

1) The TAU schools randomly had poorer average oral language classroom environments, and thus lower CSCOT scores, than the other treatment arms.
2) The CSCOT is unreliable; the lower scores for TAU and for the third data wave may not reflect a true difference in oral language classroom environments across the schools.
3) The Author became systematically more stringent when completing the CSCOT at each wave having become more experienced with the tool. This theory is supported by the data since all but 1 of the TAU schools were in the final wave. Whilst the researcher undertook training in

the completion of the CSCOT prior to the study commencing, it is possible that over time and through completing cumulatively more observations, the researcher changed the way in which the CSCOT was completed in some systematic and unintentional way, thus affecting intra-rater reliability.

Since the issue of intra-rater reliability was not addressed in the development of the CSCOT (Dockrell et al., 2012; Dockrell et al., 2015), the Author co-supervised a student research project testing the reliability of the CSCOT scores to examine this (Barlow & Stringer, 2019). Intra-rater reliability (reliability over time completed by the same researcher) was tested for the CSCOT scores over a period of four months by the same researcher. The Author completed three CSCOT observations for classrooms at time 1, and then again four months later at time 2 under the same conditions. The total proportion scores suggested a high level of intra-rater reliability across all three dimensions of the CSCOT tool. This finding suggests that change in average CSCOT scores over the three time points is less likely to be due to poor intra-rater reliability over time, and more likely to reflect either a genuine decrease in CSCOT scores, or else an issue with validity. This should be addressed in future research evaluating the tool.

It must also be noted that the CSCOT was designed as a tool for teachers and school staff to evaluate and improve the classroom oral language environment. It was not designed as a research tool for comparing different classrooms in the original studies (Dockrell et al., 2012; 2015) and this was a novel application of the tool in this study. It would be useful if future research were to address the potential of the CSCOT for this purpose. This suggests that the CSCOT shouldn’t be given too much weight at the present time, and that the lack of balance across the treatment arms at baseline does not necessarily represent a problem.

Whilst it is not possible to draw conclusions on the nature of the mismatch across treatment arms, the strength of the overall matching process can be substantiated with other measures. The IDACI scores are not statistically different from each other across the treatment arms or data waves, which is encouraging. Since the IDACI (derived from the Index of Multiple Deprivation scores) are an accepted measure of socio-economic status, it can still be tentatively argued that the treatment arms are balanced. As is discussed below however, there are also caveats associated with using IDACI scores based upon school postcodes rather than individual participant postcodes.

5.7.2.3 Engagement with Parents

Although every effort was made to produce simple information sheets and consent form for parents, some teachers during the stakeholder consultation interviews reported that parents had difficulty accessing the information. This may have affected recruitment if some parents were unable to access the forms. Because of the amount of information necessary to convey, there is not an easy solution to this issue. The Author recommended that future studies use QR codes on the information materials to link to videos explaining studies in simple language to help reach parents with low literacy or English language levels.

Poor questionnaire response rate from both teachers and parents also affected the implementation of planned analyses of social, emotional and behavioural difficulties (SDQ questionnaire) and functional communication (FOCUS questionnaire). It is unclear whether this is likely to be due to difficulty accessing the materials or for another reason. Because of this, it was not possible to measure the effects of SEBD on the response to BEST, nor to measure any potential benefits for functional communication.

It is important that future work examines the trajectory of children's functional communication since this is an important diagnostic area of DLD (Bishop et al., 2016) and interventions for DLD should therefore seek to examine this.

5.7.3 Summary

This section discussed strengths and weaknesses of the research, as well as caveats that should be taken into account when interpreting the findings. Overall this study provides robust evidence about the strengths of BEST as well as using the study limitations to develop recommendations for future research. We now move on to examine the theoretical and clinical implications of this thesis.

5.8 Implications of the Research

5.8.1 Theoretical Implications

This research has implications for the theoretical basis of language intervention and theories of language disorder more generally. The hypothesised mechanisms are discussed under the relevant research questions in this chapter. Overall the findings lend support to the hypothesised mechanisms of BEST presented in 1.5.1.4. and 1.6.5, and to the wider usage-based theory due to the progress made by children in the study, compared to those who did not receive a usage-based intervention.

Although the present study did not seek to directly test the constructivist theory, it did measure the efficacy of a usage-based language intervention on children's language outcomes and it is possible to comment on the theoretical claims for this reason. Since BEST is shown to improve outcomes for children with language difficulties, it can be argued that intervention underpinned by usage-based theory can be effective for targeting children with such difficulties. In particular, there is evidence of improved ability with analogy, suggesting that children who received the intervention harnessed ways to produce more abstract phrases from the examples that they were exposed to.

A further theme relating to usage-based theory that emerged from the literature review in chapter one is the similarities between the emergentist and usage-based accounts. These theories appear to take similar approaches but rarely acknowledge the presence of one another. This may be due to the differing fields of Speech and Language Therapy and Psychology from which they were developed, however it would be beneficial for both for research to consider their respective similarities and whether they are able to enhance one another in future work. For example, emergentist accounts provide explicit explanations of language disorder unlike Tomasello's current usage-based theory which considers mainly typical development.

The evidence from this study suggests that signing increases children's awareness of sentence morphological frame, which in turn supports analogy and therefore increases production skills. It is likely that the increase in analogy facilitates the use of abstract representations, thereby increasing the child's processing capacity and accelerating progress. Although the basis for these claims originate from the phonology literature, they can be applied to grammaticality as well (Bishop, 2014).

Whilst this study cannot provide conclusive evidence about the accuracy of the hypothesised theoretical mechanisms underpinning BEST, it does lend support to the cognitive mechanisms hypothesised to underpin each phase of language acquisition proposed by Tomasello (2003).

In light of the discussion above, there are two key aims that subsequent research should address with regards to establishing the role of theoretical underpinnings. Firstly studies should address whether intervention underpinned by usage-based theory is more effective than similar interventions underpinned by other key theories of language acquisition, such as generativist theories or the specific accounts of language disorders discussed in emergentist theories (Chiat, 2001; Evans, 2001) and interventions that are not theoretically motivated. This will determine whether the mechanisms hypothesised by these accounts are accurate when applied to a clinical context.

Secondly, studies must seek to identify the active ingredients in interventions that are found to be effective. This may involve many rounds of small-scale trials, or many study arms of larger scale trials to assess the differences between slightly varied versions of the same intervention (such as in the present study with the testing of BEST with and without sign). This is a particular difficulty in complex intervention research, since there are many potential active ingredients that could be affecting outcomes and they could also interact with one another (Craig et al., 2008; Eldridge et al., 2016). Trials evaluating active ingredients also require greater power if they are to effectively measure the differences between intervention arms. Despite the obvious difficulties however, this work is essential for developing interventions that are theoretically motivated, and for furthering theoretical understanding of language difficulties.

A further theoretical implication, which also has consequences for implementation, is the lack of praise delivered during BEST sessions. In the BEST manual, McKean et al. argue that, in line with the constructivist account of language acquisition, praise and other verbal input should be avoided during sessions to avoid disrupting the mapping of verbal input onto the visual stimuli. Whilst this was not perceived to cause issues during the study, it is possible that the lack of praise may have a detrimental effect on children's attention, particularly for those with SEBD. It is also possible that children's positive behaviours are not reinforced due to the lack of praise. Conversely, research has shown that inconsistent praise can negatively affect children's motivation (Zentall & Morris, 2010), so the opposite may also be true. Research examining the lack of praise in BEST would be required to establish its role and whether it in fact promotes or hinders learning through the BEST input.

Risk factors for language difficulties, discussed in 1.2.4, present another theoretical issue with relation to the underlying usage-based theory. BEST was in part designed to treat children from low socio-economic status backgrounds and usage-based theory allows this to be addressed by increasing the quality and quantity of input the child is exposed to- since this is hypothesised to be lower than for children from higher SES backgrounds. Further, the quality input may aid children who are at risk due to cognitive factors by supporting processing and bootstrapping. Usage-based theory does not specifically speak to genetic risk factors, although it must be acknowledged that environmental and cognitive factors may interact with these. It is therefore reasonable to argue that children who are at risk of language difficulties due to environmental and cognitive factors may particularly benefit from receiving BEST. On the other hand, further research would be required to establish whether children at risk due to genetic factors would benefit in the same way and this may have clinical as well as theoretical implications.

5.8.2 Clinical Implications

A number of important clinical implications for service delivery arise from this thesis.

At the piloting and feasibility stage of the research process, BEST is a promising intervention for young children with severe language difficulties. It is one of the few comparable interventions to have undergone rigorous piloting and feasibility work and it will continue to be evaluated at the next stage of the process in a fully powered trial (McKean et al., 2020). The evidence at this stage suggests that BEST can be recommended for clinical use, although further refinement to the target population is required when more robust data can be gathered regarding SEBD and language profile. BEST will also be added to the What Works database with clear guidance for practitioners as to the level of evidence currently available and listing the appropriate caveats and principles that should be adhered to (see Figure 5.1).

With regards to signing, the evidence at this stage suggests that BEST with additional signing shows benefits above and beyond those of BEST. These findings are preliminary and must be interpreted with caution, particularly as the treatment arms are underpowered for this analysis. Although further work is still required, at this point it seems reasonable to conclude that it is worthwhile for school staff and therapists delivering BEST to become competent delivering the signs alongside the scripted intervention, in spite of the additional time and effort.

5.8.3 Implications for Implementation

Results from this stage in the evaluation process suggest BEST is a viable intervention for children with language difficulties in the early years and future evaluation will continue to examine this (McKean et al., 2020). In light of the current difficulties surrounding funding and availability of SLT services (Bercow, 2008, 2018; McGregor, 2020), schools play a vital role in delivering additional support to children with suspected or diagnosed difficulties. BEST is an appropriate intervention for wide-scale delivery within schools, provided that treatment fidelity can be successfully implemented and maintained to a high standard.

There are several features of BEST that make it particularly suitable for delivery in schools. Firstly, BEST is designed to be delivered in groups of up to six children. Whilst the number included in any group may vary based on the individual children in the group, for example if a child has emotional or behavioural difficulties a reduced group size may be beneficial in that circumstance, targeting multiple children at once is a cost-effective and time-efficient method of delivering intervention. Despite individual variation, group therapy seems to be as effective as 1:1 intervention (Eidsvåg et al., 2019) and would reduce overall time staff spend delivering intervention.

Missed appointments waste NHS time and financial resources (Blankenstein, 2003). Children are arguably more likely to attend therapy sessions that are delivered in school and which therefore do not require parents/carers to take them to a clinic during the working day or at other times. This may result in fewer missed sessions and children receiving the appropriate dosage, as well as conserving public healthcare resources.

A further advantage of BEST is that children do not have to begin with the first session and end with the last, since sessions do not follow a linear order. BEST groups could therefore run on a frequent basis and children could begin sessions whenever they are identified as needing additional support. Given the long waiting times for some NHS SLT services following initial referral (Bercow, 2008, 2018), this is a particular advantage of BEST in that children could receive intervention while waiting to be seen by a specialist. Parents may also feel reassured that their children are able to receive additional targeted support for their children immediately following identification, rather than potentially facing a long wait for specialist support.

In spite of the advantages outlined above, BEST delivery in schools would require careful training and monitoring to avoid issues with implementation. Training would need to place emphasis on adherence to the manual, specifically on the areas that school staff most frequently deviated from in the present study, which are hypothesised to underpin outcomes. These principles are: 1) not delivering verbal praise or other language alongside the scripted input; 2) children not having to speak until they want to and 3) delivering the full intervention in line with the manual, even if children do not show progress immediately. Conveying these messages would be essential for school staff to implement BEST successfully.

One way to identify ways to promote treatment fidelity in future BEST implementation in schools is to look to specific issues that were encountered during the present study. As discussed in chapter four, there was variation between school staff in their role as ‘Adult 2’, with some staff not following the BEST principles closely. This would have a greater impact if school staff were delivering BEST with no researcher input. Recommendations 9 and 10 relating to the development of video resources for TAs delivering BEST and monitoring the effect of TA assistance in subsequent studies would go some way to improving fidelity, but this issue could also be addressed with careful messaging during and after initial BEST training is delivered. Based on the comments from staff and difficulties in delivery observed by the Author, Figure 5.1 outlines the key messages that need to be understood so that BEST is delivered in accordance with the manual. These key messages could be produced on posters and for staff to display in a prominent location where BEST is delivered to serve as a continued reminder. The use of videos alongside face-to-face training would also constitute a time-and-cost-effective way to deliver training and has the advantage that videos can be re-watched to remind staff of the key principles.

Figure 5. 1

Key messaging to be delivered to staff to ensure treatment fidelity during school-led BEST intervention

5.9 Future Research Directions

This thesis has revealed several areas for further investigation. Three major issues were beyond the scope of the present study due to poor rates of questionnaire responses. As well as implementing the recommendations set out in Table 5.1 above, these specific areas should be considered in future evaluations of BEST: 1) social, emotional and behavioural difficulties; 2) functional communication; and 3) language profile.

5.9.1 Social, emotional and behavioural difficulties

It is well established that there is an association between language difficulties and Social, Emotional and Behavioural Difficulties (SEBD), although the mechanism linking the two is not fully understood

(Beitchman & Brownlie, 2005; Benner et al., 2002; Botting & Conti-Ramsden, 2000; Clegg et al., 2015; Law, Plunkett, et al., 2012; Lindsay et al., 2012; Rice et al., 1991; Rose et al., 2018; Willinger et al., 2003). Although there is evidence of positive behavioural outcomes following language intervention (Law, Plunkett, et al., 2012), it is not clear how children with behavioural difficulties would respond to BEST.

BEST places relatively low demands on children with a focus on manipulating the input. The intervention may be more accessible to children with SEBD than interventions that place greater demands on the child. Conversely, the use of praise and other verbal input should be avoided during the sessions, which may have implications for maintaining a child's attention, particularly if they have SEBD. This issue was not apparent in the research, however. During the student randomised controlled pilot study (see chapter two), only one child was unable to complete the full intervention due to non-compliance, while during the present study no children were unable to complete BEST due to SEBD.

Initially this study sought to measure SEBD using the Strengths and Difficulties Questionnaire (SDQ) to categorise children as having high or low levels of SEBD. This was not possible due to poor questionnaire return rates from parents and teachers. Future studies following the recommendations in Table 5.1 may have greater success collecting this information. This information could be used to explore the differential progress of children with differing SEBD profiles to establish whether BEST can be recommended as a suitable intervention for children with particular SEBD. This would also further refine the target population of BEST, having important clinical implications for selecting the children most likely to benefit.

5.9.2 Functional Communication

Functional communication is the degree of success a communicator is able to achieve through language and other aspects of communication and is now considered part of the DLD diagnostic criteria (Bishop et al., 2017). Research examining functional communication is a key priority for children's entire wellbeing, since parents and children value communication as the basis for achieving other life goals including social and academic success (Dockrell et al., 2014).

This study initially sought to determine the effects of BEST on children's functional communication using the FOCUS questionnaire. A reliable measure of functional communication could not be obtained due to the poor response rate and the study was not able to answer further questions about functional communication. Future research should seek to examine the effect of BEST on functional communication, to establish whether this is a viable secondary outcome and whether the language outcomes observed in the present study influence functional communication.

5.9.3 Language Profiles

Children with receptive or mixed language profiles are at risk of poorer long term outcomes than children with expressive-only difficulties (Law et al., 2004; Nickisch & von Kries, 2009). Children with expressive-only language difficulties tend to make greater progress in intervention than those with mixed expressive and receptive language difficulties (Boyle et al., 2010). As discussed throughout this thesis, BEST aims to improve both receptive and expressive abilities by promoting the development, use and understanding of abstract sentence structures. BEST aims to target children with expressive or mixed difficulties. Whist there is clear evidence that BEST supports expressive language, it is not clear whether this is likely to extend to receptive difficulties.

This study aimed to explore the progress of children with differing language profiles receiving BEST using post-hoc analyses to further refine the intervention target population, and gain insight into the underlying mechanisms supported by BEST. In line with expected numbers, very few children had a receptive only language profile. These children also had higher than average language abilities at the beginning of the study based on NRDLS comprehension and production scores. It was therefore considered inappropriate to conduct the planned post-hoc analysis as a fair comparison could not be drawn. This should be a priority for a future fully powered trial to further refine the target population of BEST and determine whether children with receptive-only difficulties can also benefit from an intervention of this nature.

At this point the advantages and disadvantages of language profiling based on standardised tests must also be mentioned. Children were defined as having an expressive, receptive or mixed language profile based on reaching a specific cut-point (the 16th centile) on one or both of the subscales of the NRDLS. Whilst this approach has been taken in other studies (e.g. Ryan et al., 2016), the subscales provide only a snapshot of the child's performance on the day the test was administered and may be affected by external factors. This limitation does not affect the main analysis which did not account for language profile, and the sensitivity analysis which it could impact shows only minor deviations from the main analysis results. The issue of using cut-points on standardised tests for inclusion and exclusion criteria is also contentious both clinically and for research purposes; strict cut-points in standardised tests are liable to measurement error close to thresholds, posing a risk of mis-diagnosis or mis-categorisation (Le et al., 2020; McKean et al., 2017; Burchinal, 2008) and results must be interpreted in light of this.

5.9.4 Dosage

During the present study information on dosage was recorded for all children receiving BEST. Analyses were carried out to determine the differing effects of receiving 'high' (>=75%) or 'low' (<75%) dosage.

It is currently unclear what is the optimal treatment dosage required to achieve change in a child's language abilities. Evidence is sparse, and the little which does exist has mixed outcomes (Bellon-Harn et al., 2014; Justice et al., 2017; Smith-Lock et al., 2013). Due to this vacuum in evidence, Speech and Language Therapists in the UK tend to design treatment duration to fit with school term dates, most often delivering 6 weekly sessions of intervention. This is despite evidence from Law et al. (2004), that interventions delivered over eight or more sessions are associated with better clinical outcomes. More recent evidence about the ideal dose form should also be incorporated into policy and decision making around provision and funding. One notable point here is that it is extremely difficult to balance the optimal dosage and treatment fidelity, partly due to the lack of SLT time allocated to delivering intervention. The mechanisms underpinning BEST suggest that a high level of fidelity is necessary when BEST is being delivered, however whether this can realistically be achieved in the current service delivery model is contentious.

Feedback from SLTs about BEST suggests that the model of 16 sessions of intervention therefore does not fit their usual practice. Furthermore, during early development work it appears that children make faster progress learning the sentence structures in the target sentences, and longer to learn the morphology (see sentence structures and morphology section above). There is a temptation therefore for practitioners to not complete the necessary number of sessions for the children to achieve mastery of the morphological targets. Given the crucial role grammatical morphology plays in language and literacy development as children move into school (Muter et al., 2004), this could significantly impact on the effectiveness of the intervention.

By tracking children's progress in each intervention session with respect to sentence structure and morphology and completing short assessments after 4, 8, 12, and 16 sessions, future work should seek to further refine optimal dosage and delivery.

5.10 Author's Personal reflections

Through the process of conducting this research I have been extremely privileged to work alongside so many wonderful people and gain a wealth of skills and knowledge about language interventions and their evaluation. Beginning this research, I had little knowledge of many issues and debates within Speech and Language Sciences and specifically pertaining to successful intervention delivery. Through the excellent training provided to me, I began to understand the key issues for consideration when designing and implementing an intervention trial. I furthered my understanding of reliability and treatment fidelity and overcame practical barriers such as maintaining blinding when sharing a workspace with research colleagues and scheduling intervention sessions with schools. As I point out at the end of this chapter, one of my biggest take home messages has been the importance of utilising social capital and building and maintaining strong working relationships with stakeholders, schools and families. I have also come to realise that a more joined up approach to collaboration between universities and schools which nurture these relationships and reduce time spent recruiting is necessary for strong intervention studies to be produced moving forwards. It is my hope that the research community finds new ways to build and maintain relationships with schools and stakeholders, so that research can benefit from diverse collaboration and schools can see the value in taking part in research.

References

Abbott, C., & Lucey, H. (2005). Symbol communication in special schools in England: The current position and some key issues. *British Journal of Special Education*, *32*(4), 196–201. https://doi.org/10.1111/j.1467-8578.2005.00397.x

Adams, C., Lockton, E., Freed, J., Gaile, J., Earl, G., McBean, K., Nash, M., Green, J., Vail, A., & Law, J. (2012). The Social Communication Intervention Project: A randomized controlled trial of the effectiveness of speech and language therapy for school-age children who have pragmatic and social communication problems with or without autism spectrum disorder. *International Journal of Language and Communication Disorders*, *47*(3), 233–244. https://doi.org/10.1111/j.1460-6984.2011.00146.x

Adams, C., Lockton, E., Gaile, J., Earl, G., & Freed, J. (2012). Implementation of a manualized communication intervention for school-aged children with pragmatic and social communication needs in a randomized controlled trial: The Social Communication Intervention Project. *International Journal of Language and Communication Disorders*, *47*(3), 245–256. https://doi.org/10.1111/j.1460-6984.2012.00147.x

Akhtar, N., & Gernsbacher, M. A. (2007). Joint Attention and Vocabulary Development: A Critical Look. *Lang Linguist Compass*, *1*(3), 195–207. https://doi.org/10.1111/j.1749-818X.2007.00014.x.Joint

Albarracín, D., Gillette, J. C., Earl, A. N., Glasman, L. R., Durantini, M. R., & Ho, M. H. (2005). A test of major assumptions about behavior change: A comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. *Psychological Bulletin*, *131*(6), 856–897. https://doi.org/10.1037/0033-2909.131.6.856

Aldred, C., Green, J., & Adams, C. (2004). A new social communication intervention for children with autism: Pilot randomised controlled treatment study suggesting effectiveness. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/j.1469-

7610.2004.00338.x

Ambridge, B., & Lieven, E. V. (2011). Child language acquisition: contrasting theoretical approaches. In *Child Language Acquisition: {Contrasting} Theoretical Approaches*. https://doi.org/10.1017/CBO9780511975073

Ambridge, B., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2006). The distributed learning effect for children's acquisition of an abstract syntactic construction. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2005.09.003

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders. In *Arlington*. https://doi.org/10.1176/appi.books.9780890425596.744053

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion.* New Jersey: Princeton University Press.

Angrist, J. D., & Pischke, J.-S. (2014). *Mastering'metrics: The path from cause to effect*. Princeton University Press.

Archibald, L. M. D., & Gathercole, S. E. (2006). Visuospatial Immediate Memory in Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *49*(2), 265–277. https://doi.org/10.1044/1092-4388(2006/022)

Atencio, C. A., & Schreiner, C. E. (2008). Spectrotemporal processing differences between auditory cortical fast-spiking and regular-spiking neurons. *Journal of Neuroscience*, *28*(15), 3897–3910. https://doi.org/10.1523/JNEUROSCI.5366-07.2008

Barlow, L., & Stringer, H. (2019). *Investigating the Reliability and Acceptability of the CSCOT in the Early Years*.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3). https://doi.org/10.1016/j.jml.2012.11.001

Bartlett, R., Wright, T., Olarinde, T., Holmes, T., Beamon, E. R., & Wallace, D. (2017). Schools as Sites for Recruiting Participants and Implementing Research. *Journal of Community Health Nursing*, *34*(2), 80–88. https://doi.org/10.1080/07370016.2017.1304146

Bayat, M., Mindes, G., & Covitt, S. (2010). What does RTI (response to intervention) look like in preschool? *Early Childhood Education Journal*, *37*(6), 493–500. https://doi.org/10.1007/s10643-010-0372-6

Beitchman, J., & Brownlie, E. (2005). Language development and its impact on children's psychosocial and emotional development. *Encyclopedia on Early Childhood Development*, 1–7.

Beitchman, J. H., Jiang, H., Koyama, E., Johnson, C. J., Escobar, M., Atkinson, L., Brownlie, E. B., & Vida, R. (2008). Models and determinants of vocabulary growth from kindergarten to adulthood. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *49*(6), 626–634. https://doi.org/10.1111/j.1469-7610.2008.01878.x

Beitchman, J. H., Wilson, B., Johnson, C. J., Atkinson, L., Young, A., Adlaf, E., Escobar, M., & Douglas, L. (2001). Fourteen-year follow-up of speech/language-impaired and control children: Psychiatric outcome. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(1), 75–82.

Bell, N. (1991). Gestalt imagery: A critical factor in language comprehension. *Annals of Dyslexia*, *41*, 246–260. https://doi.org/10.1007/BF02648089

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

Bellon-Harn, M. L., Byers, B. A., & Lappi, J. (2014). Treatment Intensity: Effects of Interactive Book Reading on Narrative Abilities in Preschool Children With SLI. *Communication Disorders Quarterly*. https://doi.org/10.1177/1525740114524051

Bellon-Harn, M. L. (2012). Dose frequency: Comparison of language outcomes in preschool children with language impairment. *Child Language Teaching and Therapy*, *28*(2), 225–240. https://doi.org/10.1177/0265659012445936

Benner, G. J., Nelson, J. R., & Epstein, M. H. (2002). Language skills of children with EBD: A literature review. *Journal of Emotional and Behavioral Disorders*, *10*(1), 43–56.

Bercow, J. (2008). The Bercow Report. *Communication*.

Bercow, J. (2018). *Bercow: Ten Years On An independent review of provision for children and young people with speech, language and communication needs in England*. www.bercow10yearson.com

Berry, D. C., Neal, M., Hall, E. G., McMurray, R. G., Schwartz, T. A., Skelly, A. H., & Smith-Miller, C. (2013). Recruitment and Retention Strategies for a Community-Based Weight Management Study for Multi-Ethnic Elementary School Children and Their Parents. *Public Health Nursing*, *30*(1), 80–86. https://doi.org/10.1111/phn.12003

Best, W. (2005). Investigation of a new intervention for children with word-finding problems. *International Journal of Language and Communication Disorders*. https://doi.org/10.1080/13682820410001734154

Bickford-Smith, A., Wijayatilake, L., & Woods, G. (2005). Evaluating the Effectiveness of an Early Years Language Intervention. *Educational Psychology in Practice*. https://doi.org/10.1080/02667360500205859

Bishop, D. V. M., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment: neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/jslhr.4206.1295

Bishop, D. V. M. (1994). Grammatical errors in specific language impairment: Competence or performance limitations? *Applied Psycholinguistics*. https://doi.org/10.1017/S0142716400006895

Bishop, D. V. M., & McArthur, G. M. (2005). Individual differences in auditory processing in specific language impairment: A follow-up study using event-related potentials and behavioural thresholds. *Cortex*, *41*(3), 327–341. https://doi.org/10.1016/S0010-9452(08)70270-3

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2016). *CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children*. https://doi.org/10.1371/journal.pone.0158753

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Catalise-, T. (2017). *Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology*. https://doi.org/10.1111/jcpp.12721

Bishop, D. V. M., & Mogford-Bevan, K. (1993). *Language development in exceptional circumstances*. Lawrence Erlbaum Associates.

Bishop, D. V. M. (2014). Uncommon understanding: Development and disorders of language comprehension in children, classic edition. In *Uncommon Understanding: Development and Disorders of Language Comprehension in Children, Classic Edition*. Taylor and Francis. https://doi.org/10.4324/9780203381472

Bishop, D. V. M. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language and Communication Disorders*, *52*(6), 671–680. https://doi.org/10.1111/1460-6984.12335

Bishop, D. V. M. (2002). The role of genes in the etiology of specific language impairment. *Journal of Communication Disorders*, *35*(4), 311–328. https://doi.org/10.1016/S0021-9924(02)00087-4

Bishop, D. V. M. (2006). EPS mid-career award 2005: Developmental cognitive genetics: How psychology can inform genetics and vice versa. *Quarterly Journal of Experimental Psychology*, *59*(7), 1153–1168. https://doi.org/10.1080/17470210500489372

Blankenstein, R. (2003). Failed appointments – do telephone reminders always work? *Clinical Governance: An International Journal*, *8*(3), 208–212.

https://doi.org/10.1108/14777270310487011

Bolderson, S., Dosanjh, C., Milligan, C., Pring, T., & Chiat, S. (2011). Colourful semantics: A clinical investigation. *Child Language Teaching and Therapy*. https://doi.org/10.1177/0265659011412248

Botting, N., & Conti-Ramsden, G. (2000). Social and behavioural difficulties in children with language impairment. *Child Language Teaching and Therapy*. https://doi.org/10.1177/026565900001600201

Botting, N., & Conti-Ramsden, G. (2008). The role of language, social cognition, and social skill in the functional social outcomes of young adolescents with and without a history of SLI. *British Journal of Developmental Psychology*, *26*(2), 281–300. https://doi.org/10.1348/026151007X235891

Botting, N., Riches, N., Gaynor, M., & Morgan, G. (2010). Gesture production and comprehension in children with specific language impairment. *British Journal of Developmental Psychology*, *28*(1), 51–69. https://doi.org/10.1348/026151009X482642

Bowerman, M. (1988). No The '"no negative evidence"' problem: How do children avoid constructing an overly general grammar? In J. A. Hawkins (Ed.), *Explaining language universals* (pp. 73–101). Blackwell.

Bowerman, M. (1975). Early syntactic development: a cross-linguistic study with special reference to Finnish. In *Lingua*. https://doi.org/10.1016/0024-3841(75)90034-0

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., Götz, K., & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *49*(4), 422–432. https://doi.org/10.1111/j.1469-7610.2007.01849.x

Boyle, J., McCartney, E., Forbes, J., & O'Hare, A. (2007a). A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment. *Health Technology Assessment*, *11*(25). https://doi.org/10.3310/hta11250

Boyle, J., McCartney, E., Forbes, J., & O'Hare, A. (2007b). A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment. *Health Technology Assessment*, *11*(25), 1–158. https://doi.org/10.3310/hta11250

Boyle, J, Mccartney, E., O'hare, A., & Law, J. (2010). Intervention for mixed receptive-expressive language impairment: A review. *Developmental Medicine and Child Neurology*, *52*(11), 994–999. https://doi.org/10.1111/j.1469-8749.2010.03750.x

Boyle, J., Mccartney, E., O'hare, A., & Law, J. (2010). Intervention for mixed receptive-expressive language impairment: A review. *Developmental Medicine and Child Neurology*, *52*(11), 994–999. https://doi.org/10.1111/j.1469-8749.2010.03750.x

Braine, M. D. S., & Bowerman, M. (1976). Children's First Word Combinations. *Monographs of the Society for Research in Child Development*, *41*(164), 1–104. https://doi.org/10.1111/j.1540-5834.2006.00381.x

Braun, V., & Clarke, V. (2006a). Qualitative Research in Psychology Using thematic analysis in psychology Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Braun, V., & Clarke, V. (2006b). Qualitative Research in Psychology Using thematic analysis in psychology Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. http://www.tandfonline.com/action/journalInformation?journalCode=uqrp20%5Cnhttp://www.tandfonline.com/action/journalInformation?journalCode=uqrp20

Bray, J., Schlenger, W., Zarkin, G., & Galvin, D. (2009). Analyzing Data from Nonrandomized Group Studies. In *Analyzing Data from Nonrandomized Group Studies*.

https://doi.org/10.3768/rtipress.2008.mr.0008.0811

Bruns, C., Varley, R., Zimmerer, V. C., Carragher, M., Brekelmans, G., & Beeke, S. (2019). "I don't know": a usage-based approach to familiar collocations in non-fluent aphasia. *Aphasiology*, *33*(2), 140–162. https://doi.org/10.1080/02687038.2018.1535692

Bruzzese, J. M., Gallagher, R., McCann-Doyle, S., Reiss, P. T., & Wijetunga, N. A. (2009). Effective methods to improve recruitment and retention in school-based substance use prevention studies. *Journal of School Health*, *79*(9), 400–407. https://doi.org/10.1111/j.1746-1561.2009.00427.x

Bryan, A. (2008). *Colourful Semantics: Thematic Role Therapy*. https://doi.org/10.1002/9780470699157.ch10

Bryan, K., Freer, J., & Furlong, C. (2007). Language and communication difficulties in juvenile offenders. *International Journal of Language & Communication Disorders*, *42*(5), 505–520. https://doi.org/10.1080/13682820601053977

Burchinal, M. R. (2008). How measurement error affects the interpretation and understanding of effect sizes. *Child Development Perspectives*, *2*(3), 178–180. https://doi.org/10.1111/j.1750-8606.2008.00062.x

Burgoyne, K., Duff, F. J., Clarke, P. J., Buckley, S., Snowling, M. J., & Hulme, C. (2012). Efficacy of a reading and language intervention for children with Down syndrome: A randomized controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/j.1469-7610.2012.02557.x

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425–455. https://doi.org/10.1080/01690969508407111

Cable, A. L., & Domsch, C. (2011). Systematic review of the literature on the treatment of children with late language emergence. *International Journal of Language and Communication Disorders*, *46*(2), 138–154. https://doi.org/10.3109/13682822.2010.487883

Camarata, S. M., Nelson, K. E., & Camarata, M. N. (1994). Comparison of Conversational-Recasting and Imitative Procedures for Training Grammatical Structures in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *37*(6), 1414–1423. https://doi.org/10.1044/jshr.3706.1414

Camarata, S., Nelson, K. E., Gillum, H., & Camarata, M. (2009). Incidental receptive language growth associated with expressive grammar intervention in SLI. *First Language*, *29*(1), 51–63. https://doi.org/10.1177/0142723708098810

Campbell, M., Fitzpatrick, R., Haines, A., Kinmouth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *BMJ*, *321*, 694–696. https://doi.org/10.1136/bmj.321.7262.694

Cane, J., Connor, D. O., & Michie, S. (2012). Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implementation Science*, *7*(37), 1–17.

Capone, N. C., & Mcgregor, K. K. (2004). Gesture Development: A Review for Clinical and Research Practices. *Journal of Speech Language and Hearing Research*, *47*, 173–186.

Capone Singleton, N., & Anderson, L. (2020). Making Object Shape Explicit for Toddlers With Late Language Emergence. In *Journal of speech, language, and hearing research.* (Vol. 63, Issue 3, pp. 749–763). https://doi.org/10.1044/2019_JSLHR-19-00235

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- to 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*, 315–330.

Carter, A. S., Messinger, D. S., Stone, W. L., Celimli, S., Nahmias, A. S., & Yoder, P. (2011). A randomized controlled trial of Hanen's "More Than Words" in toddlers with early autism symptoms. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/j.1469-7610.2011.02395.x

Catalogue of Bias Collaboration, Nunan, D., & Heneghan, C. (2018). Lack of blinding. *Catalogue Of Bias*.

Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions and evidence. *Language and Cognitive Processes*. https://doi.org/10.1080/01690960042000012

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology*. https://doi.org/10.1037/0012-1649.37.6.739

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The M.I.T Press.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. MIT Press.

Chomsky, N. (1995). *The minimalist program*. The MIT Press.

Chomsky, N. (2006). *Language and mind* (Third edit). Cambridge. https://doi.org/10.2174/1874440001004020081

Choudhury, N., & Benasich, A. A. (2003). A family aggregation study: The influence of family history and other risk factors on language development. *Journal of Speech, Language, and Hearing Research*, *46*(2), 261–272. https://doi.org/10.1044/1092-4388(2003/021)

Clahsen, H. (1989). The Grammatical Characterization of Developmental Dysphasia. *Linguistics*, *27*(5), 897. https://doi.org/10.1515/ling.1989.27.5.897

Clahsen, H. (1991). *Child language and developmental dysphasia linguistic studies of the acquisition of German* (K. Richman (ed.)). Amsterdam : J. Benjamins Pub. Co.

Cleave, P. L., & Fey, M. E. (1997). Two Approaches to the Facilitation of Grammar in Children with Language Impairments: Rationale and Description. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/1058-0360.0601.22

Clegg, J., Law, J., Rush, R., Peters, T. J., & Roulstone, S. (2015). The contribution of early language development to children's emotional and behavioural functioning at 6 years: an analysis of data from the Children in Focus sample from the ALSPAC birth cohort. *Journal of Child Psychology* *and Psychiatry*, *56*(1), 67–75.

Coady, J. A., Kluender, K. R., & Evans, J. L. (2005). Categorical Perception of Speech by Children With Specific Language Impairments. *Journal of Speech Language and Hearing Research*, *48*(4), 944. https://doi.org/10.1044/1092-4388(2005/065)

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. In *Statistical Power Analysis for the Behavioral Sciences*. https://doi.org/10.1234/12345678

Cohen, N. J., Farnia, F., & Im-Bolter, N. (2013). Higher order language competence and adolescent mental health. *Journal of Child Psychology and Psychiatry*, *54*(7), 733–744.

Colquhoun, H. L., Squires, J. E., Kolehmainen, N., Fraser, C., & Grimshaw, J. M. (2017). Methods for designing interventions to change healthcare professionals' behaviour: A systematic review. *Implementation Science*, *12*(1), 1–11. https://doi.org/10.1186/s13012-017-0560-5

Cognitive Concepts. (1997). *Earobics: Auditory development and phonics program*. Cognitive Concepts.

Conti-Ramsden, G., & Botting, N. (2004). Social Difficulties and Victimization in Children With SLI at 11 Years of Age. *Journal of Speech Language and Hearing Research*, *47*(1), 145. https://doi.org/10.1044/1092-4388(2004/013)

Conti-Ramsden, G., & Botting, N. (2008). Emotional health in adolescents with and without a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *49*(5), 516–525. https://doi.org/10.1111/j.1469-7610.2007.01858.x

Conti-Ramsden, G., Botting, N., Simkin, Z., & Knox, E. (2001). Follow-up of children attending infant language units: outcomes at 11 years of age. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *36*(2), 207–219. https://doi.org/10.1080/1368282001001988

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language and Communication Disorders*, *53*(2), 237–255. https://doi.org/10.1111/1460-6984.12338

Conti-Ramsden, G., (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *46*(5), 1029–1037. https://doi.org/10.1044/1092-4388(2003/082)

Conti-Ramsden, G., & Durkin, K. (2016). What Factors Influence Language Impairment? Considering Resilience as well as Risk. *Folia Phoniatrica et Logopaedica*, *67*(6), 293–299. https://doi.org/10.1159/000444750

Conti-Ramsden, G., Durkin, K., Simkin, Z., & Knox, E. (2009). Specific language impairment and school outcomes. I: Identifying and explaining variability at the end of compulsory education. *International Journal of Language and Communication Disorders*, *44*(1), 15–35. https://doi.org/10.1080/13682820801921601

Conti-Ramsden, G., Mok, P. L. H., Pickles, A., & Durkin, K. (2013). Adolescents with a history of specific language impairment (SLI): Strengths and difficulties in social, emotional and behavioral functioning. *Research in Developmental Disabilities*, *34*(11), 4161–4169. https://doi.org/10.1016/j.ridd.2013.08.043

Conti-Ramsden, G., & Windfuhr, K. (2002). Productivity with word order and morphology: a comparative look at children with SLI and children with normal language abilities. *International Journal of Language & Communication Disorders*, *37*(1), 17–30. https://doi.org/10.1080/13682820110089380

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., Health, P., Unit, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: new guidance. In *Medical Reseach Council*. https://doi.org/10.1136/bmj.a1655

Crawford, J. R., Garthwaite, P. H., & Porter, S. (2010). Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cognitive Neuropsychology*. https://doi.org/10.1080/02643294.2010.513967

Croft, W. (2001). Radical Construction Grammar. *Oxford Oxford University Press Croft William 2003 Typology And*. https://doi.org/10.1093/acprof:oso/9780198299554.001.0001

Crooke, P. J., Hendrix, R. E., & Rachman, J. Y. (2008). Brief report: Measuring the effectiveness of teaching social thinking to children with Asperger syndrome (AS) and High Functioning Autism (HFA). *Journal of Autism and Developmental Disorders*, *38*(3), 581–591. https://doi.org/10.1007/s10803-007-0466-1

Dąbrowska, E., & Lieven, E. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics*. https://doi.org/10.1515/cogl.2005.16.3.437

Davies, P., Shanks, B., & Davies, K. (2004). Improving narrative skills in young children with delayed language development. *Educational Review*. https://doi.org/10.1080/0013191042000201181

Davis, E. (2016). What is the differential response of Building Early Sentences Therapy, based on pre-school children's language profiles and gender, when compared to intervention as normal? *Unpublished Undergraduate Dissertation*.

Department of Communities and Local Government. (2015). The English Indices of Deprivation 2015: Research Report. In *DCLG*.

Des Jarlais, D. C., Lyles, C., & Crepaz, N. (2004). Improving the Reporting Quality of Nonrandomized Evaluations of Behavioral and Public Health Interventions: The TREND Statement. In *American Journal of Public Health*. https://doi.org/10.2105/AJPH.94.3.361

Deverka, P. A., Lavallee, D. C., Desai, P. J., Esmail, L. C., Ramsey, S. D., Veenstra, D. L., & Tunis, S. R. (2012). Stakeholder participation in comparative effectiveness research: defining a framework for effective engagement. *Journal of Comparative Effectiveness Research*, *1*(2), 181–194. https://doi.org/10.2217/cer.12.7

Dockrell, J., Bakopoulou, I., Law, J., & Spencer, S. (2012). *Developing a communication supporting classrooms observation tool*.

Dockrell, J. E., Bakopoulou, I., Law, J., Spencer, S., & Lindsay, G. (2015). Capturing communication supporting classrooms: The development of a tool and feasibility study. *Child Language Teaching and Therapy*. https://doi.org/10.1177/0265659015572165

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. *Child and Adolescent Mental Health*, *20*(2), 116–125. https://doi.org/10.1111/camh.12072

Dockrell, J., Lindsay, G., Roulstone, S., & Law, J. (2014). Supporting children with speech, language and communication needs: An overview of the results of the Better Communication Research Programme. *International Journal of Language and Communication Disorders*, *49*(5), 543–557. https://doi.org/10.1111/1460-6984.12089

Dollaghan, C. A., & Campbell, T. F. (2009). How well do poor language scores at ages 3 and 4 predict poor language scores at age 6. *International Journal of Speech-Language Pathology*, *11*(5), 358–365. https://doi.org/10.1080/17549500903030824

Durkin, K., Simkin, Z., Knox, E., & Conti-Ramsden, G. (2009). Specific language impairment and school outcomes. II: Educational context, student satisfaction, and post-compulsory progress. *International Journal of Language and Communication Disorders*, *44*(1), 36–55. https://doi.org/10.1080/13682820801921510

Ebbels, S. (2007). Teaching grammar to school-aged children with specific language impairment using Shape Coding. *Child Language Teaching and Therapy*, *23*(1), 67–93. https://doi.org/10.1191/0265659007072143

Ebbels, S. (2014). Effectiveness of intervention for grammar in school-aged children with primary language impairments: A review of the evidence. *Child Language Teaching and Therapy*, *30*(1), 7–40. https://doi.org/10.1177/0265659013512321

Ebbels, S. H., McCartney, E., Slonims, V., Dockrell, J. E., & Norbury, C. F. (2019). Evidence-based pathways to intervention for children with language disorders. *International Journal of Language and Communication Disorders*. https://doi.org/10.1111/1460-6984.12387

Ebbels, S. H., van der Lely, H. K. J., & Dockrell, J. E. (2007). Intervention for Verb Argument Structure in Children With Persistent SLI: A Randomized Control Trial. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2007/093)

Edwards, S., Letts, C., & Sinka, I. (2011). *The New Reynell Developmental Language Scales*. GL Assessment.

Eidsvåg, S. S., Plante, E., Oglivie, T., Privette, C., & Mailend, M. L. (2019). Individual versus small group treatment of morphological errors for children with developmental language disorder. *Language, Speech, and Hearing Services in Schools*, *50*(2), 237–252. https://doi.org/10.1044/2018_LSHSS-18-0033

Eikeseth, S. (2009). Outcome of comprehensive psycho-educational interventions for young children with autism. *Research in Developmental Disabilities*, *30*, 158–178. https://doi.org/10.1016/j.ridd.2008.02.003

El Dib, R., Jorge, E. C., Kamegasawa, A., Daher, S. R., Spagnuolo, R. S., da Silva, M. P., Braga, G. P., Volpato, E., Módolo, N. S. P., Betini, M., Do Valle, A., Corrêa, I., Bazan, R., Almeida, R. A. M. B., Weber, S. A. T., Molina, S., Yoo, H., Boas, P. V., Corrente, J. E., … Do Nascimento Junior, P. (2015). Differences between the real and the desired worlds in the results of clinical trials. *Clinics*, *70*(9), 618–622. https://doi.org/10.6061/clinics/2015(09)04

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., Lancaster, G. A., Altman, D., Bretz, F., Campbell, M., Cobo, E., Craig, P., Davidson, P., Groves, T., Gumedze, F., Hewison, J., Hirst, A., Hoddinott, P., Lamb, S. E., … Tugwell, P. (2016). CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. *BMJ*. https://doi.org/10.1136/bmj.i5239

Eldridge, S. M., Lancaster, G. A., Campbell, M. J., Thabane, L., Hopewell, S., Coleman, C. L., & Bond, C. M. (2016). Defining feasibility and pilot studies in preparation for randomised controlled trials: Development of a conceptual framework. *PLoS ONE*, *11*(3). https://doi.org/10.1371/journal.pone.0150205

Eldridge, S. M., Lancaster, G. A., Campbell, M. J., Thabane, L., Hopewell, S., Coleman, C. L., & Bond, C. M. (2016). Defining feasibility and pilot studies in preparation for randomised controlled trials: Development of a conceptual framework. *PLoS ONE*, *11*(3). https://doi.org/10.1371/journal.pone.0150205

Evans, J. L. (2001). An emergent account of language impairments in children with SLI: Implications for assessment and intervention. *Journal of Communication Disorders*, *34*(1–2), 39–54. https://doi.org/10.1016/S0021-9924(00)00040-X

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *52*(2), 321–335. https://doi.org/10.1044/1092-4388(2009/07-0189)

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*(2), 234–248. https://doi.org/10.1111/desc.12019

Fey, M.E., & Finestack, L. H. (2009). Research and development in child language intervention: A five-phase model. In R. G. Schwartz (Ed.), *Handbook of Child Language Disorders: 2nd Edition* (pp. 513–529). Psychology Press.

Fey, M. E., Cleave, P. L., Long, S. H., & Hughes, D. L. (1993). Two Approaches to the Facilitation of Grammar in Children With Language Impairment: An Experimental Evaluation. *J Speech Hear Res*.

Fey, M. E., Long, S. H., & Finestack, L. H. (2003). Ten principles of grammar facilitation for children with specific language impairments. In *American Journal of Speech-Language Pathology*.

https://doi.org/10.1044/1058-0360(2003/048)

Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel Modeling Using R*. Chapman and Hall/CRC Press.

Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*, *15*, 451–474. https://doi.org/10.1016/0091-7435(86)90024-1

Flintoff, J. (2016). Children's social and emotional adjustment and response to intervention: An analysis of the "building early sentences therapy" pilot and feasibility trial. *Unpublished Undergraduate Dissertation*.

Forrester, M. A. (1992). *Development of Young Children's Social-cognitive Skills*. Hove : L.Erlbaum.

Frans, N., Post, W. J., Huisman, M., Oenema-Mostert, I. C. E., Keegstra, A. L., & Minnaert, A. E. M. G. (2017). Early identification of children at risk for academic difficulties using standardized assessment: stability and predictive validity of preschool math and language scores. *European Early Childhood Education Research Journal*, *25*(5), 698–716. https://doi.org/10.1080/1350293X.2017.1356524

Fricke, S., Bowyer-Crane, C., Haley, A. J., Hulme, C., & Snowling, M. J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.12010

Fricke, S., Burgoyne, K., Bowyer-Crane, C., Kyriacou, M., Zosimidou, A., Maxwell, L., Lervåg, A., Snowling, M. J., & Hulme, C. (2017). The efficacy of early language intervention in mainstream school settings: a randomized controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(10), 1141–1151. https://doi.org/10.1111/jcpp.12737

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language*, *31*(3), 661–681. https://doi.org/10.1017/S0305000904006269

Friedmann, N., & Novogrodsky, R. (2007). Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain and Language*, *101*(1), 50–63. https://doi.org/10.1016/j.bandl.2006.09.006

Frizelle, P., Tolonen, A. K., Tulip, J., Murphy, C. A., Saldana, D., & McKean, C. (n.d.-a). *The influence of intervention dose form on oral language outcomes for children with Developmental Language Disorder: a systematic review and narrative synthesis.*

Frizelle, P., Tolonen, A. K., Tulip, J., Murphy, C. A., Saldana, D., & McKean, C. (n.d.-b). *The influence of quantitative intervention dosage on oral language outcomes for children with Developmental Language Disorder: a systematic review and narrative synthesis.*

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-Intervention: Definitions, Evidence, and Implications for the Learning Disabilities Construct. *Learning Disabilities Research and Practice*, *18*(3), 157–171. https://doi.org/10.1111/1540-5826.00072

Gascoigne, M. T. (ed). (2012). *Better Communication: Shaping speech, language and communication services for children and young people*. RCSLT.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200–213. https://doi.org/10.1016/0749-596X(89)90044-2

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*(3), 336–360. https://doi.org/10.1016/0749-596X(90)90004-J

Gentner, D., & Markman, A. B. (1995). *Similarity is like analogy: Structural alignment in comparison*.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45.

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition*, *65*(2–3), 263–

297. https://doi.org/10.1016/S0010-0277(98)00002-X

Ghassabian, A., Rescorla, L., Henrichs, J., Jaddoe, V. W., Verhulst, F. C., & Tiemeier, H. (2014). Early lexical development and risk of verbal and nonverbal cognitive delay at school age. *Acta Paediatrica, International Journal of Paediatrics*, *103*(1), 70–80. https://doi.org/10.1111/apa.12449

Gibbard, D. (1994). Parental-based intervention with pre-school language-delayed children. *International Journal of Language & Communication Disorders*. https://doi.org/10.3109/13682829409041488

Gibbard, D., & Smith, C. (2016). A transagency approach to enabling access to parent-based intervention for language delay in areas of social disadvantage: A service evaluation. *Child Language Teaching and Therapy*, *32*(1), 19–33. https://doi.org/10.1177/0265659014567785

Gillam, R. B., Loeb, D. F., Hoffman, L. M., Bohman, T., Champlin, C. A., Thibodeau, L., Widen, J., Brandel, J., & Friel-Patti, S. (2013). The Efficacy of Fast ForWord-Language Intervention in School Aged Children with Language Impairment: A Randomized Controlled Trial. *Journal of Speech, Language, and Hearing Research*, *51*(1), 97–119. https://doi.org/10.1021/nn300902w.Release

Girolametto, L., & Pearce, P. S. (1996). Interactive focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech & Hearing Research*.

Girolametto, L., Sussman, F., & Weitzman, E. (2007). Using case study methods to investigate the effects of interactive intervention for children with autism spectrum disorders. *Journal of Communication Disorders*. https://doi.org/10.1016/j.jcomdis.2006.11.001

Girolametto, L., & Weitzman, E. (2006). Chapter: It Takes Two to Talk--The Hanen Program for Parents: Early Language Intervention Through Caregiver Training. In *Treatment of language disorders in children*.

Gladfelter, A., Wendt, O., & Subramanian, A. (2011). Evidence- based speech and language intervention techniques for the birth-to-3 population. *EBP Briefs*, *5*(5), 1–10.

Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health*, *93*(8), 1261–1267.

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health*, *89*(9), 1322–1327. https://doi.org/10.2105/AJPH.89.9.1322

Glasziou, P., Meats, E., Heneghan, C., & Shepperd, S. (2008). What is missing from descriptions of treatment in trials and reviews? *BMJ*. https://doi.org/10.1136/bmj.39590.732037.47

Goldberg, A. (1995). Constructions: A construction grammar approach to argument structure. In *Chicago University Press*. Chicago University Press. https://doi.org/10.1016/S0378-2166(97)81937-6

Gomez, R. (2002). *Variability and Detection of Invariant Structure*. *13*(5), 431–436. https://doi.org/10.1007/s10162-016-0594-4

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135. https://doi.org/10.1016/S0010-0277(99)00003-7

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. https://doi.org/10.1097/00004583-200111000-00015

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Gray, C. A., & Garand, J. D. (1993). Social Stories: Improving Responses of Students with Autism with

Accurate Social Information. *Focus on Autism and Other Developmental Disabilities*, *8*(1), 1–10. https://doi.org/10.1177/108835769300800101

Gray, S. (2003). With Specific Language Impairment : What Predicts Success ? *Journal of Speech, Language, and Hearing Research*, *46*(1), 56–67. https://doi.org/10.1044/1092-4388(2003/005)

Gray, S. (2004). Word Learning by Preschoolers With Specific Language Impairment. *Journal of Speech Language and Hearing Research*, *48*(6), 1452. https://doi.org/10.1044/1092-4388(2005/101)

Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., Le Couteur, A., Leadbitter, K., Hudry, K., Byford, S., Barrett, B., Temple, K., Macdonald, W., & Pickles, A. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *The Lancet*. https://doi.org/10.1016/S0140-6736(10)60587-9

Greenwald, P., & Cullen, J. W. (1984). The Scientific Approach to Cancer Control. *CA: A Cancer Journal for Clinicians*, *34*, 328–332. https://doi.org/10.3322/canjclin.34.6.328

Group, L. R. (2019). *Language Intervention in the Early Years*. www.research.ncl.ac.uk/lively/ainterventions/best/

Grove, N., & Walker, M. (1990). The Makaton Vocabulary: Using Manual Signs and Graphic Symbols to Develop Interpersonal Communication. *Augmentative and Alternative Communication*. https://doi.org/10.1080/07434619012331275284

Guendouzi, J. (2003). "SLI", a generic category of language impairment that emerges from specific differences: A case study of two individual linguistic profiles. *Clinical Linguistics and Phonetics*, *17*(2), 135–152. https://doi.org/10.1080/0269920031000061803

Gustason, G. (1990). Signing Exact English. In M. H. Bornstein (Ed.), *Manual Communication: Implications for Education* (pp. 108–127). Gallaudet University Press.

Haith, M. M., & Benson, J. B. (1998). Infant cognition. In W. D. (Ed.) (Ed.), *Handbook of child* *psychology: Vol. 2. Cognition, perception, and language* (pp. 199–254). John Wiley & Sons Inc. https://doi.org/10.5962/bhl.title.7369

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children* (T. R. Risley (ed.)). Baltimore, Md. : Paul H. Brookes.

Hart, S. L., & Banda, D. R. (2010). Picture exchange communication system with individuals with developmental disabilities: A meta-analysis of single subject studies. *Remedial and Special Education*. https://doi.org/10.1177/0741932509338354

Harwood, C. (2016). What are the most appropriate outcome measures for the future Building Early Sentences Therapy. *Unpublished Undergraduate Dissertation*.

Henrichs, J., Rescorl, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W. V., Hofman, A., Raat, H., Verhulst, F. C., & Tiemeier, H. (2011). Examining continuity of early expressive vocabulary development: The generation R study. *Journal of Speech, Language, and Hearing Research*, *54*(3), 854–869. https://doi.org/10.1044/1092-4388(2010/09-0255)

Hirsh-Pasek, K., Golinkoff, R. M., & Hollich, G. (2000). An Emergentist Coalition Model for Word Learning: Mapping Words to Objects Is a Product of the Interaction of Multiple Cues. *Becoming a Word Learner: A Debate on Lexical Acquisition*, 136–164. https://doi.org/10.1093/acprof:oso/9780195130324.003.006

Hodges, L. E., Özçalışkan, Ş., & Williamson, R. (2018). Type of iconicity influences children's comprehension of gesture. *Journal of Experimental Child Psychology*, *166*, 327–339. https://doi.org/10.1016/j.jecp.2017.08.009

Hollo, A., Wehby, J. H., & Oliver, R. M. (2014). Unidentified Language Deficits in Children with Emotional and Behavioral Disorders: A Meta-Analysis. *Exceptional Children*, *80*(2), 169–186.

Howell, D. (2013). *Statistical Methods for Psychology* (8th ed., I). Wadsworth.

Isherwood, C. (2016). Acceptability and feasibility of Building Early Sentences Therapy (BEST) and experimental methods for parents , nursery staff and children involved : a pilot study. *Unpublished Undergraduate Dissertation*.

Iverson, J. M., Capirci, O., Longobardi, E., & Caselli, M. C. (1999). Gesturing in mother-child interactions. *Cognitive Development*, *14*(1), 57–75. https://doi.org/10.1016/S0885-2014(99)80018-5

Janiszewski, C., Noel, H., & Sawyer, A. G. (2003). A Meta-analysis of the Spacing Effect in Verbal Learning: Implications for Research on Advertising Repetition and Consumer Memory. *Journal of Consumer Research*. https://doi.org/10.1086/374692

Joanisse, M. F. (2004). Specific Language Impairments in Children: Phonology, Semantics, and the English Past Tense. *Current Directions in Psychological Science : A Journal of the American Psychological Society*, *13*(4), 156–160.

Joanisse, M. F. (2009). Model-based Approaches to Child Language Disorders. In R. G. Schwartz (Ed.), *Handbook of Child Language Disorders* (pp. 275–296). Psychology Press.

Joanisse, M. F., & Seidenberg, M. S. (2003). Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain and Language*, *86*(1), 40–56.

Jones, K. (2019). *Case study - Using Makaton to develop a Total Communication environment* (Vol. 1, Issue). https://doi.org/10.1017/CBO9781107415324.004

Justice, L. M., Logan, J., Jiang, H., & Schmitt, M. B. (2017). Algorithm-driven dosage decisions (AD3): Optimizing treatment for children with language impairment. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/2016_AJSLP-15-0058

Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *37*(2), 418–421. https://doi.org/10.1044/jshr.3702.418

Karmiloff-Smith, A, & Thomas, M. (2003). What can developmental disorders tell us about the neurocomputational constraints that shape development? The case of Williams syndrome. *Development and Psychopathology*, *15*(4), 969–990.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*(10), 389–398.

Keller, C. L., Brady, M. P., & Taylor, R. L. (2005). Using self evaluation to improve student teacher interns' use of specific praise. *Education and Training in Developmental Disabilities*, *40*(4), 368–376.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the. *The Relationship of Verbal and Nonverbal Communication*, *25*, 207.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511807572

Kilburn, M., Cannon, J., Mattox, T., & Shaw, R. (2014). *Programs That Work, from the Promising Practices Network on Children, Families and Communities*. 1–543. http://130.154.3.8/content/dam/rand/pubs/tools/TL100/TL145/RAND_TL145.pdf

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), 4–5. https://doi.org/10.1016/S0010-0277(02)00004-5

Klatte, I. S., Harding, S., & Roulstone, S. (2019). Speech and language therapists' views on parents' engagement in Parent–Child Interaction Therapy (PCIT). *International Journal of Language and Communication Disorders*. https://doi.org/10.1111/1460-6984.12459

Klatte, I. S., & Roulstone, S. (2016). The practical side of working with parent-child interaction therapy with preschool children with language impairments. *Child Language Teaching and Therapy*, *32*(3), 345–359. https://doi.org/10.1177/0265659016641999

Klee, T., Stringer, H., & Howard, D. (2009). Teaching evidence-based practice to speech and language therapy students in the United Kingdom. *Evidence Based Communication Assessment and Intervention.*, *3*(4), 195–207.

Knapp, T. R., & Schafer, W. D. (2009). From gain score t to ANCOVA F (and vice versa). *Practical Assessment, Research and Evaluation*, *14*(6).

Knowles, W., & Masidlover, M. (1982). *The Derbyshire Language Scheme*.

Kuoch, H., & Mirenda, P. (2003). Social Story Interventions for Young Children With Autism Spectrum Disorders. *Focus on Autism and Other Developmental Disabilities*, *18*, 219–227. https://doi.org/10.1177/10883576030180040301

Lal, R., & Bali, M. (2007). Effect of visual strategies on development of communication skills in children with autism. *Asia Pacific Disability Rehabilitation Journal*, *18*, 120–130.

Lal, R. (2010). Effect of Alternative and Augmentative Communication on Language and Social Behavior of Children with Autism. *Educational Research and Reviews*.

Lancaster, G. A. (2015). Pilot and feasibility studies come of age! *Pilot and Feasibility Studies*, *1*(1). https://doi.org/10.1186/2055-5784-1-1

Lancaster, G. A., Dodd, S., & Williamson, P. R. (2004). Design and analysis of pilot studies: Recommendations for good practice. *Journal of Evaluation in Clinical Practice*, *10*(2), 307–312. https://doi.org/10.1111/j..2002.384.doc.x

Langacker, R. W. (2000). A dynamic usage-based model. In *Usage-based models of language*.

Lapointe, S. G. (1983). Some issues in the linguistic description of agrammatism. *Cognition*, *14*(1), 1–39. https://doi.org/10.1016/0010-0277(83)90025-2

Law, J., Charlton, J., Dockrell, J., Gascoigne, J., McKean, M., & Theakston, C. (2017). *Early Language Development: Needs, provision and intervention for preschool children from socio-economically disadvantage backgrounds Early Language Development: Needs, provision, and intervention for preschool children from socio-economically disadvantag*.

Law, J., Garrett, Z., & Nye, C. (2010). Speech and language therapy interventions for children with primary speech and language delay or disorder (Review). *The Cochrane Collaboration*, *5*. https://doi.org/10.1002/14651858.CD004110.www.cochranelibrary.com

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling developmental language difficulties from school entry into adulthood: Literacy, mental health, and employment outcomes. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1401–1416. https://doi.org/10.1044/1092-4388(2009/08-0142)

Law, J., Garrett, Z., Nye, C., Zoe, G., & Nye, C. (2004). The Efficacy of Treatment for Children With Developmental Speech and ... *Journal of Speech, Language, and Hearing Research*, *47*(4), 924–943. https://doi.org/10.1044/1092-4388(2004/069)

Law, J., Lee, W., Roulstone, S., Wren, Y., Zeng, B., & Lindsay, G. (2012). “What Works”: Interventions for children and young people with speech , language and communication needs : Technical Annex. In *Department of Education Research Report*.

Law, J., Plunkett, C. C., & Stringer, H. (2012). Communication interventions and their impact on behaviour in the young child: A systematic review. *Child Language Teaching and Therapy*, *28*(1), 7–23. https://doi.org/10.1177/0265659011414214

Law, J., Roulstone, S., & Lindsay, G. (2015). Integrating external evidence of intervention effectiveness with both practice and the parent perspective: Development of “What Works” for speech, language, and communication needs. In *Developmental Medicine and Child Neurology*. https://doi.org/10.1111/dmcn.12630

Law, J., Rush, R., Anandan, C., Cox, M., & Wood, R. (2012). Predicting language change between 3 and 5 years and its implications for early identification. *Pediatrics*, *130*(1). https://doi.org/10.1542/peds.2011-1673

Law, J., Schoon, I., & Parsons, S. (2009). *Modeling Developmental Language Difficulties From School Entry Into Adulthood: Literacy, Mental Health, and Employment Outcomes*. *52*(December), 1401–1416.

Law, J., Tomblin, J. B., & Zhang, X. (2008). Characterizing the growth trajectories of language-impaired children between 7 and 11 years of age. *Journal of Speech, Language, and Hearing Research*, *51*(3), 739–749. https://doi.org/10.1044/1092-4388(2008/052)

Le, H. N. D., Mensah, F., Eadie, P., McKean, C., Sciberras, E., Bavin, E. L., Reilly, S., & Gold, L. (2020). Health-related quality of life of children with low language from early childhood to adolescence: results from an Australian longitudinal population-based study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.13277

Lea, J. (1965). A language scheme for children suffering from receptive aphasia. *Speech Pathology and Therapy*, *8*, 58–68.

Lee, W., & Pring, T. (2016). Supporting language in schools: Evaluating an intervention for children with delayed language in the early school years. *Child Language Teaching and Therapy*, *32*(2), 135–146. https://doi.org/10.1177/0265659015590426

Leonard, L. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics*, *10*(2), 179–202. https://doi.org/10.1017/S0142716400008511

Leonard, L. (1996). Characterizing specific language impairment: A crosslinguistic perspective. In M L Rice (Ed.), *Toward a genetics of language* (pp. 243–256). Lawrence Erlbaum Associates.

Leonard, L. (2003). Specific language impairment: Characterizing the deficit. *Towards a Definition of SLI*, 209–231.

Leonard, L. (2014). *Children with Specific Language Impairment* (Second Edi). The MIT Press.

Leonard, L. B., McGregor, K. K., & Allen, G. D. (1992). Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech and Hearing Research*, *35*(5), 1076–1086.

Leonard, L. B., Deevy, P., Karpicke, J. D., Christ, S., Weber, C., Kueser, J. B., & Haebig, E. (2019).

Adjective learning in young typically developing children and children with developmental language disorder: A retrieval-based approach. *Journal of Speech, Language, and Hearing Research*, *62*(12), 4433–4449. https://doi.org/10.1044/2019_JSLHR-L-19-0221

Leonard, L. B. (2007). Processing limitations and the grammatical profile of children with specific language impairment. *Advances in Child Development and Behavior*, *35*, 139.

Letts, C., Edwards, S., Sinka, I., Schaefer, B., & Gibbons, W. (2013). Socio-economic status and language acquisition: Children's performance on the new Reynell Developmental Language Scales. *International Journal of Language and Communication Disorders*, *48*(2), 131–143. https://doi.org/10.1111/1460-6984.12004

Levati, S., Campbell, P., Frost, R., Dougall, N., Wells, M., Donaldson, C., & Hagen, S. (2016). Optimisation of complex health interventions prior to a randomised controlled trial: A scoping review of strategies used. *Pilot and Feasibility Studies*, *2*(1). https://doi.org/10.1186/s40814-016-0058-y

Lhakhang, P., Lippke, S., Knoll, N., & Schwarzer, R. (2015). Evaluating brief motivational and self-regulatory hand hygiene interventions: A cross-over longitudinal design. *BMC Public Health*, *15*(1), 1–9. https://doi.org/10.1186/s12889-015-1453-7

Lindsay, G., Dockrell, J., Law, J., & Roulstone, S. (2012). *The Better Communication Research Programme: Improving provision for children and young people with speech, language and communication needs*.

Lloyd, S. (1998). *The phonics handbook: A handbook for teaching, writing and spelling* (3rd editio). Jolly Learning Ltd.

Locke, A., Ginsborg, J., & Peers, I. (2002). Development and disadvantage: implications for the early years and beyond. *Int. j. Lang. Comm. Dis*, *37*(1), 3–15. https://doi.org/10.1080/1368282011008991

Locke, J. (1954). *Essays on the law of nature : the Latin text, with a translation, introd. and notes,* *together with transcripts of Locke's shorthand in his journal for 1676* (W. von (Wolfgang von) Leyden (ed.)). Oxford, Clarendon Press.

Lubliner, S., & Smetana, L. (2005). The effects of comprehensive vocabulary instruction on Title I students' metacognitive word-learning skills and reading comprehension. *Journal of Literacy Research*. https://doi.org/10.1207/s15548430jlr3702_3

Lüke, C., Ritterfeld, U., Grimminger, A., Rohlfing, K. J., & Liszkowski, U. (2020). Integrated Communication System: Gesture and Language Acquisition in Typically Developing Children and Children With LD and DLD. *Frontiers in Psychology*, *11*(February), 1–13. https://doi.org/10.3389/fpsyg.2020.00118

MacWhinney, B. (2014). What we have learned. *Journal of Child Language*, *41*(S1), 124–131. https://doi.org/DOI: 10.1017/S0305000914000142

Mancil, G. R. (2009). Milieu Therapy as a Communication Intervention: A Review of the Literature Related to Children with Autism Spectrum Disorder. *Education and Training in Developmental Disabilities*.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*(5398), 77–80. https://doi.org/10.1126/science.283.5398.77

Marentette, P., & Nicoladis, E. (2011). Preschoolers' interpretations of gesture: Label or action associate? *Cognition*, *121*(3), 386–399. https://doi.org/10.1016/j.cognition.2011.08.012

Mars, T., Ellard, D., Carnes, D., Homer, K., Underwood, M., & Taylor, S. J. C. (2013). Fidelity in complex behaviour change interventions: A standardised approach to evaluate intervention integrity. *BMJ Open*, *3*. https://doi.org/10.1136/bmjopen-2013-003555

Marton, K. (2009). Imitation of body postures and hand movements in children with specific language impairment. *Journal of Experimental Child Psychology*, *102*(1), 1–13. https://doi.org/10.1016/j.jecp.2008.07.007

McArthur, G. M., & Bishop, D. V. M. (2004). Frequency Discrimination Deficits in People With Specific Language Impairment. *Journal of Speech Language and Hearing Research*, *47*(3), 527. https://doi.org/10.1044/1092-4388(2004/041)

McCartney, E., Boyle, J., Ellis, S., Bannatyne, S., & Turnbull, M. (2011). Indirect language therapy for children with persistent language impairment in mainstream primary schools: Outcomes from a cohort intervention. *International Journal of Language and Communication Disorders*, *46*(1), 74–82. https://doi.org/10.3109/13682820903560302

McCartney, E., Ellis, S., Boyle, J., Turnbull, M., & Kerr, J. (2010). Developing a language support model for mainstream primary school teachers. *Child Language Teaching and Therapy*, *26*(3), 359–374. https://doi.org/10.1177/0265659010369306

McConachie, H., Randle, V., Hammal, D., & Le Couteur, A. (2005). A controlled trial of a training course for parents of children with suspected autism spectrum disorder. *Journal of Pediatrics*. https://doi.org/10.1016/j.jpeds.2005.03.056

McGregor, K, K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2002/081)

McGregor, K.K. (2020). How We Fail Children With Developmental Language Disorder. *Language, Speech, and Hearing Services in Schools*, 1–12.

McKean, C., Benson, K., Jack, C., Letts, C. A., Pert, S., Preston, E., Trebacz, A., Stringer, H., & Wareham, H. (2020). Language intervention in the early years - comparing the effectiveness of language intervention approaches for pre-school children with language difficulties. *ISRCTN10974028:* https://doi.org/10.1186/ISRCTN10974028

Mckean, C., Letts, C., & Howard, D. (2013). Developmental Change Is Key to Understanding Primary Language Impairment: The Case of Phonotactic Probability and Nonword Repetition. *Journal of Speech and Hearing Research*, *56*(October), 1579–1594. https://doi.org/10.1044/1092-

4388(2013/12-0066)of

McKean, C., Mensah, F. K., Eadie, P., Bavin, E. L., Bretherton, L., Cini, E., & Reilly, S. (2015). Levers for language growth: Characteristics and predictors of language trajectories between 4 and 7 years. *PLoS ONE*, *10*(8), 1–21. https://doi.org/10.1371/journal.pone.0134251

McKean, C., Stow, C., & Pert, S. (2013). *Building Early Sentences Therapy: The BEST Manual*.

McKean, C., Wraith, D., Eadie, P., Cook, F., Mensah, F., & Reilly, S. (2017). Subgroups in language trajectories from 4 to 11 years: the nature and predictors of stable, improving and decreasing language trajectory groups. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(10), 1081–1091. https://doi.org/10.1111/jcpp.12790

McKean, C., Law, J., Laing, K., Cockerill, M., Allon-Smith, J., McCartney, E., & Forbes, J. (2017). A qualitative case study in the social capital of co-professional collaborative co-practice for children with speech, language and communication needs. *International Journal of Language and Communication Disorders*, *52*(4), 514–527. https://doi.org/10.1111/1460-6984.12296

McKean, C., Pert, S., & Stow, C. (2012). Building early sentences in your language: a service evaluation study of the “BEST” home language intervention. *Poster Presentation at the Royal College of Speech and Language Therapists Conference, The Midland Hotel Manchester, UK, 11-12th September 2012*.

McKean, C., Wraith, D., Eadie, P., Cook, F., Mensah, F., & Reilly, S. (2017). Subgroups in language trajectories from 4 to 11 years: the nature and predictors of stable, improving and decreasing language trajectory groups. *Journal of Child Psychology and Psychiatry*, *58*(10), 1081–1091. https://doi.org/10.1111/jcpp.12790

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago press.

Meltzoff, A. (1995). Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children. *Developmental Psychology*, *31*, 838–850. https://doi.org/10.1038/jid.2014.371

Michie, S., Johnston, M., Abraham, C., Lawton, R., Parker, D., & Walker, A. (2005). Making psychological theory useful for implementing evidence based practice: A consensus approach. *Quality and Safety in Health Care*. https://doi.org/10.1136/qshc.2004.011155

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(42). https://doi.org/10.1186/1748-5908-6-42

Millard, S. K., Edwards, S., & Cook, F. M. (2009). Parent-child interaction therapy: Adding to the evidence. *International Journal of Speech-Language Pathology*, *11*(1), 61–76. https://doi.org/10.1080/17549500802603895

Ministry of Housing Communities and Local Government. (2015). *No Title*. http://imd-by-postcode.opendatacommunities.org/

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 (Clinical Research Ed.)*, *340*. https://doi.org/10.1136/bmj.c869

Montgomery, J., & Evans, J. (2009). Complex sentence comprehension and working memory in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *52*(2), 269–288. https://doi.org/10.1044/1092-4388(2008/07-0116)

Morrison, C. M., Hirsh, K. W., & Duggan, G. B. (2003). Age of acquisition, ageing, and verb production: Normative and experimental data. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*. https://doi.org/10.1080/02724980244000594

Motsch, H. J., & Marks, D. K. (2015). Efficacy of the Lexicon Pirate strategy therapy for improving lexical learning in school-age children: A randomized controlled trial. *Child Language Teaching and Therapy*. https://doi.org/10.1177/0265659014564678

Motsch, H. J., & Ulrich, T. (2012). Effects of the strategy therapy “lexicon pirate” on lexical deficits in preschool age: A randomized controlled trial. *Child Language Teaching and Therapy*. https://doi.org/10.1177/0265659011432943

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. Oxford University Press.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*. https://doi.org/10.1037/0012-1649.40.5.665

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. In *Biological Reviews*. https://doi.org/10.1111/j.1469-185X.2007.00027.x

Namy, L. L., Campbell, A. L., & Tomasello, M. (2004). The Changing Role of Iconicity in Non-Verbal Symbol Learning: A U-Shaped Trajectory in the Acquisition of Arbitrary Gestures. *Journal of Cognition and Development*, *5*(1), 37–57. https://doi.org/10.1207/s15327647jcd0501

Namy, L. L., & Waxman, S. R. (1998). Words and Gestures : Infants ' Interpretations of Different Forms of Symbolic Reference Author (s): Laura L . Namy and Sandra R . Waxman Published by : Wiley on behalf of the Society for Research in Child Development Stable URL : http://www.jstor.org/st. *Child Development*, *69*(2), 295–308.

National Institute for Health Research. (2012). *Involve: Briefing notes for researchers: involving the public in NHS, public health and social care research.*

National Institute for Health Research. (2018). *Feasibility and Pilot Studies*.

Nevin, J. A., Mandell, C., & Atak, J. R. (1983). The analysis of behavioural momentum. *Journal of the Experimental Analysis of Behaviour*, *39*(1), 49–59.

Nickisch, A., & von Kries, R. (2009). Short-Term Memory (STM) Constraints in Children With Specific Language Impairment (SLI): Are There Differences Between Receptive and Expressive SLI? *Journal of Speech Language and Hearing Research*, *52*(3), 578. https://doi.org/10.1044/1092-

4388(2008/07-0150)

Nind, M., & Hewett, D. (2006). *Access to Communication: Developing the basics of communication for people who have severe learning disabilities through Intensive Interaction* (2 edition). David Fulton.

Norbury, C. F., Bishop, D. V, & Briscoe, J. (2001). Production of English finite verb morphology. *Journal of Speech Language and Hearing Research*, *44*, 165. https://doi.org/10.1044/1092-4388(2001/015)

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *The Journal of Child Psychology and Psychiatry*, *57*(11), 1247–1257. https://doi.org/10.1111/jcpp.12573

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. https://doi.org/10.1073/PNAS.1708274114

Novack, M. A., Goldin-Meadow, S., & Woodward, A. L. (2015). Learning from gesture: How early does it happen? *Cognition*. https://doi.org/10.1016/j.cognition.2015.05.018

Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *International Journal of Speech-Language Pathology*, *8*(4), 364–375. https://doi.org/10.1080/14417040600919496

O'Cathain, A., Croot, L., Sworn, K., Duncan, E., Rousseau, N., Turner, K., Yardley, L., & Hoddinott, P. (2019). Taxonomy of approaches to developing interventions to improve health: A systematic methods overview. *Pilot and Feasibility Studies*, *5*(1), 1–27. https://doi.org/10.1186/s40814-019-0425-6

Oddson, B., Thomas-Stonell, N., Robertson, B., & Rosenbaum, P. (2019). Validity of a streamlined version of the Focus on the Outcomes of Communication Under Six: Process and outcome.

Child: Care, Health and Development, *45*(4), 600–605. https://doi.org/10.1111/cch.12669

Owens, G., Granader, Y., Humphrey, A., & Baron-Cohen, S. (2008). LEGO® therapy and the social use of language programme: An evaluation of two social skills interventions for children with high functioning autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *38*, 1944–1957. https://doi.org/10.1007/s10803-008-0590-6

Parsons, S., Law, J., & Gascoigne, M. (2005). Teaching receptive vocabulary to children with specific language impairment: A curriculum-based approach. *Child Language Teaching and Therapy*. https://doi.org/10.1191/0265659005ct280oa

Pascoe, M., Stackhouse, J., & Wells, B. (2005). Phonological therapy within a psycholinguistic framework: Promoting change in a child with persisting speech difficulties. *International Journal of Language and Communication Disorders*. https://doi.org/10.1080/13682820412331290979

Pawlowska, M., Leonard, L. B., Camarata, S. M., Brown, B., & Camarata, M. N. (2008). Factors accounting for the ability of children with SLI to learn agreement morphemes in intervention. *Journal of Child Language*, *35*(1), 25–53. https://doi.org/10.1017/S0305000907008227

Perniss, P., Özyürek, A., & Morgan, G. (2015). The influence of the visual modality on language structure and conventionalization: Insights from sign language and gesture. *Topics in Cognitive Science*, *7*(1), 2–11. https://doi.org/10.1111/tops.12127

Pincus, D. B., & Friedman, A. G. (2004). Improving Children's Coping with Everyday Stress: Transporting Treatment. *Clinical Child and Family Psychology Review*, *7*(4), 223–240. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15648277&retmode=ref&cmd=prlinks%0Apapers3://publication/uuid/5FCFF542-9182-4083-B71A-7DA7BB70F0DF

Pine, J. M., & Lieven, E. V. M. (1993). Reanalysing rote-learned phrases: Individual differences in the transition to multi-word speech. *Journal of Child Language*, *20*(3), 551–571. https://doi.org/10.1017/S0305000900008473

Pinker, S. (1979). *Formal models of language learning*. *7*, 217–283.

Pinker, S. (1984). Language Learnability and Language Development. In *Linguistics*. https://doi.org/10.1007/s13398-014-0173-7.2

Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: What does the task measure? *International Journal of Language and Communication Disorders*, *50*(1), 106–118. https://doi.org/10.1111/1460-6984.12126

Poncelas, A., & Murphy, G. (2007). Accessible information for people with intellectual disabilities: Do symbols really help? *Journal of Applied Research in Intellectual Disabilities*. https://doi.org/10.1111/j.1468-3148.2006.00334.x

Prathanee, B., Thinkhamrop, B., & Dechongkit, S. (2007). Factors associated with specific language impairment and later language development during early life: A literature review. *Clinical Pediatrics*, *46*(1), 22–29. https://doi.org/10.1177/0009922806297153

Preston, D., & Carter, M. (2009). A review of the efficacy of the picture exchange communication system intervention. In *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-009-0763-y

R Core Team. (2019). *R: A language and environment for statistical computing* (1.2.5001). R Foundation for Statistical COmputing. https://www.r-project.org/

Reilly, S., McKean, C., Levickis, P. (2014). Late talking: can it predict later language difficulties? *Centre for Research Excellence in Child Language, Research Snapshot 2*.

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E. L., Prior, M., Cini, E., Conway, L., Eadie, P., & Bretherton, L. (2010). Predicting Language Outcomes at 4 Years of Age: Findings From Early Language in Victoria Study. *PEDIATRICS*, *12*(6), 1530–1537. https://doi.org/10.1542/peds.2010-0254

Reilly, S., Bavin, E. L., Bretherton, L., Conway, L., Eadie, P., Cini, E., Prior, M., Ukoumunne, O. C., & Wake, M. (2009). The Early Language in Victoria Study (ELVS): A prospective, longitudinal study of communication skills and expressive vocabulary development at 8, 12 and 24 months.

International Journal of Speech-Language Pathology, *11*(5), 344–357. https://doi.org/10.1080/17549500903147560org/10.1080/17549500903147560

Reilly, S., Tomblin, B., Law, J., Mckean, C., Mensah, F. K., Morgan, A., Goldfeld, S., Nicholson, J. M., & Wake, M. (2014). *Specific language impairment : a convenient label for whom ?* 416–451. https://doi.org/10.1111/1460-6984.12102

Reynhout, G., & Carter, M. (2006). Social stories™ for children with disabilities. *Journal of Autism and Developmental Disorders*, *36*(4), 445–469. https://doi.org/10.1007/s10803-006-0086-1

Rhodes, A. (2001). *Rhodes to Language*. Stass Publications.

Riccio, C. A., & Jemison, S. J. (1998). Adhd and emergent literacy: Influence of language factors. *Reading and Writing Quarterly*, *14*(1), 43–58. https://doi.org/10.1080/1057356980140103

Rice, M. L. (2016). Children with Specific Language Impairment and Their Families: A Future View of Nature Plus Nurture and New Technologies for Comprehensive Language Intervention Strategies. *Seminars in Speech and Language*, *37*(4), 310–318. https://doi.org/10.1055/s-0036-1587701

Rice, M L, Sell, M. A, & Hadley, P. A. (1991). Social interactions of speech- and language-impaired children. *Journal of Speech and Hearing Research*. https://doi.org/10.1044/jshr.3406.1299

Rice, M. L, Smolik, F., Perpich, D., Thompson, T., Rytting, N., & Blossom, M. (2010). Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language, and Hearing Research*, *53*(2), 333–349. https://doi.org/10.1044/1092-4388(2009/08-0183)

Rice, M. L., Tomblin, J. B., Hoffman, L., Richman, W. A., & Marquis, J. (2004). Grammatical tense deficits in children with SLI and nonspecific language impairment: Relationships with nonverbal IQ over time. *Journal of Speech, Language, and Hearing Research*, *47*(4), 816–834. https://doi.org/10.1044/1092-4388(2004/061)

Rice, M. L., & Wexler, K. (1996a). A Phenotype of Specific Language Impairment: Extended Optional Infinitives. In *Towards a Genetics of Language*.

Rice, M. L., & Wexler, K. (1996b). Toward Tense as a Clinical Marker of Specific Language Impairment in English-Speaking Children. *Journal of Speech Language and Hearing Research*. https://doi.org/10.1044/jshr.3906.1239

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific Language Impairment as a Period of Extended Optional Infinitive. *Journal of Speech Language and Hearing Research*, *38*(4), 850. https://doi.org/10.1044/jshr.3804.850

Riches, N. G. (2013). Treating the passive in children with specific language impairment: A usage-based approach. *Child Language Teaching and Therapy*, *29*(2), 155–169. https://doi.org/10.1177/0265659012466667

Riches, N. G., Tomasello, M., & Conti-Ramsden, G. (2005). Verb Learning in Children With SLI: Frequency and Spacing Effects. *Journal of Speech, Language & Hearing Research*. https://doi.org/10.1044/1092-4388(2005/097)

Rippon, H. (2002). *Reception Narrative Pack*. Black Sheep Press.

Robey, R. R., & Schultz, M. C. (1998). A model for conducting clinical-outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology*, *12*(9791), 787–810. https://doi.org/10.1080/02687039808249573

Roschelle, J., Feng, M., Gallagher, H, A., Murphy, R., Harris, C., Kamdar, D., & Trinidad, G. (2014). *Recruiting Participants for Large-Scale Random Assignment Experiments in School Settings*. https://doi.org/10.13140/RG.2.1.3581.6167

Rose, E., Lehrl, S., Ebert, S., & Weinert, S. (2018). Long-term relations between children's language, the home literacy environment, and socioemotional development from ages 3 to 8. *Early Education and Development*, *29*(3), 342–356.

Rosnow, R. L., Rosenthal, R., & Rubin, D. B. (2000). *CONTRASTS AND EFFECT-SIZE ESTIMATION*. *11*(6), 446–453.

Rost, G. C., & McGregor, K. K. (2012). Miranda rights comprehension in young adults with specific language impairment. *American Journal of Speech-Language Pathology*, *21*(2), 101–108. https://doi.org/10.1044/1058-0360(2011/10-0094)

Roulstone, S., Wren, Y., Bakopoulou, I., Goodlad, S., & Lindsay, G. (2012). Exploring Interventions For Children And Young People With Speech, Language And Communication Needs: A Study Of Practice. *Better Communication Research Programme*, 86. http://dera.ioe.ac.uk/16323/1/DFE-RR247-BCRP10a.pdf

Rowe, J. (1981). The Paget-Gorman Sign System. *Brisitsh Journal of Special Education*, *8*(4), 25–27.

Royal College of Speech and Language Therapists. (2007). Good Practice for Speech and Language Therapists Working With Clients From Linguistic Minority Communities. In *Position Paper*.

Ruxton, G. D., & Beauchamp, G. (2008). Some suggestions about appropriate use of the Kruskal-Wallis test. *Animal Behaviour*, *76*(3), 1083–1087. https://doi.org/10.1016/j.anbehav.2008.04.011

Ryan, A., Gibbon, F. E., & Oshea, A. (2016). Expressive and receptive language skills in preschool children from a socially disadvantaged area. *International Journal of Speech-Language Pathology*, *18*(1), 41–52. https://doi.org/10.3109/17549507.2015.1089935

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Saxton, M. (2005). “Recast” in a new light: Insights for practice from typical language studies. In *Child*

Language Teaching and Therapy. https://doi.org/10.1191/0265659005ct279oa

Schmitt, M. B., Justice, L. M., & Logan, J. A. R. (2017). Intensity of language treatment: contribution to children's language outcomes. *International Journal of Language and Communication Disorders*. https://doi.org/10.1111/1460-6984.12254

Schoon, I., Parsons, S., Rush, R., & Law, J. (2010). Children's Language Ability and Psychosocial Development: A 29-Year Follow-up Study. *Pediatrics*, *126*(1), e73–e80. https://doi.org/10.1542/peds.2009-3282

Schroeder, A. (2001). *Time to Talk*. LDA.

Schulz, K. F., Altman, D. G., Moher, D., Barbour, V., Berlin, J. A., Boutron, I., Devereaux, P. J., Dickersin, K., Elbourne, D., Ellenberg, S., Gebski, V., Goodman, S., Gøtzsche, P. C., Groves, T., Grunberg, S., Haynes, B., Hopewell, S., James, A., Juhn, P., … Guyatt, G. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *Ann Int Med*, *152*. https://doi.org/10.3736/jcim20100702

Schütze, C., & Wexler, K. (1996). Subject case licensing and English root infinitives. *Proceedings of the 20th Annual Boston …*, *1995*, 670–681. http://linguistics.ucla.edu/people/cschutze/BU20.pdf

Seal, B. C., & De Paolis, R. A. (2014). Manual activity and onset of first words in babies exposed and not exposed to baby signing. *Sign Language Studies*, *14*(4), 444–465. https://doi.org/10.1353/sls.2014.0015

Sekhon, M., Cartwright, M., & Francis, J. J. (2017a). Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework. *BMC Health Services Research*, *17*(1), 1–13. https://doi.org/10.1186/s12913-017-2031-8

Sekhon, M., Cartwright, M., & Francis, J. J. (2017b). Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework. *BMC Health Services Research*, *17*(1), 1–13. https://doi.org/10.1186/s12913-017-2031-8

Diehl, S.F. (1999). Listen and learn? A software review of EAROBICS. *Language, Speech & Hearing Services in Schools*.

Shafer, V. L., Ponton, C., Datta, H., Morr, M. L., & Schwartz, R. G. (2007). Neurophysiological indices of attention to speech in children with specific language impairment. *Clinical Neurophysiology*. https://doi.org/10.1016/j.clinph.2007.02.023

Sheehan, J., Girolametto, L., Reilly, S., Ukoumunne, O., Price, A., Gold, L., Weitzman, E., & Wake, M. (2009). Feasibility of a language promotion program for toddlers at risk. *Early Childhood Services*.

Sheehy, K, & Rolph, S. (2004). BERA Perspectives on Makaton. *Paper Presented at British Educational Research Association Annual Conference 2004*.

Sheehy, K., & Duffy, H. (2009). Attitudes to makaton in the ages of integration and inclusion. *International Journal of Special Education*, *24*(2), 91–102.

Shimpi, P. M., & Huttenlocher, J. (2007). Redirective labels and early vocabulary development. *Journal of Child Language*, *34*(4), 845–859. https://search.proquest.com/docview/69050098?accountid=12753

Shonkoff, J. P., & Phillips, D. A. (2000). From Neurons to Neighborhoods. In *From Neurons to Neighborhoods*. National Academies Press. https://doi.org/10.17226/9824

Sisson, C. B. (2008). A meta-analytic investigation into the efficacy of Fast ForWord intervention on improving academic performance. In *ProQuest Dissertations and Theses*.

Smith-Lock, K., Leitão, S., Lambert, L., Prior, P., Dunn, A., Cronje, J., Newhouse, S., & Nickels, L. (2013). Daily or weekly? The role of treatment frequency in the effectiveness of grammar treatment for children with specific language impairment. *International Journal of Speech-Language Pathology*. https://doi.org/10.3109/17549507.2013.777851

Snowling, M. J., Bishop, D. V. M., Stothard, S. E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(8), 759–765. https://doi.org/10.1111/j.1469-7610.2006.01631.x

Snowling, M. J., Adams, J. W., Bishop, D. V. M., & Stothard, S. E. (2001). Educational attainments of school leavers with a preschool history of speech-language impairments. *International Journal of Language and Communication Disorders*, *36*(2), 173–183. https://doi.org/10.1080/13682820010019892

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019). Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds. *Child Development*, *90*(4), 1303–1318. https://doi.org/10.1111/cdev.13072

Spooner, L. (2002). Addressing expressive language disorder in children who also have severe receptive language disorder: A psycholinguistic approach. *Child Language Teaching and Therapy*, *18*(3), 289–313. https://doi.org/10.1191/0265659002ct239oa

Standbridge, L. (2016). Does initial severity of expressive language delay/disorder in pre-school children predict expressive language outcomes following “building early sentences therapy”? A pilot and feasibility study. *Unpublished Undergraduate Dissertation*.

Stow, C., & Pert, S. (2015). SLT Assessment and Intervention: Best practice for children and young people in bilingual settings. *Royal College of Speech and Language Therapists*, 1–28. https://www.rcslt.org/members/publications/publications2/slt_assessment_and_intervention_main

Stringer, H. (2010a). Changing Practice. *RCSLT Bulletin*, *673*, 28.

Stringer, H. (2010b). Evaluating the Written Evidence. *RCSLT Bulletin*, *672*, 22.

Stringer, H. (2010c). Evidence Based Practice in Action. *RCSLT Bulletin*, *674*, 20–21.

Stringer, H. (2010d). Formulating Evidence Based Practice Questions. *RCSLT Bulletin*, *670*, 22.

Stringer, H. (2010e). From the Research Strategy to Practice. *RCSLT Bulletin*, *698*, 22.

Stringer, H. (2010f). Searching for Evidence. *RCSLT Bulletin*, *671*, 22.

Stringer, H. (2010g). What is Evidence Based Practice? *RCSLT Bulletin*, *699*, 22.

Sturgis, P., Smith, P., & Hughes, G. (2006). *A Study of Suitable Methods for Raising Response Rates in School Surveys*. http://dera.ioe.ac.uk/6378/1/RR721.pdf

Sukenik, N., & Friedmann, N. (2018). ASD is not DLI: Individuals with autism and individuals with syntactic DLI show similar performance level in syntactic tasks, but different error patterns. *Frontiers in Psychology*, *9*(APR). https://doi.org/10.3389/fpsyg.2018.00279

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*(2), 182–198. https://doi.org/10.1016/0093-934X(80)90139-X

Tallal, P. (1990). Fine-grained discrimination deficits in language-learning impaired children are specific neither to the auditory modality nor to speech perception. *J. Speech Hear. Res*.

Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews. Neuroscience*, *5*(9), 721–728. https://doi.org/10.1038/nrn1499

Tallal, P., & Piercy, M. (1973). Developmental aphasia: impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, *11*(4), 389–398. https://doi.org/10.1016/0028-3932(73)90025-0

Tallal, P. (2013). Fast forWord®: The birth of the neurocognitive training revolution. In *Progress in Brain Research*. https://doi.org/10.1016/B978-0-444-63327-9.00006-0

Thomas-Stonell, N Robertson, B., Walker, J., Oddson, B., & Washington, K Rosenbaum, P. (2012). *FOCUS©: Focus on the Outcomes of Communication Under Six*. Holland Bloorview Kids Rehabilitation Hospital.

Thomas-Stonell, N. L., Oddson, B., Robertson, B., & Rosenbaum, P. L. (2010). Development of the FOCUS (Focus on the Outcomes of Communication Under Six), a communication outcome measure for preschool children. *Developmental Medicine and Child Neurology*, *52*(1), 47–53. https://doi.org/10.1111/j.1469-8749.2009.03410.x

Thomas, M., & Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *The Behavioral and Brain Sciences*, *25*(6), 727–750.

Thordardottir, E. T., & Weismer, S. E. (2002). Verb argument structure weakness in specific language impairment in relation to age and utterance length. *Clinical Linguistics & Phonetics*, *16*(4), 233–250.

Ting, J. Y., Bergeson, T. R., & Miyamoto, R. T. (2012). Effects of Simultaneous Speech and Sign on Infants' Attention to Spoken Language. *Laryngoscope*, *122*(12), 2808–2812. https://doi.org/10.1038/jid.2014.371

Tomasello, M. (2003). Constructing a language. In *A usagebased theory of language acquisition*. https://doi.org/10.1017/CBO9781107415324.004

Tomasello, M. (2009). The usage-based theory of language acquisition. *Cambridge Handbook of Child Language*, 69–88. https://doi.org/10.1017/CBO9780511576164.005

Tomasello, M., & Farrar, M. J. (1986). Joint Attention and Early Language. *Child Development*, *57*(6), 1454–1463.

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, *4*(12), 197–211. https://doi.org/10.1177/014272378300401202

Tomblin, J. B., Zhang, X., Buckwalter, P., & O'Brien, M. (2003). The Stability of Primary Language Disorder. *Journal of Speech Language and Hearing Research*, *46*(6), 1283. https://doi.org/10.1044/1092-4388(2003/100)

Tomblin, J. Bruce. (2008). Validating diagnostic standards for specific language impairment using adolescent outcomes. In *Understanding Developmental Language Disorders: From Theory to*

Practice. https://doi.org/10.4324/9780203882580

Tomblin, J., Records, N., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear Res*, *40*(6), 1245–1260.

Tomblin, J Bruce, Records, N. L., & Zhang, X. (1996). A System for the Diagnosis of Specific Language Impairment in Kindergarten Children. *Journal of Speech Language and Hearing Research*, *39*(December), 1284–1294.

Topham, E. (2016). A randomised, controlled pilot study of Building Early Sentences Therapy (BEST) for preschool children with language impairment and their social activity and participation. *Unpublished Undergraduate Dissertation*.

Tsiouri, L., & Greer, R. D. (2003). Inducing Vocal Verbal Behavior in Children with Severe Language Delays Through Rapid Motor Imitation Responding. *Journal of Behavioral Education*. https://doi.org/10.1023/A:1025508311022

Tyler, A. A., Lewis, K. E., Haskill, A., & Tolbert, L. C. (2006). Efficacy and Cross-Domain Effects of a Morphosyntax and a Phonology Intervention. *Language, Speech, and Hearing Services in Schools*. https://doi.org/10.1044/0161-1461(2002/005)

Ullman, M. T., & Pierpont, E. I. (2005). Specific Language Impairment Is Not Specific To Language: The Procedural Deficit Hypothesis. *Cortex*, *41*, 399–433.

Ullman, M. T. (2001a). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, *2*(10), 717–726. https://doi.org/10.1038/35094573

Ullman, M. T. (2001b). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, *30*(1), 37–69. https://doi.org/10.1023/A:1005204207369

van Berkel-van Hoof, L., Hermans, D., Knoors, H., & Verhoeven, L. (2019). Effects of Signs on Word Learning by Children With Developmental Language Disorder. *Journal of Speech, Language, and*

Hearing Research, *62*(6), 1798–1812.

Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, *59*(9), 920–925. https://doi.org/10.1016/j.jclinepi.2006.02.007

van der Lely, H. K. J. (1998). SLI in Children: Movement, Economy, and Deficits in the Computational-Syntactic System. *Language Acquisition*, *7*(2–4), 161–192. https://doi.org/10.1207/s15327817la0702-4_4

van der Lely, H. K. J. (2004). Evidence for and implications of a domain- specific grammatical deficit. In L. Jenkins (Ed.), *The genetics of language* (pp. 117–145). Elsevier.

van der Lely, H. K. J. (2005). Domain-specific cognitive systems: Insight from Grammatical-SLI. *Trends in Cognitive Sciences*, *9*(2), 53–59. https://doi.org/10.1016/j.tics.2004.12.002

van der Lely, H. K. J., & Stollwerck, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition*, *62*(3), 245–290. https://doi.org/10.1016/S0010-0277(96)00783-4

Van Horne, A. J. O., Curran, M., Larson, C., & Fey, M. E. (2018). Effects of a complexity-based approach on generalization of past tense –ed and related morphemes. *Language, Speech, and Hearing Services in Schools*, *49*(3S), 681–693. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0142

Van Horne, A. J. O., Fey, M., & Curran, M. (2017). Do the hard things first: A randomized controlled trial testing the effects of exemplar selection on generalization following therapy for grammatical morphology. *Journal of Speech, Language, and Hearing Research*, *60*(9), 2569–2588. https://doi.org/10.1044/2017_JSLHR-L-17-0001

Vogt, S. S., & Kauschke, C. (2017a). Observing iconic gestures enhances word learning in typically developing children and children with specific language impairment. In *Journal of Child Language* (Vol. 44, Issue 6, pp. 1458–1484). Cambridge University Press.

https://doi.org/10.1017/S0305000916000647

Vogt, S. S., & Kauschke, C. (2017b). With some help from others' hands: Iconic gesture helps semantic learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3213–3225. https://doi.org/10.1044/2017_JSLHR-L-17-0004

Volpato, D., Orton, D., & Blackburn, D. (1986). Making progress with Makaton. *Nursing Times*, *82*(18), 33–35.

Wagner, E. F., Tubman, J. G., & Gil, A. G. (2004). Implementing school-based substance abuse interventions: Methodological dilemmas and recommended solutions. *Addiction*, *99*(SUPPL. 2), 106–119. https://doi.org/10.1111/j.1360-0443.2004.00858.x

Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, *57*, 209–232. https://doi.org/10.1016/j.specom.2013.09.008

Wake, M., Tobin, S., Girolametto, L., Ukoumunne, O. C., Gold, L., Levickis, P., Sheehan, J., Goldfeld, S., & Reilly, S. (2011). Outcomes of population based language promotion for slow to talk toddlers at ages 2 and 3 years: Let's learn language cluster randomised controlled trial. *BMJ (Online)*. https://doi.org/10.1136/bmj.d4741

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

Washington, K., Thomas-Stonell, N., Oddson, B., Mcleod, S., Warr-Leeper, G., Robertson, B., & Rosenbaum, P. (2013). Construct validity of the FOCUS©(Focus on the Outcomes of Communication Under Six): A communicative participation outcome measure for preschool children. *Child: Care, Health and Development*. https://doi.org/10.1111/cch.12043

Westerveld, M. F., & Gillon, G. T. (2008). Oral narrative intervention for children with mixed reading disability. *Child Language Teaching and Therapy*, *24*, 31–54.

https://doi.org/10.1177/0265659007084567

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot & N. Hornstein (Eds.), *Verb Movement* (pp. 305–350). Cambridge University Press. https://books.google.com/books?hl=en&lr=&id=TQe85SBpNDoC&pgis=1

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, *106*(1–4), 23–79. https://doi.org/10.1016/S0024-3841(98)00029-1

Wexler, K. (2003). Lenneberg's dream: Learning, normal language development, and specific language impairment. *Language Competence across Populations: Toward a Definition of Specific Language Impairment.*

What Works Clearinghouse. (2007). *Standards Handbook*. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

White, L. S. (2012). Conducting School-based Research: Lessons Learned From an After-school Intervention Study. *Journal of Pediatric Health Care*, *26*(1), 75–79. https://doi.org/10.1016/j.pedhc.2011.05.001

White, S., & Spencer, S. (2018). A school-commissioned model of speech and language therapy. *Child Language Teaching and Therapy*, *34*(2), 141–153. https://doi.org/10.1177/0265659018780961

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. https://ggplot2.tidyverse.org

Wiig, E. H., Secord, W., & Semel, E. (1992). *Clinical evaluation of language (CELF Preschool)*. The Psychological Corporation. Harcourt Brace Jovanvich, Inc.

Willinger, U., Brunner, E., Diendorfer-Radner, G., Sams, J., Sirsch, U., & Eisenwort, B. (2003). Behaviour in children with language development disorders. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie*.

Wolfe, D. L., & Heilmann, J. (2010). Simplified and expanded input in a focused stimulation program for a child with expressive language delay (ELD). *Child Language Teaching and Therapy*. https://doi.org/10.1177/0265659010369286

World Health Organization. (2001). International Classification of Functioning, Disability and Health. *International Classification of Functioning, Disability and Health.* https://doi.org/10.1097/01.pep.0000245823.21888.71

World Health Organization. (2007). International Classification of Functioning, Disability and Health-Children and Youth Version. In *Journal of Chemical Information and Modeling*. https://doi.org/10.1017/CBO9781107415324.004

Yang, C., Crain, S., Berwick, R. C., Chomsky, N., & Bolhuis, J. J. (2017). The growth of language: Universal Grammar, experience, and principles of computation. In *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2016.12.023

Yoder, P., Camarata, S., & Gardner, E. (2005). Treatment effects on speech intelligibility and length of utterance in children with specific language and intelligibility impairments. *Journal of Early Intervention*. https://doi.org/10.1177/105381510502800105

Zambrana, I. M., Pons, F., Eadie, P., & Ystrom, E. (2014). Trajectories of language delay from age 3 to 5: Persistence, recovery and late onset. *International Journal of Language and Communication Disorders*, *49*(3), 304–316. https://doi.org/10.1111/1460-6984.12073

Zentall, S. R., & Morris, B. J. (2010). " Good job, you're so smart": The effects of inconsistency of praise type on young children's motivation. *Journal of Experimental Child Psychology*, *107*(2), 155–163. https://doi.org/10.1016/j.jecp.2010.04.015

Zipoli, R. P., & Kennedy, M. (2005). Evidence-based practice among speech-language pathologists: Attitudes, utilization, and barriers. *American Journal of Speech-Language Pathology*, *14*(3), 208–220. https://doi.org/10.1044/1058-0360(2005/021)

Zubrick, S. R., Taylor, C. L., & Christensen, D. (2015). Patterns and predictors of language and literacy abilities 4-10 years in the longitudinal study of Australian children. *PLoS ONE*, *10*(9), 1–29.

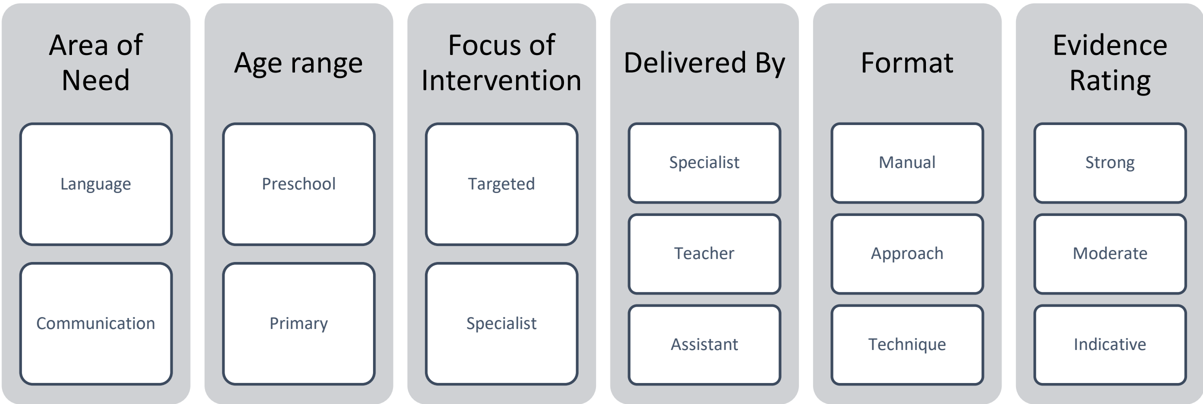
https://doi.org/10.1371/journal.pone.0135612

Appendix

Appendix 1 Search Criteria for What Works Database Review

Review of the What Works Database according to the following search terms:

Search terms applied to the What Works database to identify interventions for children with severe language difficulties in the early years



Appendix 2 Interventions resulting from the What Works Database search (also see Appendix 1) that were examined in further detail

Intervention	Relevance to BEST	Description	Reference
Active listening for active learning	1	A targeted/specialist approach addressing communication in pre-school and primary years, delivered by specialists or teachers. Supports children to identify and clarify when they do not understand something. Increases meta-awareness of the child's own language through comprehension monitoring and demonstrates that children need not feel responsible for misunderstandings. Children are taught principles of active listening and are then exposed to communication that is impossible to follow for varying reasons. Children are shown strategies to use when they do not understand. Supported by moderate evidence.	(Dollaghan & Campbell, 2009)
Becky Shanks Narrative intervention	1 Addresses storytelling principles, although the structure and build up over a number of weeks is similar to BEST, different aspects are targeted.	A targeted/specialist manualised intervention addressing language in pre-school and primary years, delivered by specialists, teachers, and assistants. The intervention focuses on the principles of storytelling, including narrative, grammar, and structure of stories. Children are told stories supported by visual aids, and explicit teaching takes place over the intervention sessions so that children understand storytelling principles and can apply them to their own storytelling. Supported by indicative evidence.	(Davies et al., 2004)
Broad Target Recast (BTR)	1 Unstructured compared to BEST; although the repetition of the correct sentences during BEST could be likened to recasting.	A targeted/specialist technique addressing language in pre-school years, delivered by specialists. Based on the recasting technique where an adult rephrases or corrects the child's utterance to demonstrate the correct grammar and vocabulary without distorting the child's meaning. During free, child-led conversation, the specialist recasts both speech and grammatical errors, and may provide additional information to increase the child's sentence length. BTR is supported by moderate evidence (Yoder et al., 2005), and there is also evidence for recasting generally (e.g. Saxton 2005)	(Saxton, 2005; Yoder et al., 2005)
Colourful Semantics	2	A specialist approach addressing language in primary years, delivered by specialists. Designed to support grammatical structures (A. Bryan, 2008), and later developed to support children with	(Bolderson et al., 2011; Bryan, 2008)

	Addresses similar targets; draws attention to underlying structures using a referent. Need to look more closely at this	vocabulary, comprehension and structure. Coloured cards correspond to aspects of sentence structure, allowing children to link syntax with meaning. Originally developed as a one to one intervention but now also used in mainstream settings (A. Bryan, 2008). Further developed into the shape coding intervention (S. H. Ebbels et al., 2007). Supported by indicative evidence.	
Cueing word-finding	1 Not relevant	A specialist approach addressing language in the primary years, delivered by specialists. A cueing aid for word-finding difficulties. Delivered one to one, in schools or specialist units. Supported by indicative evidence.	(Best, 2005)
Earobics	1	A specialist manualised intervention addressing language in pre-school and primary years, delivered by specialists. A computerised phonological awareness and language processing intervention. Children must match sounds in a variety of conditions including with and without background noise. Aims to strengthen a number of speech, language and literacy skills. Supported by moderate evidence.	(Cognitive Concepts, 1997; Diehl, 2014)
Fast ForWord	1	A specialist manualised intervention addressing language in the primary years, delivered by specialists. A reading and oral language computerised intervention based on claims that deficits in auditory temporal processing underpin language impairment (Tallal & Piercy, 1973). Controversial and generally rejected as having little efficacy (Sisson, 2008). Supported by strong evidence.	(Tallal, 2013)
Focused Stimulation	2 Similar to the principles of BEST input although is less structured. Look in more detail.	A targeted/specialist technique addressing language in pre-school years, delivered by specialists, teachers, and assistants. The technique aims to draw attention to particular vocabulary (Girolametto & Pearce, 1996) and grammatical elements. The experienced speaker uses the same construction repeatedly in sentences, varying its position, to highlight its use, stress and positioning. Supported by moderate evidence.	(Girolametto et al., 1996; Wolfe & Heilmann, 2010)
Hanen It Takes Two To Talk	1 Parent-child interaction	A specialist manualised intervention and approach addressing language and communication in the pre-school years, delivered by specialists to parents. The program teaches parents to incorporate language learning into their daily interactions with their child. It promotes turn taking, the child's motivations, confidence building and book reading. Parental training is	(Girolametto & Weitzman, 2006)

		delivered in small groups over 6-8 sessions and home visits from a Speech and Language Therapist where reflection on videoed parent-child interaction takes place. Supported by moderate evidence.	
Hanen More Than Words	1 Parent-child interaction and ASD	A specialist manualised intervention addressing language and communication in the pre-school years, delivered by specialists. Supports parents with children on the Autism Spectrum to support their communication and language. Trains parents to promote positive behaviours, understand what motivates their child to learn and increase length and meaningfulness of interactions. Parental training is delivered in small groups over 6-8 sessions and home visits from a Speech and Language Therapist where reflection on videoed parent-child interaction takes place. Supported by indicative evidence.	(Carter et al., 2011; Girolametto, Sussman, & Weitzman, 2007; McConachie et al., 2005)
Intensive Interaction	1	A specialist technique addressing communication in pre-school and primary years, delivered by specialists, teachers, and assistants. Based on caregiver interactions with babies, adults are encouraged to interact with children in a highly responsive way and to 'mirror' their behaviours and communication. Aims to teach children the benefits of social interaction. Supported by indicative evidence.	(Nind & Hewett, 2006)
Let's Learn Language	1 PCI	A universal/targeted manualised intervention addressing language in the pre-school years, delivered by assistants. The program is similar to the Hanen 'You Make The Difference' program and aims to reduce language delay through increased modelling language and child-led interactions. Parental training focuses on using the child's interests, improving quality, length and content of interactions and implementing these in everyday interactions, play and book reading. Supported by moderate evidence.	(Sheehan et al., 2009; Wake et al., 2011)
Lexicon Pirate	1 Lexical difficulties	A specialist manualised intervention and approach addressing speech and language in the pre-school and primary years, delivered by specialists. Developed for German bilingual children but may translate to English speaking children. Supports children with a variety of lexical difficulties. A puppet explores and models learning strategies such as asking for new word names, comparing semantic features of words to aid categorisation, rehearsal and repeating to aid encoding and retrieval. Supported by moderate evidence.	(Motsch & Marks, 2015; Motsch & Ulrich, 2012)

Little Talkers (Parent-based Intervention)	1	A specialist manualised intervention and approach addressing language in the pre-school years, delivered by specialists. Indirect approach by parents to improve child's expressive language from single word level to 3-4 word utterances. SLTs set parents fortnightly language goals and oversee practice activities. Parents implement strategies to achieve goals at home based on child's interests and individual daily routines. Supported by indicative evidence.	(Gibbard, 1994)
Makaton	2 Used in BEST but will discuss in use of sign section of introduction	A targeted approach addressing communication in pre-school and primary years, delivered by specialists, teachers, and assistants. Signing system for children (and adults) with learning difficulties but also applied in mainstream settings. Signing can be used as the main method of communication or to support oral language. Focus on teaching a small core vocabulary, prioritising teaching of the vocabulary based on individual circumstances and combining signing with other methods of communication. Accredited tutors teach adults who interact with the child, and it is implemented as appropriate. Database and resources are available. Can be used alongside other interventions (Bickford-Smith et al., 2005; McKean et al., 2013) Supported by moderate evidence.	(Bickford-Smith et al., 2005; Grove & Walker, 1990; Rubina Lal, 2010; Poncelas & Murphy, 2007)
Milieu Teaching Therapy	1	A targeted approach addressing language in the pre-school and primary years, delivered by specialists, teachers, and assistants. A naturalised intervention in which the treatment environment is planned to facilitate everyday interactions to improve communication and social skills through the milieu (life space). Staff provide rich interactive feedback and create new communicative situations for children to engage with. Integrates modelling, questioning and time delay to increase child's spontaneous utterances. Supported by strong evidence.	(Mancil, 2009)
Morpho-Syntactic Intervention	2 Contains rich input demonstrating target constructions, highlighting morphological endings etc (as in BEST) aims to elicit targets multiple times	A targeted approach addressing language in pre-school and primary years, delivered by specialists. Auditory awareness, focused stimulation and elicited production activities designed to promote syntax and morphology. The therapist reduces the level of support overt the three cycles. Promotes use of target constructions and therapist recasts them. (May also have positive effect on phonology (Tyler et al., 2006)). Generally overlaps other interventions. Supported by moderate evidence.	(S. M. Camarata et al., 1994; Cleave & Fey, 1997; Marc E. Fey et al., 1993)

Naturalistic Speech Intelligibility Training	1 Focus on speech sound disorder, naturalistic intervention and research focuses on specific disorders e.g. Down Syndrome	A specialist approach addressing speech and language in pre-school and primary years, delivered by specialists. Developed mainly for children with speech sound difficulties but also used with young children to target speech intelligibility and accuracy. The approach stipulates that intelligibility is determined by correct syllable structure and syntactic features as well as correct phoneme production. The environment is manipulated to promote therapist-child interaction which contains the target phonemes and grammatical elements. Supported by moderate evidence.	(S. M. Camarata et al., 1994)
Nuffield Early Language Intervention	2 Target population similar. Expressive language and grammatical competence targeted (Focus on oral language rather than phonology and reading)	A specialist manualised intervention addressing language in the pre-school and primary years, delivered by assistants. Targets multiple facets of oral language including vocabulary, listening, phoneme awareness, grammatical competence, and narrative skills and based around the interventions tested by Bowyer-Crane et al. (2008). Teaching assistants and teachers deliver the intervention following training delivered by an SLT. Can be delivered as a 10, 20, or 30 week intervention. Supported by moderate evidence.	Based on (Bowyer-Crane et al., 2008) Intervention described in (Fricke et al., 2013) RCT (Fricke et al., 2017)
Oral Language Programme	2 Elements may align with BEST focus, although this needs further investigation	A targeted manualised intervention addressing language in the primary years, delivered by assistants. Activities adapted from a number of sources including Time to Talk (Schroeder, 2001) Black Sheep Press materials (Rippon, 2002), and Rhodes to Language (Rhodes, 2001). Oral Language (OL) intervention from Bowyer-Crane et al. (2008) including vocabulary development, expressive language, listening and inferencing skills. Narrative work and independent speaking, with TAs modelling and recasting where appropriate. Supported by moderate evidence.	(Bowyer-Crane et al., 2008)
Phonology with reading programme (P+R)	1 Phonology and reading aspects aren't relevant to BEST	A targeted manualised intervention addressing language in the primary years, delivered by specialists and teachers. Phonology and Reading programme (P&R) intervention from Bowyer-Crane et al. (2008). Teaching assistants deliver three components: Letter sound knowledge based on Jolly Phonics (Lloyd, 1998), Phonological awareness via multi-sensory techniques, and book reading where children are encouraged to link their phonics and phonological knowledge for reading. Supported by moderate evidence.	(Bowyer-Crane et al., 2008)

Picture Exchange Communication System (PECS)	1 Supports functional communication but otherwise not relevant. Focus on children with ASD.	A specialist manualised intervention addressing communication in pre-school and primary years, delivered by specialists, teachers, and assistants. Children learn to exchange a picture of an item with a communication partner, who will then give them that item so they can get what they need and engage in communicative interactions with others. Originally designed for children with Autism. Children build up communication with intention cards (e.g. 'I want'), responding to questions and producing spontaneous comments. Supports functional communication. Supported by moderate evidence.	(S. L. Hart & Banda, 2010; Preston & Carter, 2009)
Pre-School Autism Communication Therapy (PACT)	1 ASD	A universal/targeted manualised intervention addressing language and communication in pre-school and primary years, delivered by specialists. For children with Autism or social communication disorders, aiming to increase their initiation of social communication and use of language. Parents are videoed interacting with their child and reflect on the recordings with an SLT. Goals are set based on the manual which suggests the skills to be targeted, and parental desires are incorporated. Parents are then encouraged to use positive interaction with their child. Supported by moderate evidence.	(Aldred et al., 2004; Green et al., 2010)
Pre-Teaching Vocabulary (PTV)	1	A universal, targeted and specialised manualised intervention and approach addressing language in the primary years, delivered by teachers and assistants. Supports and scaffolds teachers' discussion of new words through structured steps to learning vocabulary. Uses pictures and naming, question and knowledge games and songs to pre-teach vocabulary to the child or group before teaching of the material proper. Based on multiple works including Lubliner and Smetana (2005) and Parsons, Law, and Gascoigne (2005) and the principles discussed in Bowyer-Crane et al. (2008). Supported by indicative evidence.	(Bowyer-Crane et al., 2008; Lubliner & Smetana, 2005; Parsons et al., 2005) (References not available for some of the works discussed on the What Works database)
Psycholinguistic Framework	1 Speech	A targeted/specialist approach addressing speech and language in pre-school and primary years, delivered by specialists. A framework which allows SLTs to build up a picture of children's speech, including the nature of their difficulties, so that these can be targeted through therapy targets for speech and processing. Supported by moderate evidence.	(Pascoe et al., 2005)

Rapid Motor Imitation Antecedent Training (RMIA)	1 ASD	A specialist manualised intervention addressing language in the pre-school years, delivered by specialists. Designed to elicit first words from non or minimally verbal children with Autism Spectrum Disorder. Applies the principle of Behavioural Momentum (Nevin et al., 1983) to children, where asking child to complete easy tasks leads to compliance with more difficult or demanding tasks. RMIA builds upon easy motor tasks performed in succession before introducing vocalisations. Supported by indicative evidence.	(Nevin et al., 1983; Tsiouri & Greer, 2003)
Reading and Language Intervention for Children with Down Syndrome (RLI)	1 Down Syndrome & focus on reading and vocabulary learning.	A targeted manualised intervention addressing language in the primary years, delivered by assistants. Supports children with Down Syndrome with vocabulary development and use in oral and written forms. Reading strand and a language strand focusing on vocabulary learning and use. Supported by moderate evidence.	(Burgoyne et al., 2012)
Shape Coding by Susan Ebbels	2 Relevant	A targeted/specialist manualised intervention addressing language in the primary and secondary years, delivered by specialists, teachers, and assistants. Drawing on work by Bryan (2008) and Lea (1965), shape coding was developed to support children with language difficulties to understand sentence structure by linking semantics with syntax through visual representation of sentence morphology. Colours represent parts of speech, arrows represent tense and aspect, and shapes represent syntactic and argument structures (Ebbels, 2007). Shape coding is a flexible tool that is withdrawn when it is no longer required, rather than after a set number of sessions. Supported by moderate evidence.	(Ebbels, 2007)
Social Communication Intervention Programme	1 Children 6 years +, greater focus on improving functional and social communication rather than underlying structures as in BEST	A targeted/specialist manualised intervention addressing language in the primary years, delivered by specialists and assistants. Aims to support language use through semantics, inferencing, narratives and verbal reasoning (collectively Language Processing). Aims to support communication including functional communication and social understanding and social interaction (SUSI) e.g. understanding emotions and routines. Goals are set and sessions are delivered. Caregivers are involved in reviewing and setting goals. Intervention follows a set framework but is tailored for each child to support their individual needs). Supported by moderate evidence.	(Adams, Lockton, Freed, et al., 2012; Adams, Lockton, Gaile, et al., 2012)

Social Stories	1 pragmatics, tailored towards pragmatic difficulties/autism	A targeted/specialist approach addressing communication in pre-school, primary and secondary years, delivered by specialists, teachers, and assistants. Aims to support children with ASD or pragmatic difficulties understand different people's perspectives, events and appropriate communication through sung, written and computer based social stories. The stories are tailored to the individual child's social communication and behaviour, and therefore target this by presenting appropriate behaviours. The child is then asked questions based on the story and their reactions. Supported by moderate evidence.	(Gray & Garand, 1993; Kuoch & Mirenda, 2003; Reynhout & Carter, 2006)
Social Thinking	1	A targeted/specialist approach addressing communication in pre-school and primary years, delivered by specialists, teachers, and assistants. SLTs tailor intervention based on ILAUGH model to support children with social interaction difficulties. ILAUGH: Initiation of language; Listening with eyes and brain; Abstract and inferential language/communication; Understanding perspective; Getting the big picture; Humour, and Human relatedness Supported by indicative evidence.	(Crooke et al., 2008)
Social use of language programme	1	A targeted/specialist manualised intervention addressing communication the primary and secondary years, delivered by specialists. In three parts the intervention teaches and supports communication skills, applying them in challenging real-life situations and metacognition. Mainly applied to older children but can also support children with ASD and younger children. Supported by indicative evidence.	(Owens et al., 2008)
Story Grammar Intervention	1	A specialist manualised intervention addressing language in the primary years, delivered by specialists. Children are introduced to stories with strong structures, and supported to talk about what makes a good story. Children are encouraged to think more deeply about the story and use clues to make inferences. Structural elements including grammar, problems, resolution and endings are introduced systematically over the therapy sessions. Visual techniques such as story maps and labels are used to support understanding of the story structure and meaning. Supported by indicative evidence.	(Westerveld & Gillon, 2008)
Strathclyde Language	2 Uses some elements e.g. colourful	A specialist manualised intervention addressing language in the primary years, delivered by specialists and assistants. Incorporates comprehension monitoring, vocabulary learning and	(J. Boyle et al., 2007b; Elspeth

Intervention Programme	semantics that are similar to BEST.	comprehension, grammar (using a colourful semantics approach) and narrative comprehension. Supported by moderate evidence.	McCartney et al., 2010)
Task Boost KS1	2 Commonly used in BEST schools, targets sentence building and other elements possibly relevant to BEST	A targeted manualised intervention addressing language in the pre-school and primary years, delivered by teachers and assistants. Strengthens teaching staff understanding of the importance of SLCN and incorporating positive strategies into whole class teaching practice. Identifies children with language difficulties, seeks to close the language gap between children with and without delays. Focuses on Conversations, narratives, sentence building, vocabulary and listening skills. Supported by moderate evidence.	(Lee & Pring, 2016)
TeaCHH	1 ASD, SLCN not a high priority outcome	A specialist approach addressing communication across all years, delivered by specialists, teachers, and assistants. Aims to promote communication and other related skills for Autistic children through organisation, communication supports, time visualisation and repeated interventions and techniques such as rewards following the completion of work. Supported by indicative evidence.	(Eikeseth, 2009)
Visual Approaches to Support Speech and Language	2 SIGN but will discuss this in the sign section of the introduction as it is more relevant there	A universal/targeted/specialist approach addressing speech and language in pre-school and primary years, delivered by specialists, teachers, and assistants. An approach aimed at providing additional visual cues for language learning support. Can support different areas of language such as word order or grammatical elements. Can include shape coding/colourful semantics type visual representation, signing, use of objects or referents. Signing can be used to visually symbolise and represent speech sounds or elements of language. Can also include strategies such as visual timetables and other environmental stimuli. Much of the current evidence is for Autistic children. Supported by indicative evidence.	(Archibald & Gathercole, 2006; Lal & Bali, 2007)
Visualising and Verbalising	1	A targeted/specialist manualised intervention addressing language in all years, delivered by specialists, teachers, and assistants. Aims to support reportedly weak imagery in language impaired children (Bell, 1991) by teaching the child to describe, and visualise pictures and sentences, and then answer increasingly complex questions using visual cues to support answers. Moves on to inferential questions where the child must evaluate and infer from the content provided. Supported by moderate evidence.	(Bell, 1991)

MRC Pilot study definition:
"A pilot study need not be a 'scale model' of the planned mainstage evaluation, but should address the main uncertainties that have been identified in the development work. Pilot study results should be interpreted cautiously when making assumptions about the required sample size, likely response rates, etc., when the evaluation is scaled up" (Craig et al., 2008, pp. 10).

NIHR Pilot Study Definition:

"A version of the main study that is run in miniature to test whether the components of the main study can all work together. It is focused on the processes of the main study, for example to ensure recruitment, randomisation, treatment, and follow-up assessments all run smoothly. It will therefore resemble the main study in many respects, including an assessment of the primary outcome. In some cases this will be the first phase of the substantive study and data from the pilot phase may contribute to the final analysis; this can be referred to as an internal pilot. Or at the end of the pilot study the data may be analysed and set aside, a so-called external pilot"

NIHR definition of feasibility studies
"Pieces of research done before a main study in order to answer the question "Can this study be done?". They are used to estimate important parameters that are needed to design the main study. The design of a feasibility study generally involves listing those parameters which are uncertain and describing the methods for improving their precision so that the main study will have a better chance of success. Examples of such parameters include:

- standard deviation of the outcome measure, which is needed in some cases to estimate sample size
- willingness of participants to be randomised
- willingness of clinicians to recruit participants
- number of eligible patients; carers or other appropriate participants
- characteristics of the proposed outcome measure and in some cases feasibility

  studies might involve designing a suitable outcome measure
- follow-up rates, response rates to questionnaires, adherence/compliance rates, ICCs in cluster trials, etc.
- availability of data needed or the usefulness and limitations of a particular database; and
- time needed to collect and analyse data.

Feasibility studies for randomised controlled trials may not themselves be randomised. Crucially, feasibility studies do not evaluate the outcome of interest; that is left to the main study.

Feasibility studies do not evaluate the outcome of interest; that is left to the main study. Feasibility studies for randomised controlled trials may not themselves be randomised. If a feasibility study is a small randomised controlled trial, it does not necessarily need to have a primary outcome or power calculations. Instead, the sample size is often used to estimate the critical parameters (e.g. recruitment rate) to the necessary degree of precision"

Definition accessed on 29/05/2019 from https://www.nihr.ac.uk/funding-and-support/documents/funding-for-research-studies/research-programmes/PGfAR/CCF-PGfAR-Feasibility-and-Pilot-studies.pdf

Definitions accessed on 29/05/2019 from https://www.nihr.ac.uk/funding-and-support/documents/funding-for-research-studies/research-programmes/PGfAR/CCF-PGfAR-Feasibility-and-Pilot-studies.pdf

CONSORT 2010 checklist of information to include when reporting a pilot or feasibility randomized trial in a journal or conference abstract

Item	Description	Reported on line number
Title	Identification of study as randomised pilot or feasibility trial	
Authors *	Contact details for the corresponding author	
Trial design	Description of pilot trial design (eg, parallel, cluster)	
Methods		
Participants	Eligibility criteria for participants and the settings where the pilot trial was conducted	
Interventions	Interventions intended for each group	
Objective	Specific objectives of the pilot trial	
Outcome	Prespecified assessment or measurement to address the pilot trial objectives**	
Randomization	How participants were allocated to interventions	
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants screened and randomised to each group for the pilot trial objectives**	
Recruitment	Trial status†	
Numbers analysed	Number of participants analysed in each group for the pilot objectives**	
Outcome	Results for the pilot objectives, including any expressions of uncertainty**	
Harms	Important adverse events or side effects	
Conclusions	General interpretation of the results of pilot trial and their implications for the future definitive trial	
Trial registration	Registration number for pilot trial and name of trial register	
Funding	Source of funding for pilot trial	

Citation: Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 2016;355.

**this item is specific to conference abstracts*

***Space permitting, list all pilot trial objectives and give the results for each. Otherwise, report those that are a priori agreed as the most important to the decision to proceed with the future definitive RCT.*

†For conference abstracts

Evaluating the Effectiveness of Building Early Sentences Therapy

Information Sheet for Parents/Carers

A Research Project by

Ms Anastasia Trebacz
PhD Student- Speech and Language Sciences
Newcastle University

Dr Cristina McKean
Specialist Speech and Language Therapist
Newcastle University

Dr Helen Stringer
Specialist Speech and Language Therapist
Newcastle University

Professor James Law
Professor of Speech and Language Sciences
Newcastle University

With the assistance of Research Associates from Newcastle University School of Education, Communication and Language Sciences

Dear Parent(s)/Carer(s), you and your child are being invited to take part in a research project. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish.

Ask us if there is anything that is not clear or if you would like more information (Anastasia's phone number is 0191 208 6510 and Cristina's is 0191 208 6528). We would like to know whether or not you wish to take part by 21/09/2018. You can give the consent form to Anastasia, who will be visiting the school at various times, or you can return it to your child's class teacher.

What is the purpose of the project?

Lots of children have additional support with learning language, and your child's school provides support in lots of ways for your child's language development. In Newcastle University we have developed a new approach to supporting children's language development called Building Early Sentences Therapy (BEST).

BEST aims to support children to understand and use a range of sentences of increasing length. BEST is a play-based approach where children hear adults use lots of new language whilst playing with toys in a small group of children. A range of clues are used to help children to work out the rules of the sentences such as the adult using gestures, acting out sentences with toys and varying different parts of the sentence.

We want to find out:

- Does BEST improve children's progress in language development more than the usual support approaches used in schools?
- Does the adult using gestures in BEST help children more than when gesture isn't used?
- Do some children benefit more from BEST than others and why might that be?

In your child's school we will be delivering BEST with gestures, and following the progress that children make. We will be comparing their progress with children in other schools who are receiving BEST without gestures, or receiving the usual support offered by schools, so we can find out which approach is most effective.

We are asking for your permission to work with your child to include them in a BEST group and measure their progress over time.

Why have we been chosen?

You have been chosen because your child's class teacher thinks your child may benefit from some additional support for their language development. We will be inviting about 100 children to take part in this research.

What will happen in the research?

If you agree to take part, the study would have 4 steps

Step 1: Checking if your child might benefit from BEST

Your child's language, communication and social and emotional development will be checked using some play-based assessments in school and asking you and your child's teacher to fill in two short questionnaires. If you have any queries about the questionnaires we would be happy to support you to fill them in. The assessments will last about 40 minutes and are fun, play based assessments which involve talking about pictures that your child will be shown and acting out actions with toys. This will happen between 24th September-5th October.

Step 2: Receiving BEST intervention

If they are eligible for the study your child will then receive Building Early Sentences Therapy twice per week for 8 weeks in school.

BEST will be delivered by Anastasia and a familiar learning support assistant in small groups. Your child will hear sentences and see them acted out with toys. Gestures will also be used to represent the sentences. They will then be shown other actions carried out with toys, and given an opportunity to describe what is happening. They are not required to give an answer, and will not be pushed to do so, however we find that most children enjoy joining in and describing the actions. You will be given two easy homework tasks per week, which involve looking at pictures with your child and describing what is happening. Your child will continue to benefit from other support and strategies which your child's school provides and we will work closely with them to find the best way for BEST to fit in with their other approaches.

Stage 3 and 4: Measuring your child's progress

To measure how much progress in language and communication your child makes we will repeat the assessments in Step 1 and one of the questionnaires shortly after BEST has finished around 10th-15th December, and then again 6 – 8 weeks later after the holidays. Research Associates from the research team at Newcastle University will administer all of the assessments.

What else will happen?

During the study we will collect language assessment data from your children. This involves recording their responses to tests of language in booklets. We will also ask you and your child's teacher to fill in some questionnaires about your child's language. We will audio record each assessment so that we can check all of our data is accurate. If your child receives BEST as part of the study, we will also video record the third session so that a member of the research team can check that the therapy is being delivered consistently, increasing the scientific reliability of the study.

Some children will be receiving speech and language therapy, others will not. If your child is receiving Speech and Language Therapy we will ask your permission to let their therapist know your child is in the BEST study, and will agree with them what we should do next. We expect that therapists will take a break from seeing your child while the BEST study is underway. They will then pick up again once our study has finished, but this decision will be made by you and your child's therapist.

We will ask for your telephone number, in case we have any specific queries or concerns relating to your child over the course of the study. We will also ask for your postcode to use in an anonymous analysis of social factors. Both will be stored confidentially and destroyed at the end of the study. There is space to provide this information on the consent form.

Do I have to take part?

It is up to you to decide whether or not your child can take part. If you do decide they can take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect will not affect any of the usual services or support that your child receives.

What if my child doesn't want to take part?

Anastasia, Cristina and the Research Associates are experienced working with children. If you agree for your child to take part, but when they are seen they are unhappy, your child will be treated sympathetically and taken back to their class. We will let you know what happened and discuss with you what to do about the assessment. You can withdraw your child from the project at any time if you would like to.

What are the benefits to us?

Our aim is to find more effective ways to help young children to develop strong language skills. We do not know if children who receive BEST with or without gesture will make different amounts of progress, or if they will make different amounts of progress than children receiving their usual support in schools. The study aims to find out whether using BEST in schools could improve language progress for some children. If we find that BEST is beneficial to the children who receive it, we will offer to train nurseries and schools to deliver the therapy.

What will happen to the results of the research project?

Anastasia and the research team will write up the whole project to share with teachers and other educational specialists. If you would like to receive a summary of the findings please tick the box on the consent form and we will post one to you. We will also inform all schools involved of the results, and you can contact the research team if you would like more information at any point.

What about confidentiality?

We will ensure that your child's results are confidential. None of the information gathered nor the records which need to be kept for the project will have your child's name on.

The questionnaires you complete will remain completely anonymous and will only be used for statistical analysis, not referred to individually.

Your child will be given a participant number to identify them on all records. All of the information we collect about your child will be kept in a locked filing cabinet in a secure room at Newcastle University. Any video and audio recordings will be stored on a Newcastle University encrypted server, and again will be filed using an anonymised code, not your child's name. These will be deleted after the results of the project have been written up in a report. Only people who are working on the project will have access to the records. When we write up the project we will not identify your child or his/her school by name.

Additional information

This project has been given approval by Newcastle University Faculty of Humanities and Social Sciences Ethics Committee. The Head Teacher has also given permission for the research to be carried out in your child's school. If you agree to take part you will be given a copy of this information sheet and of your signed consent form. Please return your signed consent form to your child's class teacher by 21/09/18. All data is handled and stored in accordance with GDPR guidelines.

Your child's school has been chosen to deliver BEST with additional gesture.

We are asking for your permission to work with your child to include them in a BEST group and measure their progress over time.

Your child's safety

All people who work with children have a duty to ensure their safety. Sometimes this means we have to share information with other professionals and parents/carers even though the child would prefer us not to. Please note, if any child involved in the project discloses information indicating an incident of inappropriate behaviour or potential harm to the child, the researchers will follow the established procedures for reporting this information and you will be informed. All of the researchers who will work with your child are DBS checked.

Your right to Withdraw

You do not have to take part in this research. You will not be included in any research without first giving your consent. You also have the right to withdraw your consent at any time and therefore opt out of the study. To do this please contact Anastasia Trebacz (details on the front cover). In May 2019 the data collected for the study will be fully anonymised and published. You won't be able to withdraw your anonymised data after this point, but it will be fully anonymous and all of your identifiable information will be deleted.

If you have any concerns about the project

If you have a concern about any aspect of this project, you can speak to Anastasia or Cristina who will do their best to answer your questions. If you remain unhappy, you have the right to lodge a complaint with the Information Commissioner's Office at Wycliffe House, Water Lane, Wilmslow, SK9 5AF. https://ico.org.uk/

If you want more information please contact Anastasia Trebacz or Cristina McKean

Ms Anastasia Trebacz

PhD Student

School of Education Communication and Language Sciences,

King George VI Building,

Queen Victoria Road,

Newcastle upon Tyne.

NE1 7RU

Telephone: 0191 208 7510

Email: a.g.s.trebacz2@newcastle.ac.uk

Dr Cristina McKean

Senior Lecturer,

Newcastle University

School of Education Communication and Language Sciences,

King George VI Building,

Queen Victoria Road,

Newcastle upon Tyne.

NE1 7RU

Telephone: 0191 208 6528

Email: cristina.mckean@newcastle.ac.uk

The data protection officer for this study is:

Maureen Wilkinson,

Newcastle University,

Newcastle upon Tyne,

NE1 7RU

Telephone: 0191 208 6000

Email: rec-man@ncl.ac.uk

5

Evaluating the Effectiveness of Building Early Sentences Therapy

Ms Anastasia Trebacz, Dr Cristina McKean, Dr Helen Stringer, and Professor James Law

Consent Form for Parents/Carers

Name of Child		Date of birth
Name of Parent/ Guardian		
Postcode		

Please initial box

1	I confirm that I have read and understood the information sheet dated June 2018 for the above project, have asked any questions I wanted to, and I voluntarily consent for my child and me to take part in this research project.	☐
2	I consent for the information outlined in the information booklet to be collected. I understand what information is being collected, how it is being collected, how it will be stored and why it is being collected, and that all participation is voluntary.	☐
3	I understand that I have the right to withdraw my child at any time, without giving a reason before May 2019. After that all personal identifiable information will be destroyed. All information will be kept fully confidential at all times.	☐
4	I consent for my child to be audio and video recorded, and understand that these recordings will be stored securely and confidentially, and destroyed by the end of the study in accordance with GDPR guidelines.	☐
5	I believe I understand the purpose, extent and possible risks of my child's and my own involvement in this project.	☐
6	I understand that if my child discloses information that indicates he/she may be at risk it will be followed up through established channels.	☐
7	I would like to receive a summary of the outcomes of the study in 2019 after the study is complete (optional)	☐
8		☐

If my child is receiving any speech and language therapy from NHS services I give permission for the researchers to tell the therapists that my child is involved in this study.

Please turn over

Children already receiving Speech and Language Therapy

My child is receiving Speech and Language Therapy (please circle): **Yes / No / Not Sure**

If **yes,** do you give your permission for us to let the SLT services know that your child is involved in the study so they can plan the best way to work with you and your child whilst they are involved in the study? **Yes / No**

What is the name of your speech and language therapist and where do you see them?..

....

....

Language spoken in the home

Which languages are spoken in your home? (please list)

What is the **main** language spoken in your home?

What is the **main** language spoken to your child?

What percentage of the time would you estimate that they hear the **main** language? ___%

Parent/Guardian Signature		Date
Researcher signature	(Anastasia Trebacz)	
Telephone number (in case we have any specific queries relating to your child)		

Please give the form to Anastasia Trebacz or return it to your child's class teacher. You will be given a copy of this form to keep.

Appendix 7 Additional Form to Introduce New GDPR Rules to Parents

Dear Parent/Carer,

Thank you very much for volunteering to be a part of the Building Early Sentences Therapy (BEST) research study!
We have carried out some language assessments with your child, and they are eligible to take part in the rest of the study. We are looking forward to working with your child over the coming weeks.

We would like to provide you with some additional information about how we will look after the information we collect form you and your child in the study.

You may be aware of the new General Data Protection Regulation (GDPR) rules that will now govern how data is managed from the 25th May 2018. We wanted to give you some additional information to explain how we will follow these rules for this study.

The previous information you have received about the study is correct. Please see below some additional information we would like to make you aware of.

<u>What information are we collecting from you?</u>

We are collecting your name and telephone number so we can contact you with any specific queries or concerns about your child. We are collecting your postcode so that we can understand the range of areas where the children in the study live. We are collecting information about whether or not your child is receiving speech and language therapy, and what languages they hear at home, to determine whether they are eligible to take part in the study, and whether the study is best suited to their individual needs. We also ask for your child's name and data of birth, so that we can remove their data from the study if you request this. This will be stored securely, and separate to the rest of their data so all data will be impossible for anyone other than the study team to link to your child.

During the study we will collect language assessment data from your children. This involves recording their responses to tests of language in booklets. We will also ask you and your child's teacher to fill in some questionnaires about your child's language.

We will audio record each assessment so that we can check all of our data is accurate. If your child receives BEST as part of the study, we will also video record the third session so that a member of the research team can check that the therapy is being delivered consistently, increasing the scientific reliability of the study.

Any identifiable data we collect, such as your or your child's name, date of birth and post code, will be stored separately and securely, away from all other data collected during the study (such as the assessment data), so that it is anonymous.

We will not collect any other information that is not discussed here and everything will be kept securely.

No information will be transferred outside of the Newcastle University research team.

Only members of the research team will have access to data collected in this study, and all data will be fully anonymised. We will analyse the data and publish summary results of what we find in research reports but no information regarding individual children will be published.

All identifiable data, including video and audio recordings will be stored securely in a University password protected online repository or in a locked filing cabinet in secure University offices. Video and audio recordings will be destroyed by the end of the study (July 2019), or before then if we no longer need them.

The fully anonymised data will be kept for 5 years after the end of the study whilst we analyse and publish reports about the research and then destroyed (July 2024).

Your right to withdraw

You do not have to take part in this research. You will not be included in any research without first giving your consent. You also have the right to withdraw your consent at any time and therefore opt out of the study. To do this please contact Anastasia Trebacz T: 0191 2086510

E: a.g.s.trebacz2@newcastle.ac.uk

A: School of Education, Communication and Language Sciences,

Newcastle University,

King George VI Building,

Queen Victoria Road,

Newcastle upon Tyne,

NE1 7RU

If you request to withdraw from the study all your information will be destroyed. You will not be disadvantaged in any way if you decide to withdraw, and your child will continue to receive their excellent usual support from school.

In May 2019 the data collected for the study will be fully anonymised and published. You won't be able to withdraw your anonymised data after this point, but it will be fully anonymous and no all of your identifiable information will be deleted.

Who is conducting this research?

This study is being conducted by researchers from Newcastle University. Anastasia Trebacz is the main contact, and her details are above.

The Data Protection Officer is:

Maureen Wilkinson,

Newcastle University,

Newcastle upon Tyne,

NE1 7RU

T: 0191 208 6000

E: rec-man@ncl.ac.uk

Why are we collecting certain information from you?

The purpose of this study is to test whether Building Early Sentences Therapy (BEST) is effective. We are conducting a trial to compare the therapy with usual interventions that the children receive in school. The lawful basis for carrying out this study under GDPR is as a 'Task in the Public Interest, (Article 6,1e) as research is cited as part of the Universities duties. Processing of any special categories of personal data will be Scientific Research (Article 9, 2j)

What to do if you are unhappy

If you are not happy with the way your information is being handled, or with any response received from us, you have the right to lodge a complaint with the Information Commissioner's Office at Wycliffe House, Water Lane, Wilmslow, SK9 5AF

https://ico.org.uk/

Please could you sign and return the form below to your child's teacher if you are happy with the above information.

Please contact Anastasia if you have any queries or questions relating to how we are using your data, or about the study in general.

Name of Child		Date of birth
Name of Parent/ Guardian		

Please initial box

1 I understand which information is being collected, how it is being collected, and how it will be securely stored. I understand why it is being collected. ☐

2 I consent for the information outlined above to be collected, and stored as described above. ☐

3 I consent for my child to be audio and video recorded, and understand that these recordings will be stored securely and confidentially, and destroyed by the end of the study. ☐

4 I understand that I have the right to withdraw my child at any time, without giving a reason before May 2019. After that all personal identifiable information will be destroyed. All information will be kept fully confidential at all times. ☐

You will be provided with a copy of this form to keep.

With very many thanks

Anastasia Trebacz and the BEST Research Team

Appendix 8 CSCOT Accessed From

https://www.thecommunicationtrust.org.uk/resources/resources/resources-for-practitioners/communication-supporting-classroom-observation-tool.aspx

Appendix 9 SDQ 2-4 Questionnaires (parent and educator versions) Accessed From

http://www.sdqinfo.org/py/sdqinfo/b3.py?language=Englishqz(UK)

Appendix 10 FOCUS 34 Questionnaires and FOCUS Instruction Sheet Accessed From Flintbox

https://research.hollandbloorview.ca/outcomemeasures/focus

https://research.hollandbloorview.ca/Assets/Research/Documents/Parent_FOCUS_Instruction_Sheet.pdf

Baseline Assessment

Targeted BEST Assessment Sentences	Generalised BEST Assessment Sentences
The teddy is sitting	The frog is standing
The boy is walking	The king is running
The woman/lady is laughing	The witch is crying
The woman/lady is jumping	The witch is swimming
The boy is riding the/a bike	The king is climbing the/a wall
The baby is smelling the/a sock	The fairy is opening the/a door
The teddy is washing the/a banana	The frog is reading the/a book
The man is eating the/an apple	The pirate is drinking the water
The girl is kicking the/a train	The queen is cutting the/a cake
The woman/lady is brushing the/a cat	The witch is painting the/a picture
The girl is hugging the/a cat	The queen is tickling the/a frog
The boy is kissing the/a teddy	The king is pushing the/a frog
The man is putting the/a cup on the/a table	The pirate is dropping the/a plate on the floor
The baby is pouring milk into the/a cup	The fairy is spilling milkshake on the floor
The girl is giving the/a banana to the/a baby	The queen is feeding the/a pizza to the/a fairy
The boy is throwing the/a banana to the/a man	The king is bouncing the/a ball to the/a pirate

Outcome Assessment

Targeted BEST Assessment Sentences	Generalised BEST Assessment Sentences
The man is laughing	The pirate is crying
The man is sitting	The pirate is standing
The teddy is walking	The frog is running
The girl is jumping	The queen is swimming
The boy is riding the/a horse	The king is climbing the/a tree
The woman/lady is brushing the/a teddy	The witch is painting the/a house
The man is eating the/an orange	The pirate is drinking the coffee
The girl is kissing the/a baby	The queen is pushing the/a fairy
The baby is kicking the/an apple	The fairy is cutting the/a pizza
The baby is smelling the/a banana	The fairy is opening the window
The girl is hugging the/a baby	The queen is tickling the fairy
The girl is washing the/an apple	The queen is reading the/a newspaper
The man is pouring milk into the/a shoe	The pirate is spilling water onto the sofa
The baby is putting the/a key on the/a table	The fairy is dropping the/a bowl on the floor
The boy is giving the/a banana to the/a woman/lady	The king is feeding the bread to the/a witch
The girl is throwing the/a ball to the/a man	The queen is bouncing the/a balloon to the pirate

Follow-up Assessment

Targeted BEST Assessment Sentences	Generalised BEST Assessment Sentences
The baby is sitting	The fairy is standing
The girl is walking	The queen is running
The teddy is laughing	The frog is crying
The boy is jumping	The king is swimming
The boy is riding the/a cat	The king is climbing the/a ladder

The baby is smelling the/an orange	The fairy is opening the/a box
The teddy is washing the/a spoon	The frog is reading the/a letter
The teddy is eating the/a banana	The frog is drinking the/a milkshake
The woman/lady is kicking the/an apple	The witch is cutting the/a cake
The boy is brushing the/a cat	The king is painting the/a car
The girl is hugging the/a teddy	The queen is tickling the/a frog
The boy is kissing the/a horse	The king is pushing the/a rabbit
The man is putting the/a spoon on the/a bed	The pirate is dropping the/a box on the floor
The baby is pouring juice into the/a box	The fairy is spilling coffee onto the sofa
The girl is giving the/a phone to the/a woman/lady	The queen is feeding the cheese to the/a witch
The boy is throwing the/a banana to the/a baby	The king is bouncing the/a ball to the/a pirate

Building Early Sentences Therapy

Training Materials for Teaching Assistants

Anastasia Trebacz

Newcastle University

a.g.s.trebacz2@newcastle.ac.uk

School:

BEST Intervention:

Materials:

Manual

Therapy recording booklet

Laminated prompt sheet

Toy checklist

The Research Study

This study is evaluating a new language therapy to see whether it is effective when compared to other interventions that are currently being delivered in schools in North Tyneside. Your school has been selected to deliver the therapy (with or without additional gestures), and you have been invited to take part in delivering the therapy alongside the researcher.

This information is designed to explain the principles of the therapy, so that you can deliver it alongside the research team.

What is BEST?

Building Early Sentences Therapy (BEST) is a new language intervention that has been designed to support 3-6 year olds with significant language delays or disorders. The therapy helps children to use and understand 2, 3, and 4 clause-element sentences by highlighting the underlying structure of sentences, and therefore improving their ability to abstract across when producing new sentences.

Who is BEST designed for?

BEST has been designed to help children around 3-6 years old who have difficulties with their language. The therapy can be used to help a wide range of children, but for the study we will be working with children who speak English as their main language, and who do not have a diagnosed learning disability. This does not mean that the therapy can't be used for other children who do not fit into this group, but to keep the research scientific we will focus on this group in the present study.

How is BEST delivered?

You will work in partnership with Anastasia, the researcher working on this project, to deliver the therapy. The therapy is delivered twice per week over 8 weeks (totaling 16 sessions). Each session lasts roughly 20 minutes, although this can vary depending on the number of children and their concentration levels.

Each therapy session consists of three small blocks. In each block there are two parts: Part A where we model sentences for the child and Part B where we give the child an opportunity to produce their own sentence.

The BEST Toys

Part A- Modelling the Sentence

To deliver Part A first ensure you have the correct toys for the session using the **toy checklist** (Toys will be provided for each session). Select the toys for the first sentence listed on the **therapy recording booklet** and act out the sentence using the toys (e.g. for the sentence *'the boy is running'*, make the boy figure run along the table). The researcher will say the sentence as you act it out, and provide additional gestures if relevant. Then move on to the second sentence by selecting the toys and acting out the sentence. Continue until you have been through all of the sentences in Part A. There will be small variations between all of the sentences in Part A. The purpose of this is to provide model sentences so that children start to learn which parts stay the same and which parts change within a sentence.

Part B- The child's sentence

To give the child an opportunity to produce a sentence based on actions you act out, select the toys for the appropriate sentence then act out the sentence and, addressing the child by name say "*What's happening*?" Make sure the child can clearly see what you are doing. The researcher will record any responses that the child makes. When the child has had the opportunity to produce a sentence (even if they haven't said anything), move on to acting out the next sentence and asking the child "*What's happening*?". The child does not have to respond, and you do not need to encourage them. We want to wait until the child is ready to respond by themselves. It is important that we do not interrupt the input that they are hearing, so you do not need to praise the child while the session is taking place. The child will feel rewarded by taking part and describing the sentences, and it is not necessary to give additional verbal praise until the end.

At the end of each block put all the toys into the middle of the table and allow children a few minutes' free play. Comment on what they are doing, and provide additional model sentences, for example if they act out the man stroking the horse then say: "Oh look, the man is stroking the horse!"

<u>Some tips for creating the best therapy session:</u>

- It is important that therapy sessions are consistent across the experiment, so please follow the above directions closely
- You do not need to encourage the child to participate or produce sentences if they do not want to. The most important thing is that they hear the input we are providing, so don't worry if they aren't keen to join in yet. (Children usually become more comfortable with the structure after a couple of sessions anyway, so you will probably notice that they will join in on their own, given time)
- Although it may seem unnatural at first, you do not need to praise the child while the therapy session is taking place. They will be rewarded by taking part, describing the sentences and having the opportunity to interact with the toys. It is the aim that they are exposed to the language input without disruption where possible, so try to keep other talking to a minimum while the session is taking place. If you need to reengage the child and draw them back to the task then it is fine to do so, especially by using their name.
- If you have questions at any point please speak to Anastasia or contact her using the details below
- The sessions are designed to be fun and enjoyable for children, and will run smoothly if children are engaged and able to take turns.

Contact information

If you have any questions, or feel that any of the information in the manual is not clear, please let me know. Some of the information may also become clearer once I have demonstrated a therapy session.

If you have any questions at any point please email me at a.g.s.trebacz2@newcastle.ac.uk

Research Associates' Protocol- Refining and evaluating the efficacy of Building Early Sentences Therapy (BEST): a constructivist language intervention for children with language difficulties in the early years

Anastasia Trebacz

a.g.s.trebacz2@newcastle.ac.uk

07990930765

Room G16 KGVI

Contents

Timeline

It is possible that timings could change, and you will be kept up to date. Please let me know if there are any problems with the prospective dates.

First Data Wave

2/3/18 Last day for recruitment

16/4/18-4/5/18 BASELINE MEASURES (excluding Thursdays) (2 weeks + an extra week for 'mopping up')

25/6/18-6/7/18 OUTCOME MEASURES (excluding Thursdays)

Second Data Wave

ASAP- Follow-Up Measures (Wave 1)

ASAP- Baseline Measures (Wave 2)

Expenses

Please submit all receipts, bus tickets etc to me and provide me with your bank details.

Data Collection

Materials

For each assessment you will need:

- BEST eligibility assessment
- NRDLS assessment
- NRDLS manual
- NRDLS recording form
- BEST Assessment picture booklet
- BEST Assessment picture booklet (Alternative Pictures Version)
- Language Probe picture booklet
- BEST Therapy recording form
- Language Probe recording form
- SDQ teacher version
- FOCUS teacher version

To give to teachers to complete

- SDQ parent version
- FOCUS parent version

To ask teachers to send home (in an envelope addressed to child's parent/carers)

- Warmup activity selection
- Stickers
- Certificate for Stickers

All materials are labelled on my desk. I will also upload electronic copies to Onedrive. You can take multiple blank forms at a time and keep hold of the assessment materials so that you have them ready for each assessment. Please let me know if anything is running low, or if there are any issues at all with materials. Sarah has a selection of warm-up activities which she is happy to use.

Procedure

You will be given a full list of schools (blinded to treatment arm) and children taking part (with participant codes). Contact details will be provided for the person most involved with the study from each school (this is generally the SENCO, although some are Early Years Teaching Assistants/Head teachers).

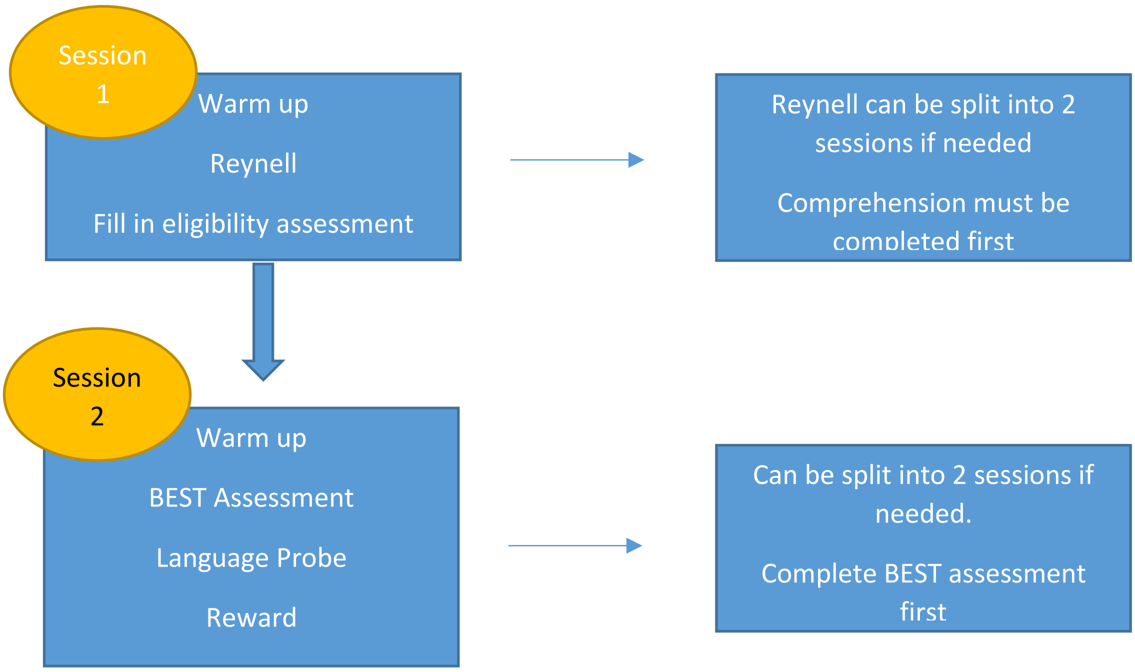
You will liaise with schools to decide on the best time to visit each. I can help as much as you need with this, but generally if you organise your own timetables it will be easiest for you.

Please update the timetable spreadsheet on the Onedrive BEST group so that we all know where everybody is and when. This will also mean we can contact schools in the event of your absence.

Discuss with teachers whether there is a separate room or quiet space, away from other children, where you can complete the assessments. Most have already said that there are specific rooms you can use, but may need reminding! Also discuss the order in which you see children, and if there are some children who are best seen at a particular time due to concentration etc. The more flexible we can be, the better!

Use whatever warm-up activities you would like to. Remember to keep them short, fun and let the child win!

The assessment sessions will be split into two parts for each child.



Data Transcribing

General Guidelines

Transcribe everything as you go along. This is especially important for comprehension items where there will not be a verbal response, as they won't be recorded to go back to. Try to record as much detail as possible, including the full response if it is different from the target and any self-corrections. Also record if you had to repeat any items to keep the child on task (these should then be marked as incorrect). Try not to feel rushed at the expense of good transcribing; if needed ask children to help put the toys away, turn the page etc. so you have more time to transcribe accurately. Flag any items you are unsure of with an asterisk to go back to later. These can be discussed with each other, and Ana/Cristina if necessary.

Please arrange with schools the best days and times to visit each. Contact details will be provided. Ana will help wherever necessary, but maintaining your own schedule will make things easier for you. Please add your timetable to the Onedrive document so we know where you are and when.

Complete the sessions over two days wherever possible, or at least over a morning and afternoon if children are in for full days, as this will be better for the child's concentration. Use the sticker certificates to encourage children to participate, but cease the session if necessary.

Remember to check the audio recording device is working before each session, and audio record each session. Start each new recording by stating the participant code, school, date and session number e.g. "This is GW, Bailey Green, 13th May, session number two". (Please don't use your phone as we don't have ethical approval to do so).

At the end of each day of assessments in schools, please go over all your recordings and check them against your transcriptions. Correct any if necessary, and discuss with each other/Ana if any are ambiguous/unclear.

Reynell Guidelines

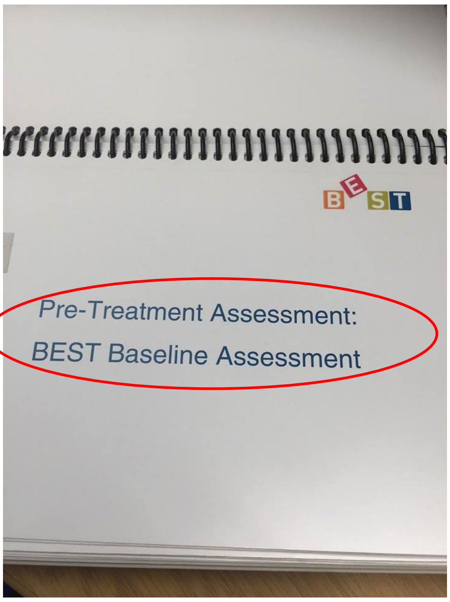
Remember to check the audio recording device is working and audio record each session. Start each new recording with the participant code, school, date and session number e.g. "This is GW, Bailey Green, 13th May, session number two".

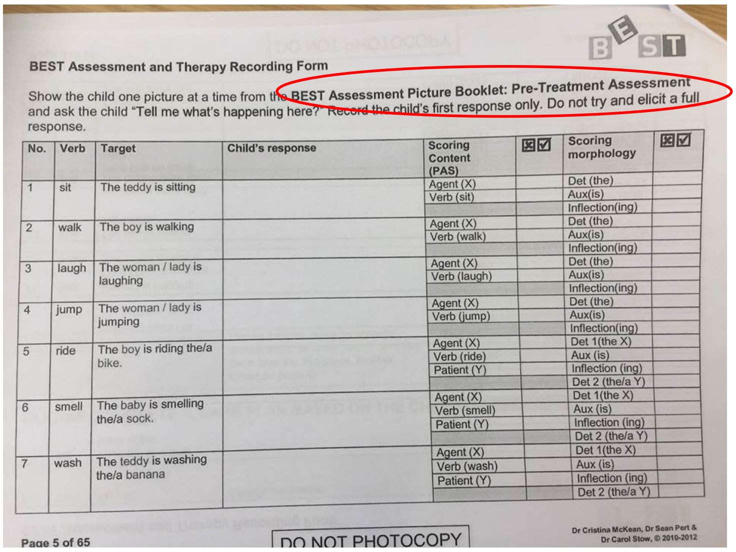
Administer the Reynell according to the instructions in the manual, and refer to them in the first instance. There are also some answers to additional queries at the back of this document. You can use the practice items to ensure children understand the task, but please then stick rigidly to the script in the manual/recording form for the test items. If there are further queries that are not answered please ask Ana, and if necessary Cristina.

BEST Assessment & Language Probe Guidelines

Remember to check the audio recording device is working and audio record each session. Start each new recording with the participant code, school, date and session number e.g. "This is GW, Bailey Green, 13th May, session number two".

Ensure you have the correct assessment and this matches the assessment recording form (see images below). The same format will be used for the Language Probe.





BEST Assessment and Therapy Recording Form

Show the child one picture at a time from the BEST Assessment Picture Booklet: Pre-Treatment Assessment and ask the child "Tell me what's happening here?" Record the child's first response only. Do not try and elicit a full response.

No.	Verb	Target	Child's response	Scoring Content (PAS)	☒☑	Scoring morphology	☒☑
1	sit	The teddy is sitting		Agent (X) Verb (sit)		Det (the) Aux(is) Inflection(ing)	
2	walk	The boy is walking		Agent (X) Verb (walk)		Det (the) Aux(is) Inflection(ing)	
3	laugh	The woman / lady is laughing		Agent (X) Verb (laugh)		Det (the) Aux(is) Inflection(ing)	
4	jump	The woman / lady is jumping		Agent (X) Verb (jump)		Det (the) Aux(is) Inflection(ing)	
5	ride	The boy is riding the/a bike.		Agent (X) Verb (ride) Patient (Y)		Det 1(the X) Aux (is) Inflection (ing) Det 2 (the/a Y)	
6	smell	The baby is smelling the/a sock.		Agent (X) Verb (smell) Patient (Y)		Det 1(the X) Aux (is) Inflection (ing) Det 2 (the/a Y)	
7	wash	The teddy is washing the/a banana		Agent (X) Verb (wash) Patient (Y)		Det 1(the X) Aux (is) Inflection (ing) Det 2 (the/a Y)	

Page 5 of 65 DO NOT PHOTOCOPY Dr Cristina McKean, Dr Sean Pert & Dr Carol Stow, © 2010-2012

BEST Picture Assessment Booklet *BEST Assessment Recording Form*

Before carrying out the assessments, please familiarise yourself with the instructions below. Show the child one picture at a time and say "Tell me what's happening here". Record their full utterance, including any self-correction in the 'Child's Response' space on the form. Mark any items you are unsure of, or that you cannot understand from the child's response with an asterisk. Go back to these later, refer to the manual in the first instance, and discuss with each other/Ana if there are any ambiguities.

Scoring Protocols for the BEST assessment and Language Probe can be found below. Please note that they are very similar, as they are scored in the same way, with the target items being the only difference.

BEST Assessment and Language Probe Scoring Protocol

BEST Assessment and Language Probe Baseline Assessment

The BEST Assessment and Language Probe baseline assessments will be completed at the beginning of the study. They are picture book assessments of the child's expressive language. For both assessments, the child is shown one picture at a time. The Researcher says "Tell me what's happening here". The child's response (if any) is then recorded.

It is not necessary to try and elicit full responses, nor to repair, remodel or in any way attempt to support the child to produce longer spoken sentences. This is a baseline assessment and will be compared with the later assessments to establish whether the child has made progress. The aim of the baseline is to capture the child's spontaneous unsupported expressive language ability.

Scoring the baseline assessment

The target spoken sentences are listed on the assessment form. Scoring is divided into four columns: Content (Signalled and Correct) and Morphology (Signalled and Correct).

Content verbs are nouns and verbs.

Morphology includes:

- Determiners such as 'a' and 'the'
- Auxiliaries such as 'is'
- Inflections such as '-ing'

The assessments are looking at two things: Firstly whether or not the child is able to use the correct argument structure. This is determined by whether or not children 'Signal' a response, by inserting a relevant (although not necessarily correct) item into the correct place in the sentence (such as inserting an agent into the agent slot, or a Determiner into a Determiner slot). Secondly, whether the child can correctly produce the target sentence as it appears on the Assessment recording form.

Scoring 'Signalled' and 'Correct' Responses

Tick the 'signalled' response box if the child includes an item relevant to the category being marked (see below for more information). Cross the box if the child does not signal a relevant word category

Tick the 'Correct' box if the child produces the target item that appears in the Target column of the form. Cross the box if they do not produce the target response. Note it is possible for children to score a point for 'signalled' and not score a point for 'correct' on an item. If the child does not score a point for 'signalled', it would then be impossible for them to score a point on 'correct' for that item.

BEST Assessment Scoring Examples

ssment and Language Probe Recording Form

ild one picture at a time from the **BEST Assessment Picture Booklet: Pre-Treatment Assessment**
child "Tell me what's happening here?" Record the child's first response only. Do not try and elicit a full

Target	Child's response	Scoring Content (PAS)	Sign alled	Corr ect	Scoring morphology	Sign alled	Cor rect
The teddy is sitting	*The man is sitting*	Agent (X)	✓	X	Det (the)		
		Verb (sit)			Aux(is)		
					Inflection(ing)		

For example in the item above the **agent** is **teddy**. If the child said "The **man** is sitting", this would be marked with a tick in the 'signalled' column for **agent**, because the child attempted to use an **agent** in the correct 'slot' in the sentence.

Of course, in the 'Correct' column, it would be marked with a cross because the child did not use the correct item 'teddy'

No.	Target	Child's response	Scoring Content (PAS)	Signalled	Correct	Scoring morphology	Signalled	Correct
1	The girl is kissing the/a baby	*Girl kiss the boy*	Agent (X)	✓	✓	Det 1 (the X)	X	X
			Verb (kiss)	✓	✓	Aux (is)	X	X
			Patient (Y)	✓	X	Inflection (ing)	X	X
						Det 2 (the/a Y)	✓	✓
2	The baby is kicking the/an apple	*Baby is hitting the ball*	Agent (X)	✓	✓	Det 1(the X)	X	X
			Verb (kick)	✓	X	Aux (is)	✓	✓
			Patient (Y)	✓	X	Inflection (ing)	✓	✓
						Det 2 (the/an Y)	✓	✓

In the above example (1):

The child did not signal Det 1, and was therefore also incorrect

The child signalled an Agent (girl), and was also correct

The child did not signal the Auxiliary, and was therefore also incorrect

The child signalled the verb (kiss), and was also correct

The child did not signal the inflection, and was therefore also incorrect

The child signalled Det 2 (the), and was also correct

The child signalled a Patient (boy), but was incorrect (the target was baby)

In the example (2) above:

The child did not signal Det 1, and was therefore also incorrect

The child signalled the Agent (baby), and was also correct

The child signalled the Auxiliary (is), and was also correct

The child signalled a verb (hit), but was incorrect (the target was kick)

The child signalled the Auxiliary ('-ing'), and was also correct

The child signalled Det 2, and was also correct

The child signalled a Patient (Ball), but was incorrect (the patient was apple)

Language Probe Scoring Examples

BEST Assessment and Language Probe Recording Form

Show the child one picture at a time from the **Language Probe Picture Booklet: Pre-Treatment Assessment** and ask the child "Tell me what's happening here?" Record the child's first response only. Do not try and elicit a full response.

No.	Verb	Target	Child's response	Scoring Content (PAS)	Signalled	Correct	Scoring morphology	Signalled	Correct
1	stand	The frog is standing	*The rabbit is standing*	Agent (X)	✓	X	Det (the)		
				Verb (stand)			Aux(is)		
							Inflection(ing)		

For example in the item above the **agent** is **frog**. If the child said “The **rabbit** is standing”, this would be marked with a tick in the ‘signalled’ column for **agent**, because the child attempted to use an **agent** in the correct ‘slot’ in the sentence.

Of course, in the ‘Correct’ column, it would be marked with a cross because the child did not use the correct item ‘frog’

	Target	Child’s response	Scoring Content (PAS)	Signalled	Correct	Scoring morphology	Signalled	Correct
1	The queen is pushing the/a fairy	*The lady hit the fairy*	Agent (X)	✓	X	Det 1 (the X)	✓	✓
			Verb (push)	✓	X	Aux (is)	X	X
			Patient (Y)	✓	✓	Inflection (ing)	X	X
						Det 2 (the/a Y)	✓	✓
2	The fairy is cutting the/a pizza	*One fairy was cutted that pizza*	Agent (X)	✓	✓	Det 1(the X)	✓	X
			Verb (cut)	✓	✓	Aux (is)	✓	X
			Patient (Y)	✓	✓	Inflection (ing)	✓	X
						Det 2 (the/a Y)	✓	X

In the above example (1):

The child signalled Det 1 and was also correct

The child signalled an Agent (lady) but was incorrect (the target was ‘fairy’)

The child did not signal an Auxiliary, and was therefore also incorrect

The child signalled a verb (hit), but was incorrect (the target was ‘push’)

The child did not signal an inflection, and was therefore also incorrect

The child signalled Det 2 and was also correct

The child signalled the Patient (fairy), and was correct

In the above example (2):

The child signalled Det 1 (One), but was incorrect (the target was ‘the’)

The child signalled an Agent (fairy), and was also correct

The child signalled an Auxiliary (was), but was incorrect (the target was ‘is’)

The child signalled a verb (cut), and was also correct

The child signalled an inflection (‘-ed’), but was incorrect (the target was ‘-ing’)

The child signalled Det 2 (‘that’), but was incorrect (the target was ‘the/a’)

The child signalled the Patient (pizza), and was also correct

Further Scoring Information

- Off-topic, irrelevant and other such spoken utterances are scored as incorrect.
- Count the number of ticks for ‘signalled’ and ‘correct’ and calculate the percentage correct in the row at the end of the assessment form. Transfer scores to the front of the recording booklet.

- Plot the converted percentage scores onto the second graph (page 5). Please use a red pen to plot the 'Signalled' responses, and a black pen to plot the 'correct' responses.
- Acceptable determiners for the agent are 'that', 'the' and 'a'.
- Acceptable determiners for the patient are 'the', 'a', 'his/her'.
- Where a noun argument is marked but the wrong lexical item is used e.g. 'dog' for 'cat' this is scored as 'signalled' but not 'correct'.
- Where a subject pronoun 'he or she' is used for the agent this is accepted as 'signalled' and 'correct'.
- Where a child makes many attempts or produces separate clause elements as separate utterances, only the first attempt/utterance is analysed.
- Where a child makes a false start and self corrects part way through an utterance the complete corrected sentence is analysed e.g. 'the boy… **the man is eating the apple'**.

BEST Assessment and Language Probe Outcome Assessments

The Outcome Assessments will be completed post-treatment or TAU period.

Each assessment consists of sixteen pictures. Record and score the assessment in the same way as the BEST assessment and Language Probe baseline assessments. Transfer the scores to the tables on the front cover and pages 2-3. Plot the child's progress on the progress tracker graph on page 4 for the BEST assessment and page 5 for the language probe. Use red ink to plot the 'signalled' scores and black ink to plot the 'correct' scores.

BEST Assessment and Language Probe Follow-Up Assessments

The Follow-Up Assessments will be completed six weeks after the Outcome Assessments.

The assessments consists of sixteen pictures each. Record and score the assessment in the same way as the BEST Assessment and Language Probe baseline and outcome assessments. Transfer the scores to the tables on the front cover and pages 2-3. Plot the child's progress on the progress tracker graph on page 4 for the BEST Assessment and page 5 for the Language Probe. Use red ink to plot the 'signalled' scores and black ink to plot the 'correct' scores. The form should now be complete.

Blinding

Please try to remind teachers that you are not supposed to know which schools are in which intervention arms! In practice this will be difficult, but we can try our best! Writing a line to this effect as your email sign-off might help remind teachers not to disclose their treatment arm. Also if checking anything with Ana, please don't identify any particular children so that blinding is maintained around children's assessment data. If blinding is not maintained for any schools please record this with reasons.

Data Scoring and Enhancement

General Guidelines

You will have a copy of the participant codes which must be stored securely along with any of the sensitive data. If any parent or teacher questionnaires are returned to you they must also be stored securely in the locked filing cabinet, as they will contain identifiable child data. Record all sessions with the provided audio recording device.

For any items you are unsure of, you can check the manuals for information. If there are still queries, please ask Ana (but don't reveal the identity of the child!), and we can check with Cristina or Helen if necessary.

Reynell Guidelines

Follow all manual guidelines in the first instance. Score all items according to the guidelines in the assessment recording form. These are straightforward and presented for each individual section. Record raw scores, and use the conversion tables in the appendix of the manual to convert raw scores to standard scores, record confidence bands, percentile ranks and age equivalents. Another researcher should then check 20% (every 5 assessments) to and identify any issues with scoring.

BEST Assessment and Language Probe Guidelines

Follow the manual guidelines in the first instance. Score all items according to the guidelines in the assessment recording forms. They are generally straightforward, but some ambiguities may arise. The manual details additional acceptable responses, as well as some responses which would not be accepted. If you are still unsure, please check with each other or Ana/Cristina if necessary. Create percentage scores for content and morphology by totalling the number of correct items for each. These can be recorded on the assessment recording forms. Then score and total the morphology and argument structure scores, and record these in the spaces provided.

Reliability Checks

Initially you will cross check 20% of each other's assessments against the audio recordings. If we find that reliability Cronbach's alphas are not sufficiently high we will discuss what to do with Cristina, and continue to check a higher percentage. Use a red pen to make any corrections, as this allows us to look back and be aware of any errors.

Data Entry

Spreadsheets will be set up in secure folders on the onedrive BEST group for data entry. We will only enter data into the spreadsheets when reliability checks have been carried out for both the transcriptions and the final adding and converting of scores, and any queries regarding responses have been addressed. Spreadsheets will be clearly labelled with instructions on how to fill them in, but ask Ana if there are any queries. The Reynell scores, and more detailed BEST and Language Probe data will be entered.

Data Storage

All Audio recordings should be uploaded to the onedrive folder called 'Research Assistant Audio Recordings' and then deleted from the audio recording devices. Each file should be named with the child code, school name, date of recording, and the assessment part number in the format current part/total of all parts

For example, DH-Monkhouse-26.02.18-Part1/3

Hard copies of forms must be brought back to the University at the end of each day. They should be labelled with the child code, school code and date. They will be kept in the filing cabinet next to Joe's desk in G16. You will keep the key between you, and Ana will not have access to this data.

FAQs

NRDLS FAQs

Questions asked to Carolyn Letts.

1. Comprehension Section Ci: VERBS

- Make Monkey run. A girl held Monkey and dragged it on the table back and forth. To me the action was ambiguous. What counts as running? Should there be definitely leg movements?

Answer: Carolyn said that for situations such as these, first refer to the manual (pp.29-39) which lists some common ambiguities that may occur in each section. If the ambiguity you encounter is not mentioned, Carolyn said that provided the action is distinguished adequately from other possibilities, and makes intuitive sense it is probably okay.

She recommends balancing is used in these situations, i.e. when you give benefit of the doubt once, do not give it the next time, and so on (so give benefit of the doubt every other time). Carolyn also pointed out that you can never be 100% sure in such situations, so balancing when you mark items as right or wrong is a good way to deal with this.

2. Production Section Ci: VERBS

- Item 23: I dropped Monkey from above and a boy said 'he dropped…he falled' (target response is fall). Should I make the action of 'falling' clearer? If so, how?

Answer: Pre-empting difficulties can be important here, and obviously the more experience we have the easier it will become! Carolyn suggested placing monkey on the edge of a table and making him 'fall off', which would be a good way to differentiate from 'dropping'. (Could you also possibly use a 'Weeeeee' sound effect?)

If any other difficulties arise like this, please do mention them and perhaps we can come up with ways of making them clearer to the child.

3. Production Section Ci: VERBS

- Tester makes Monkey jump. Answer: Monkey is bouncing (instead of Monkey is jumping) **Practice item*

Is this acceptable?

Answer: Some verbs were deliberately selected as practice items instead of target items because they frequently threw up ambiguities such as these.

Because it is a practice item, they are not being scored, and the child was quite clearly on the right track, it is probably not a big issue if they produce 'bouncing' instead of 'jumping'. If any specific difficulties arise with the target items however, we can use the audio recordings to go back and deal with them on a case by case basis, so please do flag them up.

If however a child gets an object or word wrong, see the question below.

4. Production Section Dii: SENTENCE BUILDING

- Monkey's kicking the table. Answer: Monkey's kicking the chair **Practice item*

Is this acceptable?

Answer: If a child makes a mistake on a word or object which is not reasonable (i.e. bouncing vs jumping above) then use the practice items to help them as much as possible. Carolyn said this may involve some teaching. You could say something like 'that's not a chair, it's a …. (pause and let the child answer if possible)'. The script does not therefore have to be adhered to strictly if it is necessary to provide a bit more help on the practice items, but for the test items normal procedures apply.

Intervene as much as possible so that the child fully grasps the task before embarking on the test items.

For target items this would just be marked as wrong.

5. Production Section E: VERB MORPHOLOGY

- Item 43

Target response: Salutes

What if the child has speech sound problems? Difficult to understand and know whether he/she is correct.

Answer: If you already know through speaking to the teacher, or if it becomes apparent through the test that a child does not yet produce e.g. final fricatives/final consonant clusters then Carolyn feels it is not really fair to complete the relevant sections. Alternatively they could be completed but with the caveat of a known speech sound difficulty which should be clearly reported.

If the child is using them on some occasions but not all, then the sections should be completed, and incorrect items marked as wrong. Some judgement may need to be used here, but all relevant sections will be audio recorded so flag them up if unsure and we can go over them. The same goes for any responses where you struggle to understand the child's response.

6. Stopping Procedures: What to do?

The manual details stopping procedures (pp.39). However Carolyn suggested that some judgement should be used.

Answer: If a child fails a whole section then do try a few items from the next section. If children are visibly struggling and clearly only getting one or two right due to chance over a couple of sections then you can stop the test. In this respect, not every item must be failed to constitute failing a block, but as the sections measure different facets of language that do not necessarily develop cumulatively, trying a few items from the next sections is still very important.

Adhere to this procedure for both comprehension and production sections. In reality, even if you did continue they would probably only get a couple of extra points through chance and this will not have a huge bearing on the overall score. Please record clearly when and why the test was stopped.

6. BEST assessment pluralisation: Is it acceptable?

If the child produces 'The (Boy) is smelling socks instead of the target 'smelling the sock', this is not correct, since the picture depicts only one sock. The child would therefore be marked as incorrect for the Det 2 (the/a).

Further ambiguities that have arisen throughout the first data wave. Refer to these and then to me if your query is not answered here

BEST Scoring Ambiguities

1. Child says boy instead of pirates, but he said the word pirate before. This is marked incorrect.

2. 'it's the queen' for ''the fairy is dropping the bowl on the floor" - marked as if said 'the queen'

3. Put the milk – decided as excess information

4. "The boy is riding the/a bike" – response: "A bike"

a. Patient is scored as signalled and correct

5. The teddy is washing the/a banana – response: "He's washing a banana"

6. Cuddle instead of hug is accepted

7. "That" without referring to the agent is not accepted

8. Princess instead of Queen is accepted

9. Giggle instead of laugh is accepted

10. Because the animals are portrayed with human behaviours (e.g. reading), “He/she” instead of the “the bear/ frog” is signalled and correct.

11. “Drink” scored as signalled and correct instead of milkshake

12. Accept “that” for benefactive as well as agent – signalled and correct

13. "Mug” instead of “cup” – marked as signalled and correct (No. 14) - coffee or tea instead of cup? Should have been pen or watch or something that can't be misconstrued

14. “He’s throwing a banana to him.” Instead of “The boy is throwing the/a banana to the/a man” – Benefactive signalled and correct. (No. 16)

15. If the child says “he” instead of “the boy/man/ king” – Agent signalled and correct, Determiner X not signalled and incorrect.

16. "Biking" instead of "riding the/a bike" - Verb signalled but incorrect.

17. 'in' for 'into' is ok

18. Toast for bread = ok

19. "Pot" instead of "bowl" - Patient signalled and correct

20. 'flower' for 'picture' when picture is depicting a flower is OK.

21. For a reciprocal sentence (giving/throwing when the picture is ambiguous), allow agent and benefactive to be switched.

22. For throwing/giving sentences that involve 'to the (benefactive)', they are grammatically correct if said 'the boy is giving the woman the banana'. These are marked as fully correct, minus the preposition 'to'.

23. In determiner 'the king is feeding the bread to the witch' - without 'the' in 'the bread'. Would've been better to use something that cannot be of an undetermined amount, e.g. flower, key, book, watch, crown, balloon, box. We have scored as incorrect

24. 'the big boy/little boy' for man and baby, not scored as correct, but signalled.

25. 'a pirate drinking a drink' - 'DRINK' not accepted for cup of tea/coffee

26. Pluralising – e.g. 'kicking apples' when should be 'kicking the apple' - scored as patient signalled and correct, but no determiner

27. 'He/she' instead of the baby is signalled and correct.

28. BEST outcome assessment picture 14 is ambiguous - 'The fairy's putting some feathers in the bowl' patient (feathers) signalled but incorrect and locative (bowl) signalled and correct.

29. The lady is hugging the baby – can carrying or holding be allowed? BGEC

30. 'the girl is throwing the ball to the man' - response BGEC 'the man is playing catch with the little girl' marked as correct and signalled agent, det1, ing, det3 and benefactive, signalled but incorrect verb and prep, and incorrect and unsignalled patient and det2.

31. BEST outcome assessment item no. 13 - 'tipping milk' instead of 'pouring milk' is signalled and correct. (WHHG)

32. 'Passing' instead of 'throwing' is accepted – signalled and correct.

Appendix 14 Signing Manual Displaying all Signs Used in the Study

PGSS and Makaton Signs for the BEST Trial

‘Refining and evaluating the efficacy of Building Early Sentences Therapy (BEST): a constructivist language intervention for children with language difficulties in the early years’

Anastasia Trebacz

Newcastle University

a.g.s.trebacz2@newcastle.ac.uk

Welcome to the sign manual for this evaluation of Building Early Sentences Therapy!

This manual contains the all the signs you need to sign every sentence in BEST.

You will notice that two different signing systems are used throughout this trial. One is called Paget Gorman Signed Speech (PGSS). You can find out more about PGSS on their website

https://www.pagetgorman.org/

In this study we are using PGSS to sign the grammatical words such as The, A/An, Is, and the verb ending –Ing.

The second signing system is called Makaton. This is a popular signing system used in many schools. You can find out more about Makaton on their website

https://www.makaton.org/

In this study we are using Makaton to sign the content words, as well as the agents and patients (or characters and objects) in the BEST sentences.

The signs in this manual are categorised into three groups:

There is also an index page at the back to help you find the sign you need quickly.

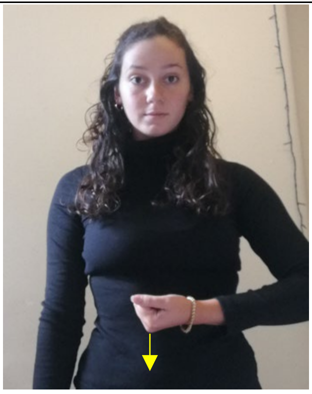
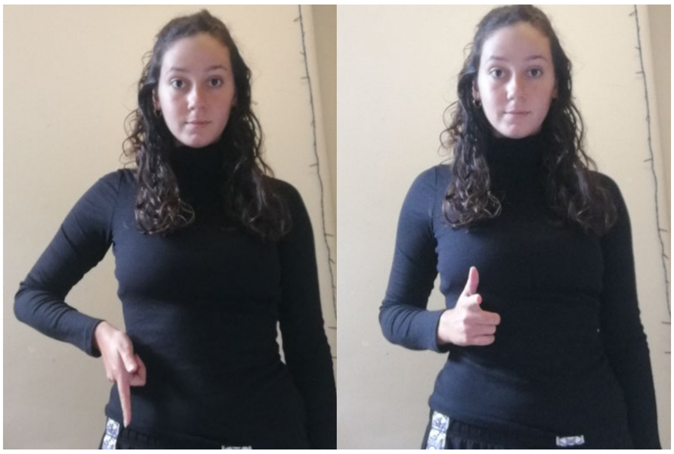
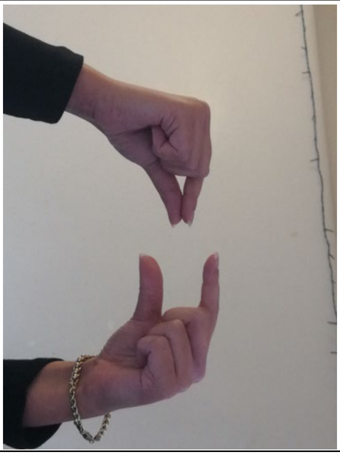
If you have any questions about the signs used in this study, or want to find out more about using the signs please visit the websites above, or contact Anastasia Trebacz at

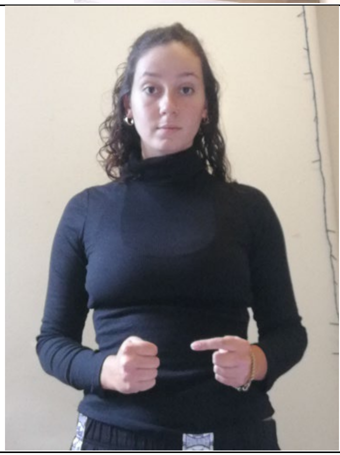
a.g.s.trebacz2@newcastle.ac.uk

Happy Signing!

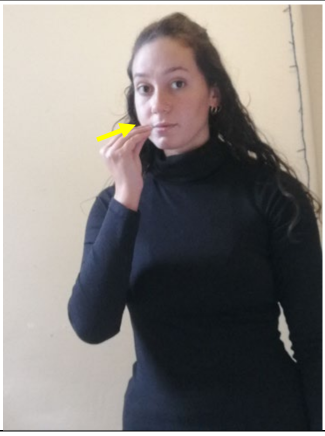
Morphology Signs

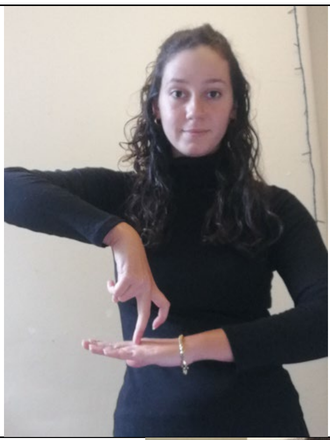
Word	Signing System	Sign

The	PGSS	
A/An	PGSS	
Is	PGSS	
-Ing	PGSS	

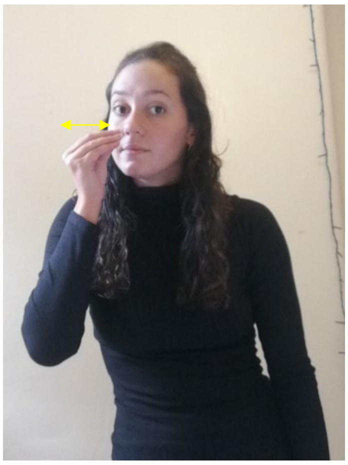
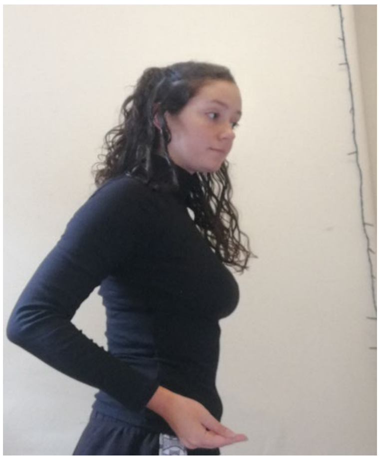
On	PGSS	
To	PGSS	
Into	PGSS	 

Verb Signs

To Brush	Makaton	
To Eat	Makaton	
To Give	Makaton	

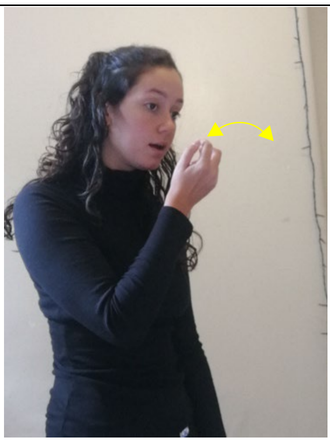
To Hug	Makaton	
To Jump	Makaton	
To Kick	Makaton	
To Kiss	Makaton	

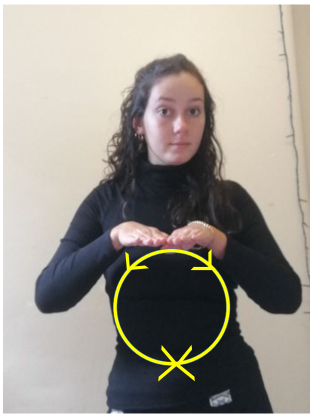
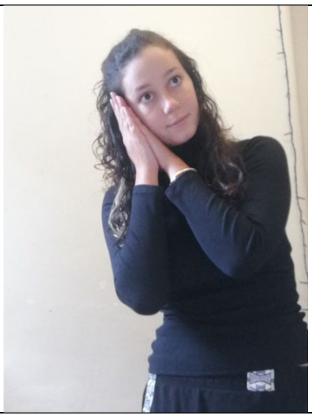
To Laugh	Makaton	
To Pour	Makaton	
To Put	Makaton	

To Ride	Makaton	
To Sit	Makaton	
To Smell	Makaton	
To Throw	Makaton	

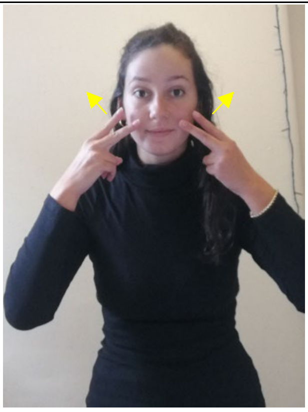
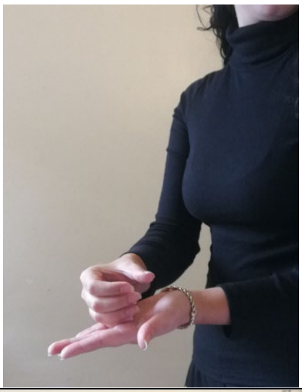
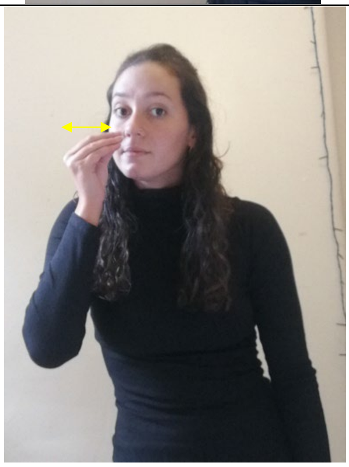
To Walk	Makaton	
To Wash	Makaton	

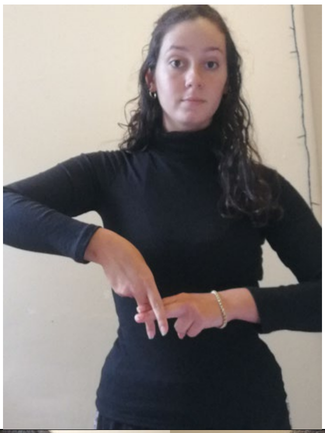
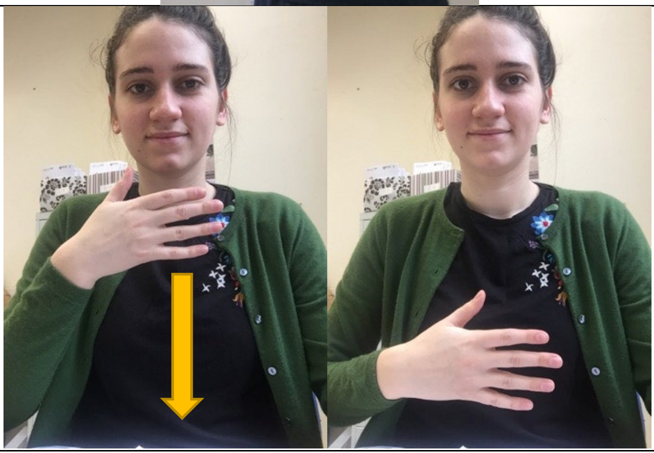
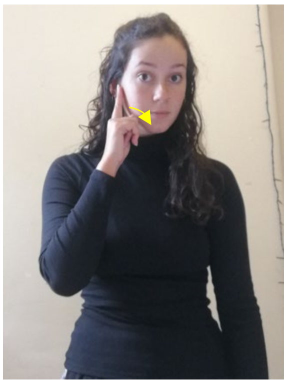
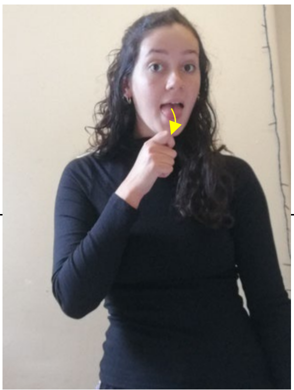
Content Word Signs

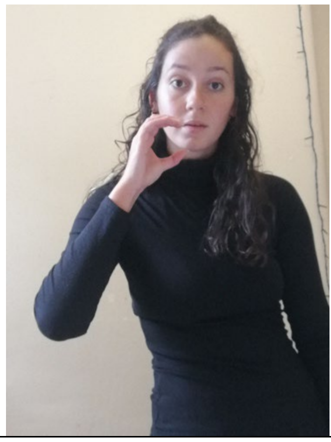
Apple	Makaton	

Baby	Makaton	
Ball	Makaton	
Banana	Makaton	
Bed	Makaton	

Bike	Makaton	
Box	Makaton	
Boy	Makaton	
Carrot	Makaton	

Cat	Makaton	
Cup	Makaton	
Flower	Makaton	
Girl	Makaton	

Horse	Makaton	
Juice	Makaton	
Key	Makaton	
Lady	Makaton	
Lolly	Makaton	

Man	Makaton	
Milk	Makaton	
Orange	Makaton	

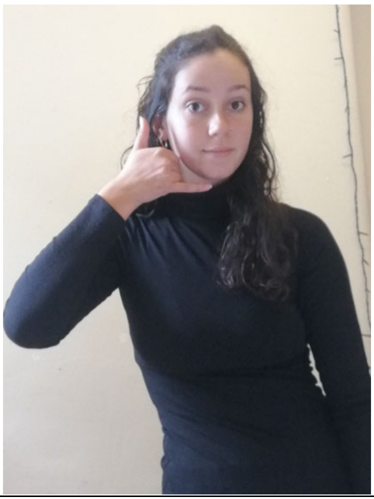
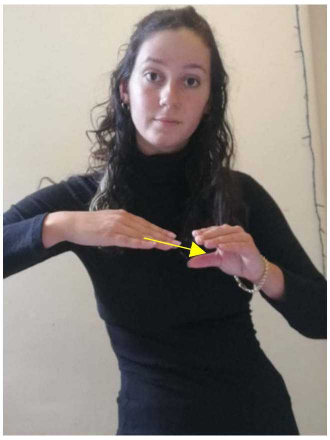
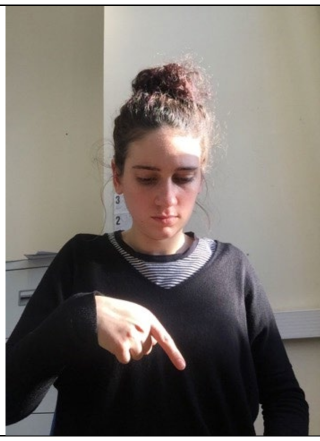
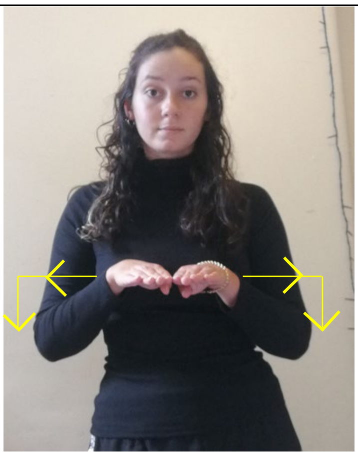
Phone	Makaton	
Shoe	Makaton	
Sock	Makaton	 
Spoon	Makaton	

Table	Makaton	
Teddy	Makaton	
Train	Makaton	

Index

Question/area of discussion with teacher	**Responses from teachers who completed the questionnaire/discussion.**
Which **interventions** are currently in use in the nursery/school classroom to support language?	BLAST, Nursery Narrative, Reception Narrative, Nursery SLT Toolbox for speech sound/language comprehension, in-house 1:1 support and interventions e.g. for non-verbal child with suspected ASD, Time to Talk, 1:1 Speech and Language Therapy, fine motor skills e.g. pencil control, number recognition, phonics- small groups, listening and attention skills
Are there certain **activities** carried out in the classroom that particularly aim to support/target language development?	Communication support. Visual strategies such as now and next boards for children, games and materials rich in language (thought is given when selecting them, frequent recaps, building on instructions). Phonics table in the classroom with activities available to children during different periods of the day. Interactive games. Real/nonce word distinguishing games, Ipad based activities. Talk for writing, Circle time, Story time sessions, milk time discussions, worship, home communication books, daily chat, reading record books, Drawing and art-based activities

Are there particular **strategies** that you and/or other staff members use to help support language?	Screening children to identify need, grouping children with similar needs and delivering in-house help before referral comes through (trying to intervene with some support as quickly as possible). Using group time, discussion time, and snack times as an opportunity to get children to speak and share ideas etc, and even those children who would usually be quieter. Widget (communication in print), Picture and word signs clearly displayed around classroom, read writing chart, talk for writing, action story maps used to help with storytelling, Ipad app 'Keezy' allows children to record a sentence before they write it, to help with literacy (audio rehearsal). Teachers don't 'baby' children using simpler words, but rather use complex words, teach meaning and checking understanding. Word bank and vocabulary wall displays current relevant vocabulary. Carry through language from stories and discussions into writing. Now and next boards, symbols and signs in the classroom, regimented structure helps children to know what happens next, routine board at the front that they plan out first thing Visual symbols, pictures, signs (mixture of Makaton and signs which have been created by the teacher and children and displayed as photos around the classroom, particularly for transitional/time words and phrases). Story reading, discreet daily phonics (+ vocabulary with it). Linking actions to words and using pictures/educational videos. Signing, working in partnership with parents and speech and language therapists. Visual timetable, repetitive routine (key phrases), effective questioning
Who receives intervention?	Certain identified as having additional need receive nursery narrative and BLAST. One school also puts some of the strongest children in the class into the BLAST group to model language, but attention is paid to make sure that they do not prevent the children who need it from taking part. Children not meeting/working below age related expectations are identified and then receive intervention
How do you identify children in need of extra language support?	Talking to the child, parent consultation, baseline assessment, NHS Speech and Language Screener, intuition, teacher tracker system to identify those below age related expectations. Observing social interaction.

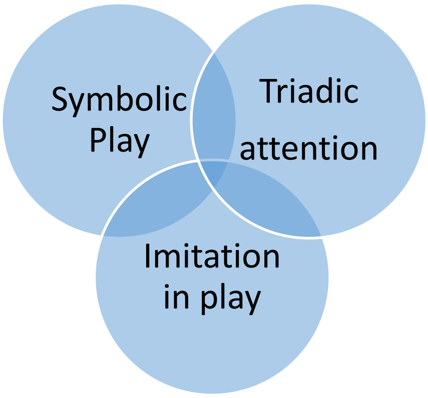
What proportion of children are identified as needing extra language support?	Estimates were as high as 20-24% and varied by class. Some teachers were less sure or said it varied year on year.
How are interventions/strategies delivered? (Group size, frequency, staff member etc?	1:1 or in small groups/whole classes, depending on the nature of the intervention. In one school the class teacher delivers BLAST because she 'happened to be the one to attend the training'. Play based, small group delivery In one school staff try to deliver BLAST daily to all children, but this is sometimes not possible due to timings, so they focus mainly on the children with greater need Teacher, SENCO, TA or SLT Assistant deliver interventions Generally daily for 10-15 minutes. Assigned staff member delivers programmes sent by SLT
What is the Speech Therapist's input?	Children are highlighted then referred to SLT if deemed appropriate. The language and communication team (NHS) used to do 1:1 work with children, but are now mainly advising teachers and training them to do the interventions instead. Children are also identified through teacher's intuition and observed social interaction. In some schools the SLT works 1:1 with particular children and, leaves programmes of work to be completed with children. SLT assesses the child (the referral can take a while). The SLT leaves a 6 week programme of work to be delivered by an SLTA, then the SLT reassesses the child. Teachers are given resources and continue to deliver various things SLTs also train teachers to deliver the interventions, and work fortnightly 1:1 with particular children. One school said input was very minimal
If SLTs leave activities for children who delivers them and how often?	Teachers, TAs, SLTAs. Depends on availability and training. Usually it is one or several assigned members of staff.

How do you judge the level of need in the classroom?	Screening, low SES, intuition with new intakes, through assessment against the profile, sharing knowledge in the Early Years Team

Name: **dob:** **NHS/ID:**	**Address:**

To benefit from the BEST approach a child needs to have the following skills and language

Skills



been

Expressive Language

No recognisable single words

or

Single words heard in clinic or reported at home

or

2 – 3 word utterances with a maximum of 2 clause elements (e.g. SV; VO)

BEST Eligibility Assessment Summary

<table>
<tr><th>Skill</th><th colspan="2">Passes Criterion? Y/N</th><th>Date</th></tr>
<tr><td>Triadic attention</td><td colspan="2"></td><td></td></tr>
<tr><td>Symbolic Play</td><td colspan="2"></td><td></td></tr>
<tr><td>Imitation in play</td><td colspan="2"></td><td></td></tr>
<tr><td>Expressive Language</td><td colspan="2"></td><td></td></tr>
<tr><td>Hearing Test Result</td><td colspan="2"></td><td></td></tr>
<tr><td>Eligible for BEST?
*all 5 criteria above must be met</td><td colspan="3"></td></tr>
<tr><td>Recommended Pathway</td><td>BEST</td><td>Hearing Test then BEST</td><td>Other (please specify)</td></tr>
</table>

Outcome ~ Time Point + Intervention Arm + SDQ score + Dosage (n) + Language Profile + Age + Intervention Arm : Language Profile + Intervention Arm : SDQ score

random = list(~ OUTCOMETime + Dosage + AgeInMonthsAtEntry + SDQscore + LanguageProfile | School,

~ OUTCOMETime + SDQscore + LanguageProfile | PptNo), data = bestLong, na.action = na.omit,

method="ML", control = list(opt="optim", correlation = corCAR1()))

Appendix 18 Description of Variables Included in Multi-Level Models

Variable name	Description	Time Invariant/Time variant	Notes
DataWave		Time Invariant	1 = data wave 1 2 = data wave 2 3 = data wave 3
InterventionArm		Time Invariant	0 = Treatment as Usual 1 = BEST without sign 2 = BEST with sign
PooledTreatment		Time Invariant	0 = Treatment as Usual 1 = BEST (with and without sign)
PptNo		Time Invariant	Numeric participant differentiator
ChildCode		Time Invariant	String participant differentiator
LanguageProfile	Expressive/Receptive/ Mixed language profile	Time Invariant	Assigned based on falling on or below the 16^{th} centile for expressive and/or receptive language
AgeInMonthsAtEntry	Child's age in months when entering the study	Time Invariant	Calculated from date of birth data at the time of entry to the relevant data wave

Dosage	Number of sessions received		0 = TAU 1-16 = Sessions received out of a total 16
School	Code for school number		
SEBD	Social, emotional and behavioural difficulties	Time Invariant	Total difficulties score (total score minus prosocial score) out of a possible 40 calculated from the teacher's version of the Strengths and difficulties questionnaire (Robert Goodman, 1997)
comprehensionScore_Centred	Comprehension standard scores	Time Variant	Comprehension standard scores at T0, T1 and T2 calculated from the NRDLS (Edwards et al., 2011)
productionScore_Centred		Time Variant	Production standard scores at T0, T1 and T2 calculated from the NRDLS (Edwards et al., 2011)
ContentCorrect_Centred		Time Variant	Content correct percentage score at T0, T1 and T2 rounded to 2 decimal places
MorphologyCorrect_Centred		Time Variant	Morphology correct percentage score at T0, T1 and T2 Rounded to 2 decimal places

FOCUSScore_Centred	Time Variant	Total FOCUS score (total of part 1 and part 2) calculated from the teacher's version of the Focus on the Outcomes of Communication Under Six 36 item questionnaire (Thomas-Stonell, N Robertson et al., 2012)
Comp0, Comp1, Comp2	Time Variant	Timepoints 0, 1 and 2 for comprehension scores
Prod0, Prod1, Prod2	Time Variant	Timepoints 0, 1 and 2 for production scores
Content0, Content1, Content2	Time Variant	Timepoints 0, 1 and 2 for content scores
Morph0, Morph1, Morph2	Time Variant	Timepoints 0, 1 and 2 for morphology scores
FOCUS0, FOCUS1, FOCUS2	Time Variant	Timepoints 0, 1 and 2 for FOCUS scores

Appendix 19 Rationale for included variables

The first key variable is DataWave, which represents whether children were in data wave 1, 2, or 3. This is a time invariant numeric variable and was not expected to impact scores, since all data waves were kept as similar as possible. It was still important to control for data wave, in case factors such as the researcher improving BEST delivery over the course of the study, or different RAs carrying out the assessments, had a systematic impact on children's scores.

The next variable entered into the model is InterventionArm, representing the therapy that children received. This is a theoretically essential factor with 3 levels, representing whether children received BEST with sign, BEST without sign, or were in the TAU arm. Similarly the factor PooledTreatment has 2 levels to code for therapy or no therapy, without accounting for the difference between the sign or no sign condition. These are both time invariant variables since children remain in the same treatment arm throughout the study.

The factor PptNo codes for each individual child, and again is a time invariant factor. ChildCode is the string data version which represents the same information. These are included to separate the children from eachother.

The 3 level factor LanguageProfile describes whether children have expressive, receptive, or mixed (expressive and receptive) language difficulties as determined by their scored on the NRDLS. This is treated as a time invariant factor calculated from scores at baseline. This factor is included to establish any variance in scores due to children's differing baseline language profiles.

AgeInMonthsAtEntry is a time invariant numeric variable for age calculated from the child's date of birth at the time of entry to the relevant data wave.

Dosage is the number of BEST sessions completed by the child out of a total of 16, and is recorded as a numeric variable. 0 sessions indicates that the child was assigned to the Treatment as Usual arm. Importantly, no children who were in the BEST with sign or BEST without sign arms received 0 sessions.

School is a factor with 13 levels representing each of the schools involved in the trial. It was important to differentiate between the schools to control for any systematic differences between the children's environments.

SEBD is a time invariant numeric variable calculated from the Strengths and Difficulties Questionnaire completed at baseline. The teacher's scores were used due to the poor return rate of the parent forms. The total difficulties score out of a possible 40 was (total score minus prosocial score) was calculated (Robert Goodman, 1997). SEBD was included to establish any associations between underlying social, emotional and behavioural difficulties and progress made on outcome measures over the course of the study.

CompTime, ProdTime, ContentTime, MorphTime, and FOCUSTime are time variant factors with three levels each (0, 1, 2). These represent the 3 timepoints at which each measurement was taken.

comprehensionScore_Centred, productionScore_Centred, ContentCorrect_Centred, MorphologyCorrect_Centred and FOCUSScore_Centred are all time variant numeric variables indicating the standard scores on the NRDLS for comprehension and production, the percentage correct scores for content and morphology on the BEST Assessments, and the teacher total FOCUS score at each of the 3 timepoints (0, 1, 2). These are the outcome variables used to address the research questions at the beginning of this chapter.

Kruskall-Wallace tests of difference at baseline across treatment arm

Kruskal-Wallis tests of difference at baseline across treatment arm for non-parametric distributions			
	Kruskal-Wallis chi-square	df	p-value
IDACI Score	2.23	2	0.328
CSCOT Proportion Total Score	9.21	2	0.01*

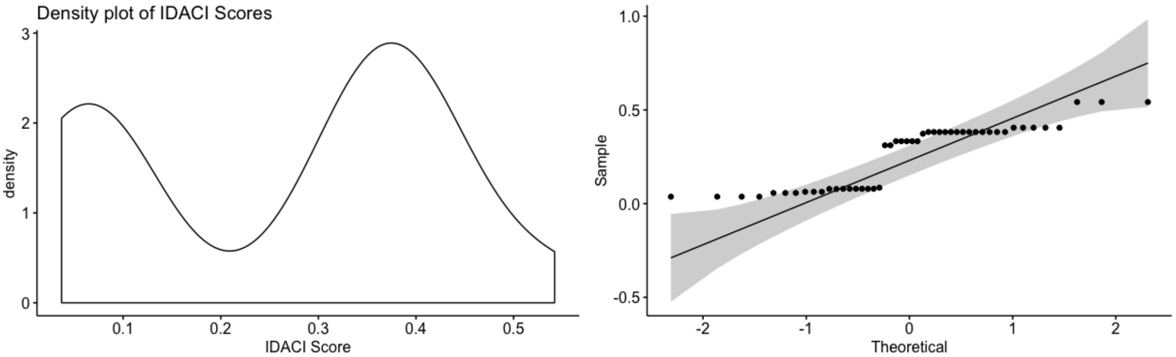
Kruskal-Wallis tests of difference conducted on non-parametric outcomes across intervention arms

Kruskall-Wallace tests of difference at baseline across data wave

Kruskal-Wallis tests of difference at baseline across data wave for non-parametric distributions				
	Kruskal-Wallis chi-square	df	p-value	conclusion
IDACI Score	1.11	2	0.5742	No significant difference between groups
IDACI Decile	3.67	2	0.16	No significant difference between groups
CSCOT Proportion Total Score	21.36	2	<0.001	Significant difference present

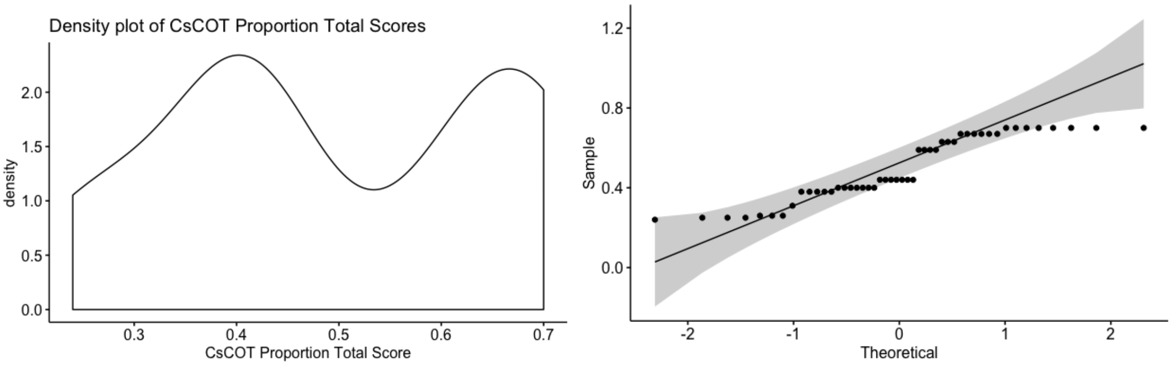
Kruskal-Wallis tests of difference conducted on non-parametric outcomes across data waves

K-W tests showing non-normal distribution of the data



Density plot and qq (quantile-quantile) plot for IDACI score

Density and qq plots suggest that the data are not normally distributed for IDACI score. A Shapiro-Wilk normality test confirmed that the data differs significantly from a normal distribution (W=0.81, p<0.001).



Density plot and qq (quantile-quantile) plot for IDACI score

Density and qq plots suggest that the data are not normally distributed for CSCOT proportion total score. A Shapiro-Wilk normality test confirmed that the data differs significantly from a normal distribution (W=0.88, p<0.001).

Kruskall-Wallace (K-W) is the non-parametric test of difference equivalent to ANOVA. ANOVAs assess means, whilst the K-W test uses medians. Both will be presented for comparison as there are known issues with the K-W test (Ruxton & Beauchamp, 2008).

Descriptive statistics across data wave and school

The table above shows the baseline descriptive statistics for age, IDACI score and decile, and CSCOT Proportion total scores across the total sample and the three data waves.

The IDACI score, decile, and CSCOT total proportion score appear to show more variation. A higher IDACI score, and lower IDACI rank indicates a postcode area of higher deprivation. The study was designed using quasi-experimental methods to produce three matched treatment arms (see chapter 2 for more information), therefore it was anticipated that all groups would be equal at baseline. Analysis of the three data waves show that the average level of deprivation increased across each. This is not necessarily an issue, since we see in the descriptive table for intervention arm above that the thee arms were balanced across the whole study. The CSCOT scores were significantly different at baseline for BEST without sign and TAU. The differences across data wave may explain this, as the average CSCOT scores decreased at each data wave, however further work would be needed to determine this.

It is not clear whether the CSCOT scores decreased at each data wave due to the increasing average level of deprivation, or because of poor reliability with the measure. This is discussed in more depth in chapter 5. A between-groups ANOVA is required to establish whether the differences are significant.

A between-groups ANOVA confirmed that there was a significant difference between the dosage received across data waves $F(2, 45)= 43.96$, $p = <.001$. Visual inspection of the mean dosage indicated that the third wave was responsible for the difference, since it was a treatment as usual only wave. Tukey's HSD post-hoc testing confirmed that wave one and wave three differed significantly ($p = <.001$) and wave two and wave three differed significantly ($p = <.001$). There was no significant difference between waves one and two ($p = >.71$).

There were no significant differences between data waves for age [$F(2,45) = 0.187$, $p = .84$] or IDACI score [$F(2,45) = 2.226$, $p = .12$]. There was a significant difference between groups for the CSCOT proportion total score [$F(2.45) = 20.65$, $p = <.001$]. Post-hoc tests were therefore carried out to establish where the differences lie for CSCOT proportion total scores. A Tukey's HSD was carried out for CSCOT proportion total score which revealed a significant difference between data wave 1 and 2 ($p = <.001$), and a significant difference between data wave 1 and 3 ($p =<.001$). The difference between data wave 2 and 3 was non-significant ($p = >.144$).

It can be assumed that the IDACI score and rank do not differ significantly across intervention arms, while CSCOT score does.

Descriptive statistics across Full Sample and Individual Data Waves (Wave one, two and three)

	Full Sample	Data Wave			Significant Difference Between Groups[e]
		One	Two	Three	
N	48	23	14	11	
Sex					
Male	28	11	11	6	
Female	20	12	3	5	
Language Profile[a]					
Expressive	15	8	5	2	
Receptive	5	4	1	-	
Mixed	28	11	8	9	
Age in months at entry	46.96 (6.13)	46.57(6.30)	46.86(6.97)	47.91(4.97)	-
School socio-economic status[b]	0.26 (0.17)	0.22(0.17)	0.26(0.16)	0.35(0.16)	-
Classroom oral language environment[c]	0.49 (0.16)	0.60(0.12)	0.43(0.15)	0.34(0.08)	*

Dosage[d]	9.52 (6.47)	11.96 (5.24)	13 (2.22)	-	*

Note. Table displays the n, means and standard deviations for children's sex, language profile, age, school IDACI score, CSCOT total proportion score and dosage across the three data waves.

[a] Mixed language profile defined as scoring on or below the 16th centile of the production *and* comprehension subscales of the NRDLS. Expressive-only language difficulties defined as scoring on or below the 16th centile of the production subscale only. Receptive-only difficulties defined as scoring on or below the 16th centile of the comprehension subscale only.

[b] Income Deprivation Affecting Children Index (IDACI) Score. IDACI score was calculated using the school postcode at: https://lginform.local.gov.uk/reports/lgastandard?mod-metric=3910&mod-area=E92000001&mod-group=AllRegions_England&mod-type=namedComparisonGroup

A higher IDACI score indicates a postcode area of higher deprivation (see the measures section of chapter 2 for more information) (Department of Communities and Local Government, 2015).

[c] Communication Supporting Classrooms Observation Tool (CSCOT) Proportion Total Score (Dockrell et l., 2012). Total proportion score calculated from a one hour classroom observation.

[d] Dosage refers to the average number of BEST sessions received out of a total possible 16. Children in the TAU arm do not have a dosage score as they did not receive the intervention.

[e] Significant difference between groups (BEST with sign, BEST without sign, TAU) indicated by *. Calculated using ANOVA tests of difference, reported below.

The table below presents descriptive statistics for number of children in each school, IDACI score and decile, CSCOT proportion total score, mean age in months at entry to the study and mean dosage for each school. Schools are not expected to be individually similar, however the groups (intervention arms) are statistically equivalent across the covariates (with the exception of CSCOT score and dosage, as discussed above). This table demonstrates the range of schools involved in the study. The table also provides the breakdown of dosage by school, showing that all children in the BEST conditions received 11-16 BEST sessions (with or without sign). Dosage is addressed in the Dosage Analysis section below.

Descriptive statistics across schools by treatment arm (BEST with sign, BEST without sign and Treatment as Usual)

	School grouped by treatment arm												
	BEST with sign (n = 4)				BEST without sign (n = 4)				Treatment as Usual (n = 5)				
	School 4	School 6	School 8	School 9	School 1	School 2	School 5	School 7	School 3	School 10	School 11	School 12	School 13
N	8	5	3	1	5	6	4	2	3	2	3	5	1
IDACI Score[a]	0.38	0.38	0.06	0.09	0.41	0.08	0.04	0.31	0.06	0.08	0.54	0.33	0.37
CSCOT Proportion Total Score[b]	0.70	0.44	0.26	0.31	0.40	0.67	0.59	0.40	0.63	0.44	0.25	0.38	0.24
Age in months at entry	42.38 (6.61)	50.40 (5.94)	50.00 (5.29)	52.00	50.00 (3.74)	48.67 (6.09)	47.25 (5.85)	40.00 (4.24)	41.00 (6.93)	47.00 (5.66)	50.00 (6.11)	47.20 (5.12)	45.00
Dosage[c]	13.63 (2.88)	14.60 (1.14)	11.33 (2.08)	16.00	12.40 (2.30)	14.67 (2.34)	12.25 (1.71)	13.00 (1.41)	-	-	-	-	-

Note. Table displays the n, means and standard deviations for school IDACI score, age and CSCOT total proportion score across the 13 schools.

[a] Income Deprivation Affecting Children Index (IDACI) Score. IDACI score was calculated using the school postcode at: https://lginform.local.gov.uk/reports/lgastandard?mod-metric=3910&mod-area=E92000001&mod-group=AllRegions_England&mod-type=namedComparisonGroup

A higher IDACI score indicates a postcode area of higher deprivation (see the measures section of chapter 2 for more information) (Department of Communities and Local Government, 2015).

[b] Communication Supporting Classrooms Observation Tool (CSCOT) Proportion Total Score (Dockrell et l., 2012). Total proportion score calculated from a one hour classroom observation.

[c] Dosage refers to the average number of BEST sessions received out of a total possible 16. Children in the TAU arm do not have a dosage score as they did not receive the intervention.

Appendix 21 ANCOVA Alternative Analysis

This study initially sought to conduct analysis using Longitudinal Multi-Level Models to account for the clustered nature of the data, where timepoints are clustered within participants who are in turn clustered within schools. The analysis was completed however, possibly due to the study being underpowered, a number of the models would not converge and the models that did converge were difficult to interpret due to the absence of others. In order to answer the research questions as thoroughly as possible, an alternative method of analysis was used. Despite this, the application of Multi-Level Models is an important part of the piloting and feasibility trial as it demonstrates the necessity of sufficient power and completion of measures. It is therefore included at the end of this chapter, and discussed in chapter 5.

There is an ongoing debate in the statistical literature about the merits and weaknesses of performing gain score ANOVAs, where gain score is calculated by subtracting T0 from T1, compared to performing ANCOVA on outcome scores while including baseline scores as a covariate in the model (Knapp & Schafer, 2009; Van Breukelen, 2006). Both analyses have been carried out, and are presented below, followed by the results of longitudinal multi-level modelling. The following mean change score analysis was carried out to provide an alternative view of the data. It has its own caveats: mainly that analyses of change scores are a crude representation of the underlying variance in the data. The following ANOVA and ANCOVA analyses could add value if taken alongside the graphs presented throughout this chapter and the conservative MLMs, and interpreted with caution.

In order to test whether the gain scores achieved by children in the differing treatment arms differed from each other, mean change scores were first calculated for each outcome measure (production, comprehension, content and morphology scores) between:

1. Baseline and outcome
2. Baseline and follow-up
3. Outcome and follow-up

Gain scores were calculated by subtracting the baseline from the outcome, the baseline from the follow-up and the outcome from the follow-up respectively. Analysis of Variance (ANOVA) was carried out on gain scores for each measure to identify whether any differences were present between the scores. Since the ANOVA results do not indicate where the differences lie, Tukey HSD post-hoc tests from the R package Stats (R Core Team, 2019). Gain scores for each timepoint and the corresponding results of the ANOVA analyses are presented below.

Descriptive statistics for change scores

	TAU N=14	BEST without sign N=17	BEST with sign N=17
Production Outcome Change Score	0.43(7.05)	5.11(7.37)	10.11(10.36)
Comprehension Outcome Change Score	2.57(9.86)	6.41(11.71)	9.18(9.64)
Content Outcome Change Score	0.87(18.06)	34.24(20.33)	32.05(12.99)
Morphology Outcome Change Score	12.29(16.30)	34.51(23.98)	35.55(19.57)
APV Content Outcome Change Score	5.30(13.83)	22.92(11.78)	27.53(13.45)
APV Morphology Outcome Change Score	7.33(17.37)	32.84(27.44)	35.44(19.49)
ELP Content Outcome Change Score	0.75(13.06)	8.19(15.62)	9.59(10.61)
ELP Morphology Outcome Change Score	13.43(14.61)	20.99(26.38)	23.80(17.08)
Age in months at entry	46.43(5.92)	47.71(5.67)	46.65(6.99)
IDACI Score	0.29(0.19)	0.19(0.16)	0.31(0.14)
IDACI Decile	4.14(3.48)	5.71(3.90)	3.59(2.96)
CSCOT Proportion Total Score	0.40(0.14)	0.54(0.12)	0.52(0.18)

Means and Standard Deviations for change scores across the three treatment arms

	Change Scores						
Outcome Measure and Time period	TAU	BEST without sign	BEST with sign	df	Residual df	F value	P value
Production							
Baseline -Outcome	0.43	5.12	10.12	2	45	5.049	0.01
Baseline – Follow-Up	3.71	4.71	12.75	2	44	3.464	0.04
Outcome- Follow-Up	3.29	-0.41	2.19	2	44	0.717	0.495
Comprehension							

Baseline -Outcome	2.57	6.41	9.18	2	45	1.527	0.228
Baseline – Follow-Up	6.50	5.88	10.81	2	44	0.819	0.447
Outcome- Follow-Up	3.92	-0.53	1.19	2	44	0.87	0.426
BEST Assessment Content							
Baseline -Outcome	0.87	34.24	32.05	2	43	16.88	<0.001
Baseline – Follow-Up	9.65	29.98	29.34	2	39	5.81	<0.01
Outcome- Follow-Up	8.78	-3.39	-2.89	2	40	3.821	0.03
BEST Assessment Morphology							
Baseline -Outcome	12.29	34.51	35.55	2	43	6.086	<0.01
Baseline – Follow-Up	19.28	37.07	34.19	2	39	2.639	0.0841
Outcome- Follow-Up	6.99	3.42	-1.30	2	40	0.663	0.521

Mean change scores between: Baseline and Outcome, Baseline and Follow-up, Outcome and Follow-up by intervention arm arm ANOVAs were carried out

Post-hoc tests were carried out on the mean change scores to identify where the significant differences lay. The mean score differences and p-values are reported. For production and comprehension the unit is scores on the NRDLS (indicated by points in the text below), and for sentence content and sentence morphology the unit is percentage.

NRDLS Production standard scores

Between baseline and outcome mean change scores were significantly higher for BEST with sign than TAU (mean difference = 9.69 points, p = <0.01). The mean change score was higher for BEST without signthan TAU, but these did not differ significantly (mean difference = 4.69 points, p=0.29). The mean change score was higher for BEST with sign than BEST without signbut these did not differ significantly (mean difference = 5 points, p = 0.21).

Between baseline and follow-up the mean change score was higher for BEST with sign than TAU, which approached significance (mean difference = 9.04 points, p = 0.06) however this was not statistically significant at the 0.05 cut-off level. The mean change score was higher for BEST without signthan TAU but did not differ significantly (mean difference = 0.99 points, p = 0.96), and the mean change score was higher for BEST with sign than BEST without signbut did not differ significantly (mean difference = 8.04 points, p = 0.08).

Change scores from outcome to follow-up were non-significantly different for any of the intervention groups, thus did not warrant post-hoc testing.

NRDLS Comprehension

There were no significant differences between change scores for any intervention groups between any of the time periods for comprehension scores on the NRDLS, thus no further testing for was warranted.

BEST Assessment Content Score

Between baseline and outcome mean change scores were significantly higher for BEST with sign than TAU (mean difference = 31.18%, p = <0.001). The mean change scores were significantly higher for BEST without sign than TAU (mean difference = 33.38%, p = <0.001). The mean change score was higher for BEST without sign than BEST with sign, but these did not differ significantly (mean difference = 2.2%, p = 0.93).

Between baseline and follow-up mean change scores were significantly higher for BEST with sign than TAU (mean difference = 19.69%, p = 0.02). The mean change scores were significantly higher for BEST without sign than TAU (mean difference = 20.34%, p = 0.01). The mean change score was higher for BEST without sign than BEST with sign, but these did not differ significantly (mean difference = 0.65%, p = >0.99).

Between outcome and follow-up mean change scores were significantly higher for TAU than for BEST without sign(mean difference = 12.17%, p = 0.04). The mean change scores were higher for TAU than BEST with sign (mean difference = 11.67%, p = 0.07). The mean change score was higher for BEST without sign than BEST with sign, which approached (mean difference = 0.65%, p = >0.99), however this was not statistically significant at the 0.05 cut-off level. The mean change score was higher for BEST with sign than BEST without sign but these did not differ significantly (mean difference = 0.5%, p = >0.99).

Morphology

Between baseline and outcome mean change scores were significantly higher for BEST with sign than TAU (mean difference = 23.26%, p = <0.01). The mean change scores were significantly higher for BEST without sign than TAU (mean difference = 22.22%, p = 0.01). The mean change score was higher for BEST with sign than BEST without sign but these did not differ significantly (mean difference = 1.04%, p = 0.99).

Change scores between baseline to follow-up, and between outcome to follow-up were non-significantly different thus did not warrant further testing.

Mean Change Score Analysis: Conclusion

The change score analysis supports the graphical trends displayed throughout this chapter. The mean scores for children in the BEST arms show significant increases compared to the means for the Treatment as Usual arm on the NRDLS production subscale, and sentence content percentage and sentence morphology percentage on the BEST Assessment. Post-hoc tests indicate that the mean change score for children in the BEST with sign arm was significantly higher than TAU on production scores between baseline and outcome. The mean sentence content change scores were significantly higher in both BEST arms than TAU between baseline and outcome, and between baseline and follow-up. The mean sentence morphology change scores for both BEST arms were significantly higher than TAU between baseline and outcome.

The mean content change scores were significantly higher for children in TAU than in the BEST arms between outcome and follow-up, however the follow-up TAU score still did not reach that of the BEST arms, and this is likely due to ceiling effects preventing the mean from increasing at Follow-up due to already high scores at Outcome for children in the BEST arms. This is discussed further in chapter 5.

Fitting the ANCOVA Models

Testing the Assumption that Covariates are not Highly Correlated

	AgeInMonthsAtEntry	SEBD	cscotTotal	IDACIscore	LanguageProfile	Dosage
AgeInMonthsAtEntry	1	0.10	-0.31	0.02	0.23	0.02

SEBD	0.10	1	0.09	0.45	0.41	-0.03
cscotTotal	-0.31	0.09	1	-0.28	-0.10	0.40
IDACIscore	0.02	0.45	-0.28	1	0.21	-0.08
LanguageProfilev	0.23	0.41	-0.10	0.21	1	-0.23
Dosage	0.02	-0.03	0.40	-0.08	-0.23	1

Correlation matrix between potential covariates

A Pearson's correlation matrix suggests that no covariates are highly correlated, and this assumption of the ANCOVA is therefore met.

Testing the Assumption that Residuals are Normally Distributed

Levene's tests were used to check the assumption of homogeneity of variance for outcomes across the three treatment arms.

Outcome	df	F Value	P-Value[1]
Production Outcome Change Score	2	1.04	0.36
Comprehension Outcome Change Score	2	0.15	0.86
Content Outcome Change Score	2	0.91	0.41
Morphology Outcome Change Score	2	1.81	0.18
ELP Content Outcome Change Score	2	0.15	0.86
ELP Morphology Outcome Change Score	2	1.24	0.30

[1] A p-value above 0.05 indicates homogeneity of variance can be assumed. A p-value below 0.05 indicates the assumption of homogeneity of variance has been violated.

Levene's tests were used to check the assumption of homogeneity of variance for covariates across the three treatment arms.

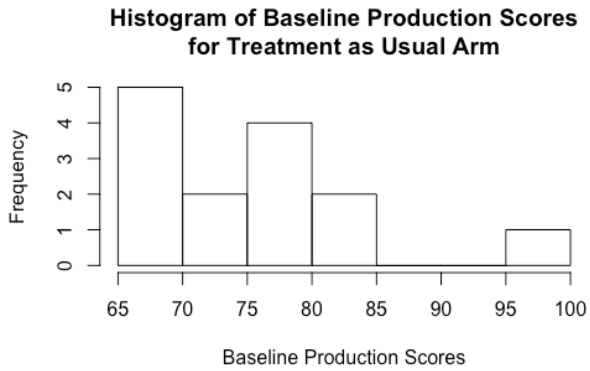
Covariate	df	F Value	P-Value[1www]
Age	2	0.56	0.58
IDACI Score	2	1.136	0.33
Language Profile	2	0.74	0.48
CSCOT Score	2	1.95	0.15
Dosage	2	8.14	<0.001

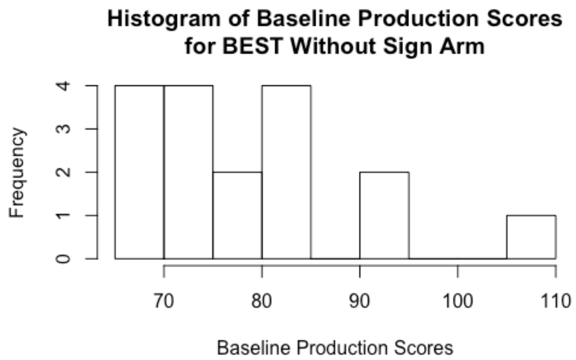
[1] A p-value above 0.05 indicates homogeneity of variance can be assumed. A p-value below 0.05 indicates the assumption of homogeneity of variance has been violated.

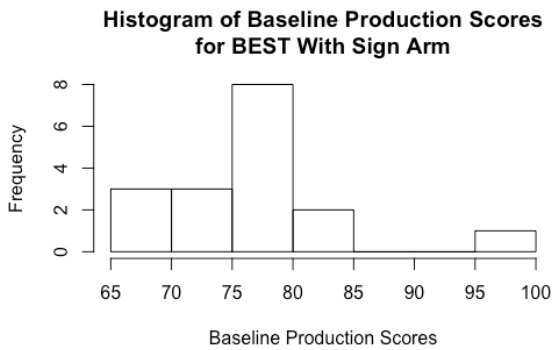
The covariates are all suitable for use in the ANCOVA model apart from dosage (due to children either receiving relatively high dosage or zero for TAU). A separate analysis was therefore carried out to explore the role of dosage on change scores between baseline and outcome.

Checking the distributions for baseline scores

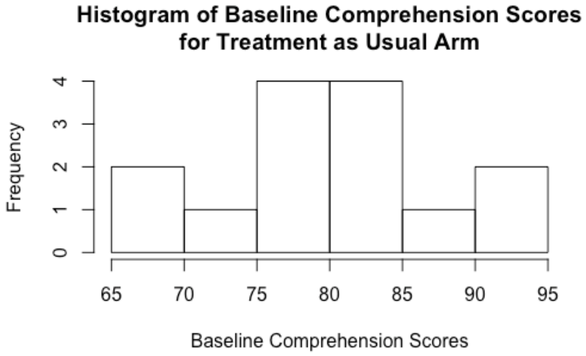
The distribution of the baseline scores were checked to ensure that they were similar at baseline. They showed the expected pattern for production and comprehension whereby the lowest score was 70, and there were several high scores, with the majority of scores between 65-75.

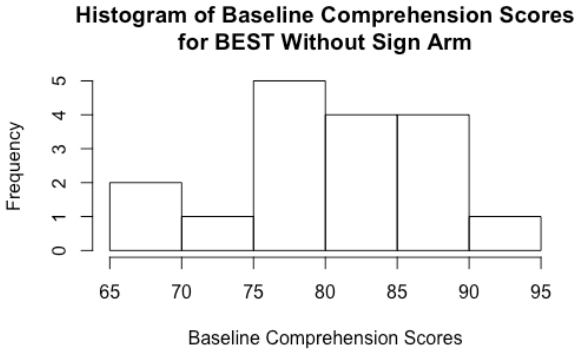


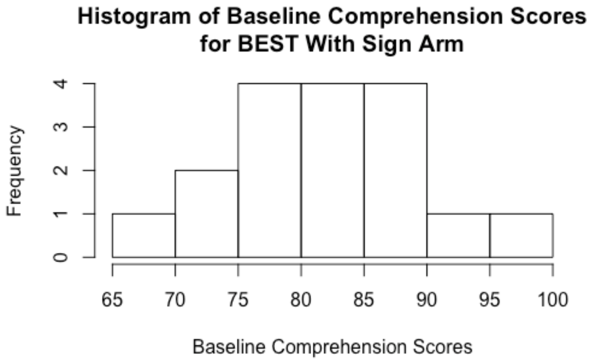




Histograms for TAU, BEST without sign and BEST with sign for comprehension scores







Histograms for TAU, BEST without sign and BEST with sign for comprehension scores

ANOVAs were conducted to ensure that the predictor variable (Intervention Arm) and covariates (Age, CSCOT score, Language Profile, IDACI Score) were independent.

The average age in months at entry did not differ significantly across the three treatment arms at baseline $F(2,45) = 1.94$, $p = 0.83$. The average CSCOT Score differed significantly across the three treatment arms $F(2,45) = 3.51$, $p = 0.04$. Language profile did not differ significantly across the three treatment arms $F(2, 45) = 0.68$, 0.51. The average IDACI Score did not differ significantly across the intervention arms $F(2, 45) = 2.38$, $p = 0.11$.

The models will therefore exclude dosage and CSCOT Scores since they are not independent from the predictor variable, and dosage as it there is not homogeneity of variance across the treatment arm

ANCOVA Type III sum of squares (SS) was used as the data are unbalanced (there are 14 participants in TAU and 17 in each BEST arm) and because there was a possibility of interactions in the results.

As type III SS was used, the order of the model specification was not important. The interaction is stipulated by the *.

Helmert orthogonal contrasts were employed in order that the type III sum of squares (SS) model could be

computed correctly. For the first contrast the TAU group was coded as -2 and each BEST group was coded as 1. For the second contrast TAU was coded as 0, BEST without signwas coded as -1 an BEST with sign was coded as 1. These follow the conventions of helmert contrasts, and result in the same outputs. The models were run for each of the dependent variables, which are reported below under the relevant section headings. Adjusted means and standard errors were produced for each ANCOVA model. Planned contrasts which ask more focussed questions of the data (Rosnow et al., 2000)

Effect sizes were calculated in accordance with Rosnow et al. (2000) using the equation

$$r_{\text{effect size}} = \sqrt{\frac{F_{\text{contrast}}}{F_{\text{contrast}} + F_{\text{noncontrast}}(df_{\text{noncontrast}}) + df_{\text{within}}}}$$

Ppendix

Two separate analyses were completed. The first analysis used dummy variables to code for intervention arm (TAU/ BEST without sign/ BEST with sign) and language profile (espressive/receptive/mixed). The model was specified as follows:

```
comp_Anova <- aov(CompOutcomeChangeScore~
            IDACIscore+
            AgeInMonths(centred)+
            DummyWithSign+
            DummyWithoutSign)
```

The intercept represented TAU expressive language profile. A benefit of the dummy coded variable model is that the results were more transparent than when using Helmert planned contrasts.

Analysis of Outcome Scores Controlling for Baseline Scores and Covariates: ANCOVA

The role of the ANCOVA is to account for variables that are not necessarily the focus of the main research question, but may still have an effect on the dependent variable. In the present research the independent variable is Intervention Arm (TAU/ BEST without sign/ BEST with sign), and the main covariates of concern are Age, SEBD, CSCOT score, IDACI score, Language profile (dummy variable) and dosage. The ANCOVA is an extension of the ANOVA which controls for variance in the covariates, rather than assuming groups are the same at baseline, as is the case with ANOVA.

The assumptions of the ANCOVA are that covariates are not highly correlated ($r>0.8$), residuals are normally distributed and that there is homogeneity of variance for all groups.

ANCOVAS are carried out on NRDLS, BEST Assessment and Generalised BEST Assessment outcome measures. Only 50% of children had complete data for the Alternative Picture Version of the BEST Assessment due to its late introduction into the study, and it was therefore not examined in the ANCOVA analysis.

Intervention Arm Analysis

The second analysis compared TAU with BEST (with and without sign, pooled), and BEST with sign and BEST without Sign.

The model was specified as follows:

```
contrasts(best$InterventionArm) <- cbind(c(-2,1,1), c(0, -1, 1))

ProdANCOVA <- aov(best$ProdOutcomeChangeScore~

            best$InterventionArm*best$LanguageProfile+

            best$AgeInMonthsAtEntry+

            best$IDACIscore)

Anova(ProdANCOVA, type = "III")
```

Outcomes: ANCOVA Results

Regression tables were created to interpret the effect of intervention arm on outcome variable scores accounting for baseline scores, IDACI score, age and language profile. The intercept value represents a child in the TAU arm with a mixed language profile. Scores were centred. This number is displayed in the top left of the regression table.

To calculate the change scores for children in differing intervention arms and with differing language profiles the betas (coefficient estimates) were added for the effect in question, and the interaction coefficient was also added when both the intervention arm and language profile varied (these cells are not shaded in the table).

Production and comprehension scores were centred to ease their interpretability. Content and Morphology scores were not centred as they are easier to interpret as percentage scores.

Age and IDACI scores were also centred so the models all assume a whole sample average age and IDACI score to aid comparability.

The intercept change score was calculated using the equation

Y = intercept + IDACI estimate + Age in Months estimate

The change score estimates were calculated using the equation:

Y = intercept + IDACI estimate + Age in Months estimate + language profile estimate + Intervention Arm Estimate + Interaction Estimate if Applicable

ANCOVA Models (Production, Comprehension, Content, Morphology)

ANCOVA Results for Production, Comprehension, Content and Morphology Outcomes (Coefficient Estimates, t-values and p-values)

	Regression models			
	Production Model	Comprehension Model	Content Model	Morphology Model
	(1)	(2)	(3)	(4)
Intercept	-5.113**	-3.686	35.781***	28.343***
	(2.342)	(2.780)	(10.963)	(7.967)
Production Score	0.823***			
	(0.171)			
Comprehension Score		0.560**		
		(0.217)		
Content Score			0.303	
			(0.202)	
Morphology Score				0.590***
				(0.162)
IDACI Score (centred)	-8.384	-5.858	-16.752	-23.812
	(8.377)	(9.885)	(16.683)	(19.563)
Dosage	0.050		-0.310	-0.314
	(0.703)		(1.201)	(1.445)
Age in Months (centred)	-0.039	0.029	0.659	0.858*
	(0.211)	(0.248)	(0.397)	(0.472)

BEST with Sign	9.289	7.248*	32.908*	28.208
	(10.046)	(3.753)	(17.381)	(20.812)
BEST without Sign	3.794	3.158	31.270*	22.239
	(9.801)	(3.867)	(17.442)	(20.632)
Observations	48	48	46	46
R^2	0.527	0.235	0.514	0.505
Adjusted R^2	0.458	0.144	0.439	0.429
Residual Std. Error	8.697 (= 41)	10.347 (df = 42)	15.607 (df = 39)	18.940 (df = 39)
F Statistic	7.626*** (df = 6; 41) (p = 0.00002)	2.586** (df = 5; 42) (p = 0.040)	6.879*** (df = 6; 39) (p = 0.00005)	6.624*** (df = 6; 39) (p = 0.0001)

Note: *p<0.1; **p<0.05; ***p<0.01

Production

The covariate baseline score was significantly related to the outcome score $F(1, 36) = 20.95$, $p = <0.001$. There was a significant effect of intervention arm after controlling for the covariates $F(2, 36) = 6.03$, $p = <0.01$. There was a significant interaction between intervention arm and language profile $F(4, 36) = 4.85$, $p = <0.01$.

Planned contrasts using the dummy variable model revealed that the interaction of BEST with sign and receptive language profile significantly increased production scores $t(1, 36) = 2.11$, $p = 0.04$. The interaction between BEST with sign and Expressive language approached significance and decreased production scores $t(1, 36) = -1.88$, $p = 0.07$. BEST with sign significantly increased production scores compared to the reference $t(1, 36) = 3.30$, $p = <0.01$. BEST without signsignificantly increased production scores compared to the reference $t(1, 36) = 2.43$, $p = 0.02$. Expressive language profile significantly increased production scores compared to the reference $t(1,36) = 2.02$, $p = 0.03$.

Comprehension

Intervention arm was non-significantly related to comprehension outcome scores after controlling for covariates $F(2, 36) = 0.25$, $p = 0.78$. There were no other significant predictors of comprehension scores.

Content

There was a significant effect of intervention arm for content percentage scores after controlling for covariates $F(2, 34) = 8.96$, p

$= <0.001$. The covariate age in months at entry approached significance for content percentage scores $F(1, 34) = 3.77$, $p = 0.06$.

Planned contrasts using the dummy variable model revealed that BEST with sign significantly increased content scores compared to the reference group $t(1, 34) = 4.18$, $p = <0.001$. BEST without signalso significantly increased content percentage scores compared to the reference group $t(1, 34) = 3.41$, $p = 0.02$.

Morphology

The covariate baseline score was significantly related to the outcome morphology percentage score $F(1, 34) = 6.49$, $p = 0.02$. The covariate age in months at entry was also significantly related to the outcome morphology percentage score $F(1, 34) = 4.43$, $p = 0.04$. There was a significant effect of intervention arm after controlling for covariates $F(2, 34) = 7.93$, $p = 0.001$.

Planned contrasts revealed that BEST with sign significantly increased morphology percentage scores $t(1, 34) = 3.94$, $p = <0.001$. BEST without signdid non-significantly increase scores $t(1, 34) = 0.97$, $p = 0.34$.

ANCOVA Results for ELP Content and Morphology Outcomes (Beta, SE)

	Regression Models	
	ELP Content Model	ELP Morphology Model
	(1)	(2)
Intercept	22.681***	24.022***
	(6.013)	(7.732)
ELP Content Score	0.471***	
	(0.126)	
ELP Morphology Score		0.715***
		(0.168)
IDACI Score (centred)	-9.313	-21.391
	(11.107)	(19.831)
Dosage	-0.535	-1.064
	(0.880)	(1.488)
Age in Months (centred)	0.568**	0.965**
	(0.273)	(0.459)
BEST with Sign	13.843	25.242
	(12.752)	(21.444)
BEST without Sign	8.861	18.716
	(12.917)	(21.444)
Observations	46	46
R^2	0.429	0.468

Adjusted R^2	0.341	0.387
Residual Std. Error (df = 39)	11.053	18.914
F Statistic (df = 6; 39)	4.876*** (p = 0.001)	5.729*** (p = 0.0003)
Note:		*p<0.1; **p<0.05; ***p<0.01

Sentence Content: Generalised BEST Assessment

The covariate baseline score was significantly related to ELP content percentage score F(1, 34) = 16.97, p = <0.001. The covariate age in months at entry to the study approached significance F(1, 34) = 3.07, p = 0.09.

Planned contrasts revealed that BEST with sign improved ELP content scores compared to the reference group, which showed a trend towards significance t(1, 34) = 0.08.

The covariate baseline score was significantly related to the morphology percentage outcome score F(1, 34) = 10.91, p = <0.01. The covariate age was also significantly related to ELP morphology outcomes F(1, 34) = 4.93, p = 0.03.

Planned contrasts also revealed a trend towards significance for BEST with sign, which increased ELP morphology scores t(1, 34) = 1.20, p = 0.05.

Maintenance following Therapy

Maintenance following intervention was measured using ANCOVA. The analysis examined scores at follow-up, controlling for covariates including baseline score, age and language profile.

ANCOVA Results for Maintenance of Production, Comprehension, Content and Morphology Outcomes (Coefficient Estimates and Standard Errors)

	Regression Models			
	Production Model	Comprehension Model	Content Model	Morphology Model
	(1)	(2)	(3)	(4)
Intercept	-3.673	-1.213	51.744***	35.764***
	(2.703)	(3.110)	(10.437)	(8.791)
Production Score	0.824***			

	(0.199)			
Comprehension Score		0.454*		
		(0.245)		
Content Score			0.153	
			(0.194)	
Morphology Score				0.545***
				(0.182)
IDACI Score (centred)	-3.776	-14.789	-4.472	11.191
	(9.747)	(11.118)	(15.545)	(21.576)
Dosage	-1.945**	-0.281	-0.487	-0.771
	(0.883)	(0.950)	(1.248)	(1.798)
Age in Months (centred)	-0.163	-0.202	0.890**	0.303
	(0.244)	(0.279)	(0.381)	(0.547)
BEST with Sign	36.369***	9.297	25.609	28.275
	(12.773)	(13.863)	(18.342)	(26.464)
BEST without Sign	27.093**	1.846	20.282	27.283
	(12.228)	(13.438)	(17.624)	(25.203)
Observations	47	47	42	42
R^2	0.423	0.215	0.388	0.339
Adjusted R^2	0.337	0.097	0.283	0.226
Residual Std. Error	10.037 (df = 40)	11.575 (df = 40)	14.258 (df = 35)	20.534 (df = 35)
F Statistic	4.889*** (df = 6; 40) (p = 0.001)	1.821 (df = 6; 40) (p = 0.120)	3.700*** (df = 6; 35) (p = 0.006)	2.991** (df = 6; 35) (p = 0.019)
Note:				*p<0.1; **p<0.05; ***p<0.01

Production

	TAU	BEST without sign	BEST with sign
Mixed Profile	-13.66*	-7.38	0.20**
Expressive Profile	0.04*	-8.10	-4.07*
Receptive Profile	5.74	-2.21	26.71

None of the covariates were significantly related to production score at follow-up.

Intervention arm was significantly related to production score F(2, 35) = 4.78, p = 0.01.

Planned contrasts revealed a significant negative effect of the interaction between BEST with sign and an expressive language profile t(1, 35) = -2.15, p = 0.04. BEST with sign significantly improved production score at follow-up t(1, 35) = 3.09, p = <0.01. An expressive language profile significantly improved scores compared to the reference group t(1, 35) = 2.11, p = 0.04.

Comprehension

	TAU	BEST without sign	BEST with sign
Mixed Profile	-12.45	-11.13	-8.30
Expressive Profile	-13.84	-19.38	-10.18
Receptive Profile	-13.1	-11.52	13.84

Intervention arm was non-significantly related to comprehension outcome scores F(2, 36) = 0.31, p = 0.74. There were no other significant predictors of comprehension scores.

Content

	TAU	BEST without sign	BEST with sign
Mixed Profile	50.91***	56.89	67.29*
Expressive Profile	50.87	69.14	70.63
Receptive Profile	35.12	70.89	70.84

The covariate age in months was significantly associated with content percentage scores at follow-up F(1,30) = 4.42, p = 0.04. Intervention arm was non-significantly related to content scores F(2, 30) = 2.81, p = 0.08.

Planned contrasts revealed that BEST with sign significantly increased content scores compared to the reference category t(1, 30) = 2.37, p = 0.02. Age significantly increased content score t(1, 30) = 2.10, p = 0.04.

Morphology

	TAU	BEST without sign	BEST with sign
Mixed Profile	52.96***	58.95	64.58

Expressive Profile	31.87	60.8	65.67
Receptive Profile	35.01	81.94	72.91

The covariate baseline score was significantly related to morphology score at follow-up F(1, 5.05, p = 0.03.

Planned contrasts revealed that baseline score significantly increased the follow-up score t(1, 30) = 2.25, p = 0.03.

Maintenance of Generalisation

ANCOVA Results for Maintenance of Generalisation to Generalised BEST Assessment for Sentence Content and Morphology (Coefficient Estimates and Standard Errors)

	Regression models	
	ELP Content Model (Maintenance)	ELP Morphology Model (Maintenance)
	(1)	(2)
Intercept	18.071*	31.564***
	(8.902)	(8.822)
ELP Content Score	0.545***	
	(0.189)	
ELP Morphology Score		0.586***
		(0.200)
IDACI Score (centred)	13.557	9.601
	(16.179)	(22.249)
Dosage	-1.285	-1.994
	(1.438)	(1.894)
Age in Months (centred)	0.324	0.541
	(0.418)	(0.561)
BEST with Sign	32.511	40.968
	(21.073)	(27.728)
BEST without Sign	31.320	38.892

	(20.609)	(26.806)
Observations	41	41
R^2	0.315	0.313
Adjusted R^2	0.194	0.192
Residual Std. Error (df = 34)	15.658	20.702
F Statistic (df = 6; 34)	2.602** (p = 0.035)	2.579** (p = 0.037)
Note:		*p<0.1; **p<0.05; ***p<0.01

Maintenance of Generalisation Analysis: ANCOVA Results

ELP Sentence Content

	TAU	BEST without sign	BEST with sign
Mixed Profile	35.60*	41.76	41.55
Expressive Profile	25.67	47.09	52.54
Receptive Profile	-2.72*	55.54*	53.88*

The covariate baseline content score was significantly related to sentence content scores at follow-up F(1, 29) = 10.58, p = <0.01. Intervention arm was non-significantly related to follow-up content scores F(29) = 0.48, p = 0.62.

Planned contrasts revealed that baseline score significantly increased follow-up scores t(11, 29) = 3.25, p = <0.01. There interaction between BEST with sign and a receptive language profile significantly increased content scores t(29) = 0.04. The interaction between BEST without sign and receptive language profile also significantly increased scores at follow-up t(29) = 2.65, p = 0.01.

Receptive language profile significantly decreased scores at follow-up compared to the reference category t(29) = -2.26, p 0.03.

ELP Sentence Morphology

	TAU	BEST without sign	BEST with sign
Mixed Profile	50.73***	49.81	49.58
Expressive Profile	24.61*	52.72	62.47*
Receptive Profile	-2.88*	66.01**	72.85*

The covariate baseline score was significantly related to the follow-up score F(1, 29) = 7.67, p = <0.01. The interaction between intervention arm and language profile was significantly related to follow-up scores F(4, 29) = 3.19, p = 0.03.

Planned contrasts revealed that age significantly improved scores t(29) = 2.77, p = <0.01. The interaction between BEST with sign and receptive language profile significantly increased scores compared to the reference group t(29) = 2.60, p = 0.01. The interaction between BEST without sign and receptive language profile significantly increased scores t(29) = 2.84, p= <0.01. The interaction between BEST with sign and expressive language profile also significantly increased scores at follow-up t(29) = 2.24, p = 0.03.

There was a negative significant effect of receptive language profile on follow-up scores t(29) = -2.53, p = 0.02. There was also a negative significant effect of expressive language profile t(26) = -2.06, p = <0.05.

ANCOVA Analysis Conclusion

Several key findings emerged from the ANCOVA analysis.

There was evidence for the efficacy of BEST on the main outcome variables. For production scores from baseline to outcome there was a positive effect of the interaction between BEST with sign receptive language profile on outcome scores. No significant relationships explained the comprehension scores. For BEST Assessment content scores there were significant main effects of both BEST without sign and BEST with sign. For the BEST Assessment morphology scores there was a significant main effect of BEST with sign.

There was some weak evidence of generalisation to ELP scores; for both content and morphology outcomes there were non-significant positive trends for children in the BEST with sign arm, however no significant relationships were identified.

There was some mixed evidence of maintenance following the intervention. For production there was a significant negative effect of the interaction between BEST with sign and expressive language. No evidence was found for comprehension scores. For BEST Assessment content scores there was a significant positive effect of BEST with sign, and also of age on follow-up scores. There was a significant effect of baseline score only on follow-up score.

There was positive evidence for maintenance of the generalised ELP language scores. For the ELP content scores the interaction between BEST with sign and receptive language, and also between BEST without sign and receptive language profile significantly improved scores. Finally for the ELP morphology score the interactions between BEST with sign and receptive language, BEST with sign and expressive language, and BEST without sign and receptive language all significantly increased scores.

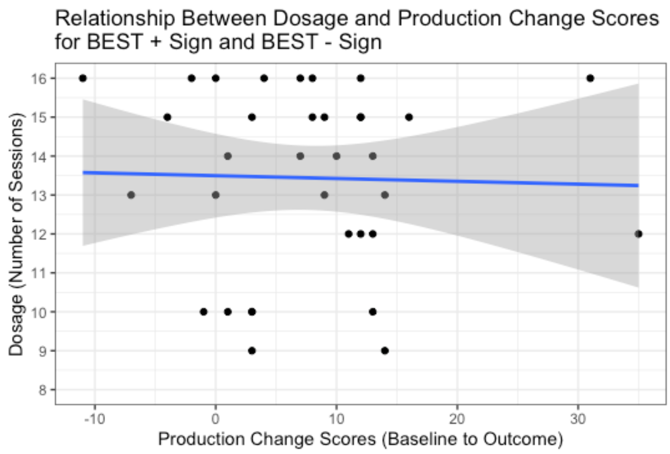
Multilevel Results for Sensitivity Analysis Outcomes Beta(SE)

	Multilevel Model					
	Production	Comprehension	Targeted Content	Targeted Morphology	Generalised Content	Generalised Morphology
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	74.548***	76.616***	-12.682***	-17.466***	-2.938	-11.825***
	(1.521)	(2.398)	(3.105)	(4.219)	(3.872)	(4.384)
Time 1	0.991					
	(2.201)					
Time 2	3.835					
	(3.087)					
BEST without Sign		5.466				
		(3.989)				
BEST with Sign		8.465**				
		(3.818)				
Age (centred)			-0.256		1.081	
			(4.636)		(3.381)	
Dosage			11.831**		2.513	
			(4.649)		(4.045)	
IDACI (centred)				11.620**		11.834**
				(5.589)		(4.853)
Expressive Profile				23.043***		21.694***
				(6.028)		(5.063)
Receptive Profile	-3.249	-0.863	-5.328	4.908	-16.290	-7.978
	(5.152)	(6.046)	(11.298)	(16.296)	(12.778)	(16.213)
BEST without Sign Time 1	-1.685	-1.187	-2.963	15.011	-10.504	2.834
	(5.199)	(5.865)	(11.512)	(16.505)	(12.745)	(16.239)
BEST without Sign Time 2	-0.309**	0.132	0.850**	0.316	0.771	0.425
	(0.127)	(0.178)	(0.357)	(0.728)	(0.455)	(0.727)
BEST with Sign Time 1	0.267	0.166	-0.243	-0.960	0.658	-0.060

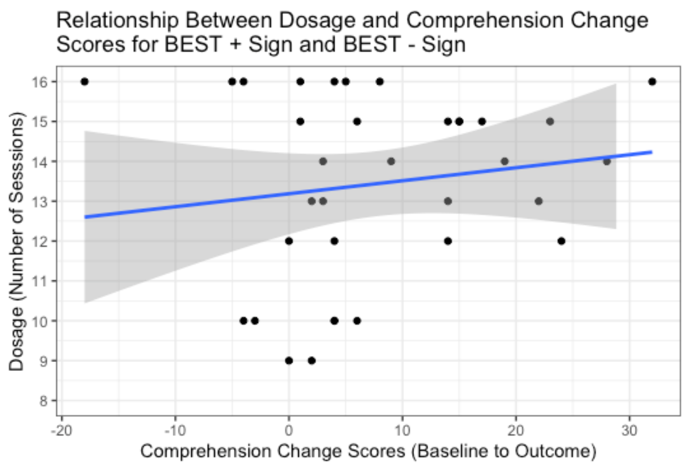
	(0.373)	(0.391)	(0.812)	(1.183)	(0.891)	(1.164)
BEST with Sign Time 2	-6.682	-11.129	-27.096**	-45.388**	-23.537*	-42.837**
	(4.665)	(7.329)	(9.849)	(14.416)	(10.928)	(14.087)
LanguageProfileExpressive	0.263	5.461	2.073	-10.668**	-1.909	-10.953**
	(1.537)	(4.415)	(3.236)	(5.011)	(3.775)	(4.967)
ProdTimeprod1:InterventionArm1	6.731**					
	(3.072)					
ProdTimeprod2:InterventionArm1	2.642					
	(4.312)					
ProdTimeprod1:InterventionArm2	8.469***					
	(3.019)					
ProdTimeprod2:InterventionArm2	8.548**					
	(4.279)					
LanguageProfileReceptive		0.372				
		(2.356)				
CompTimecomp1:InterventionArm1		3.850				
		(5.592)				
CompTimecomp2:InterventionArm1		1.404				
		(5.320)				
CompTimecomp1:InterventionArm2		6.421				
		(5.569)				
CompTimecomp2:InterventionArm2		5.781				
		(5.442)				
ContentTimeContent1:InterventionArm1			33.436***		2.844	
			(6.470)		(4.808)	
ContentTimeContent2:InterventionArm1			17.522***		12.573**	
			(6.550)		(5.794)	

ContentTimeContent1:InterventionArm2			33.283***		7.433	
			(6.318)		(4.519)	
ContentTimeContent2:InterventionArm2			19.873***		13.678**	
			(6.480)		(5.698)	
MorphologyTimeMorph1:InterventionArm1				21.387***		7.967
				(7.817)		(6.847)
MorphologyTimeMorph2:InterventionArm1				12.328		8.452
				(8.490)		(7.252)
MorphologyTimeMorph1:InterventionArm2				24.474***		11.299*
				(7.623)		(6.529)
MorphologyTimeMorph2:InterventionArm2				13.456		6.816
				(8.426)		(7.242)
Observations	128	101	122	122	122	122
Log Likelihood	-415.374	-335.125	-472.766	-515.958	-467.608	-508.659
Akaike Inf. Crit.	900.748	742.251	1,015.532	1,101.916	1,005.215	1,087.317
Bayesian Inf. Crit.	1,000.569	836.395	1,113.673	1,200.057	1,103.356	1,185.458
Note:				*p<0.1; **p<0.05; ***p<0.01		

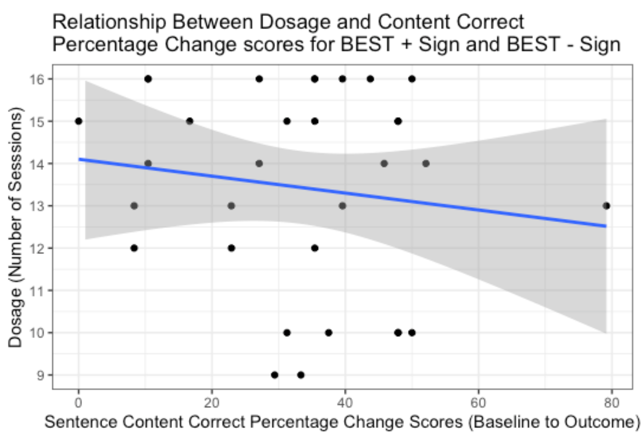
Appendix 23: Correlation Between Dosage and Growth Additional Information



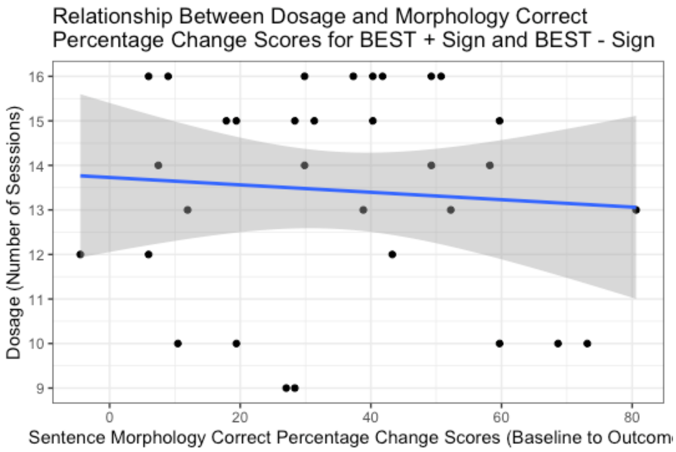
[r (32) = -0.03, p = 0.87]



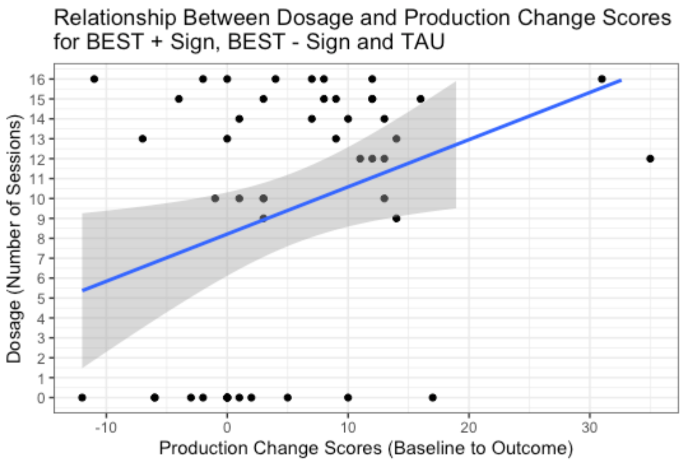
[r (32) = 0.15, P = .40]



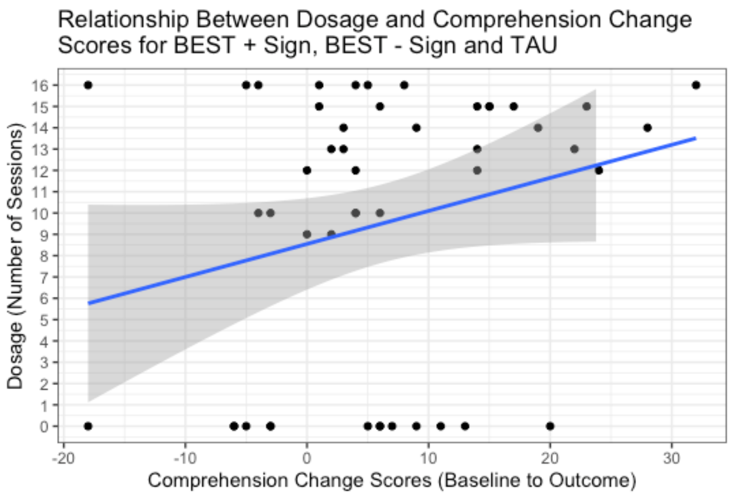
[r (30) = -0.14, p = .44]



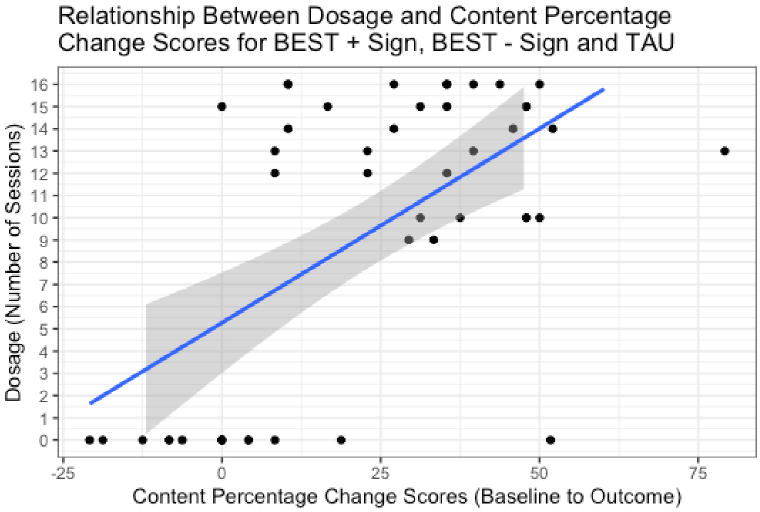
[r (30) = -0.08, p = 0.68].



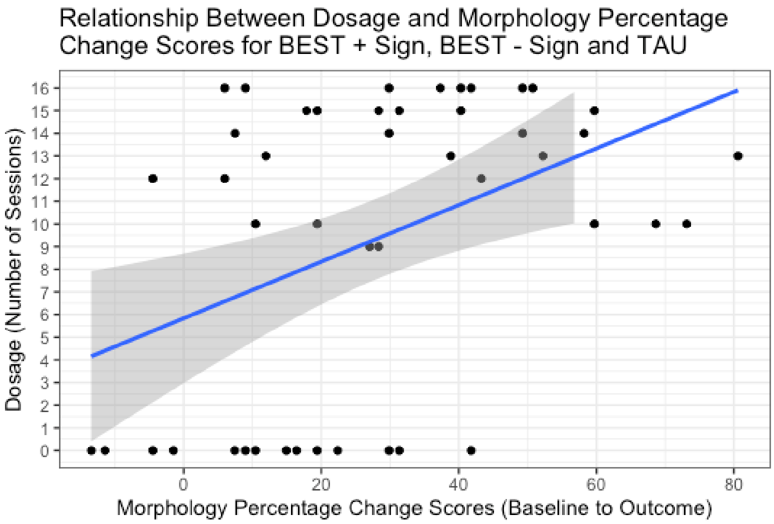
[r (46) = 0.34, p = .02]



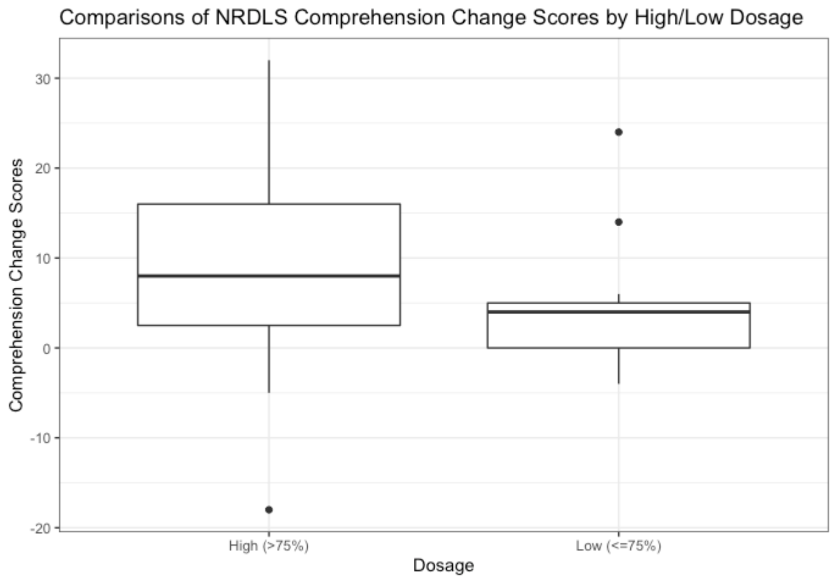
[r (46) = 0.25, p = 0.08]



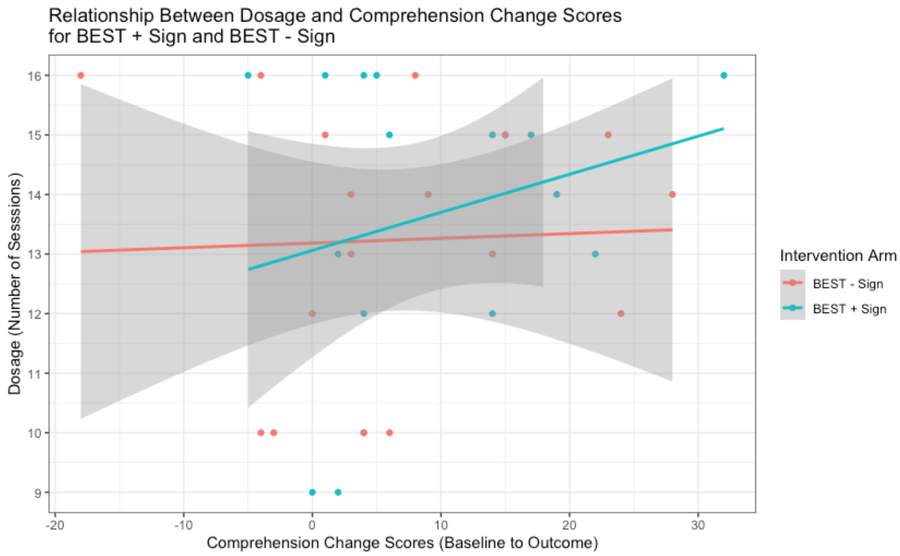
[r (44) = 0.61, p = <.001]



[r (44) = 0.43, p = <.01]



Comprehension scores by high and low dosage



showing the relationship between comprehension change scores and dosage

Correlation

The relationship between dosage (<=75% />75%) and comprehension change scores from baseline to outcome was non-significantly correlated. Pearson's correlation coefficient was -0.15 (p = 0.40).

For high dosage (>75%) there was not a significant correlation between dosage and comprehension change score from baseline to outcome. Pearson's correlation coefficient -0.29 (p= 0.19)

For low dosage (<=75%) there was not a significant correlation between dosage and comprehension change score from baseline to outcome. Pearson's correlation coefficient 0.55, (p= 0.08)

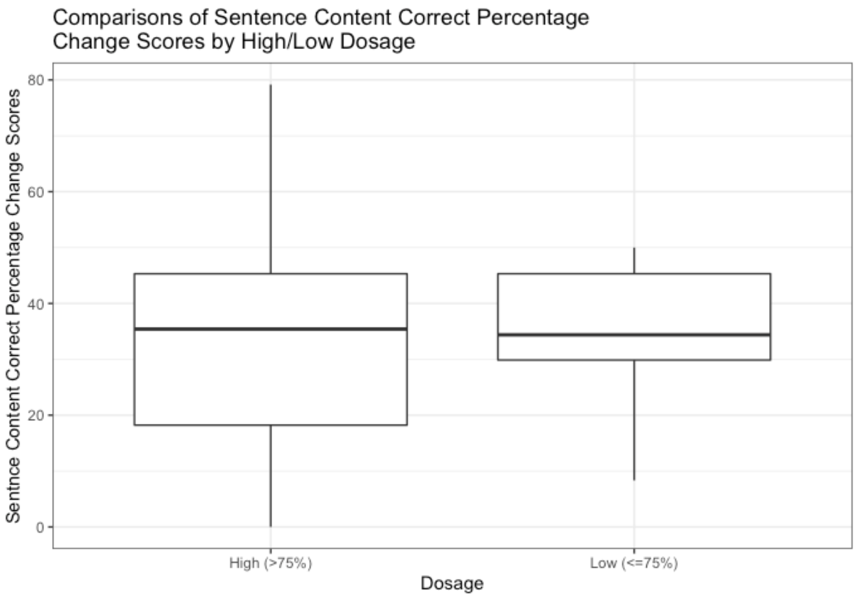
By Dosage and Intervention Arm

For high dosage (>75%) BEST with sign arm there was not a significant correlation between dosage and comprehension change score from baseline to outcome. Pearson's correlation coefficient -0.23, (p= 0.48)

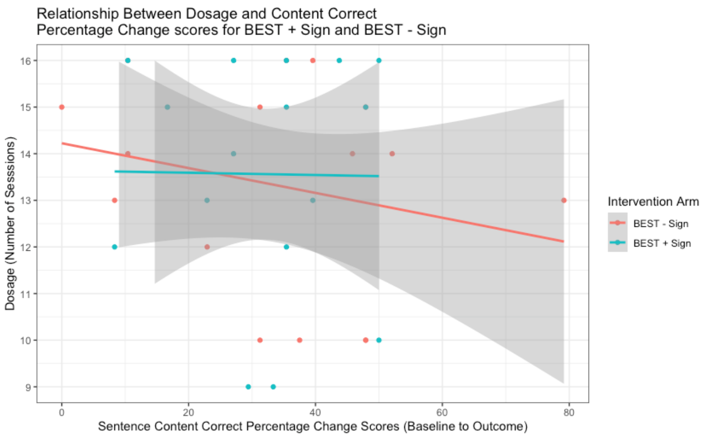
For high dosage (>75%) BEST without sign arm there was not a significant correlation between dosage and comprehension change score from baseline to outcome. Pearson's correlation coefficient -0.41, (p= 0.21)

For low dosage (<=75%) BEST with sign arm there was not a significant correlation between dosage and comprehension change score from baseline to outcome. Pearson's correlation coefficient 0.74, (p= 0.15)

For low dosage (<=75%) BEST without sign arm there was not a significant correlation between dosage and comprehension change score from baseline to outcome. Pearson's correlation coefficient 0.56, (p= 0.24)



Content correct percentage change scores by high and low dosage



The relationship between dosage (<=75% />75%) and sentence content correct percentage change scores from baseline to outcome was non-significantly correlated. Pearson's correlation coefficient was -0.14 (p = 0.45).

By Dosage

For high dosage (>75%) there was not a significant correlation between dosage and content correct percentage change scores from baseline to outcome. Pearson's correlation coefficient -0.12, (p= 0.60)

For low dosage (<=75%) there was not a significant correlation between dosage and content correct percentage change scores from baseline to outcome. Pearson's correlation coefficient -0.50, (p= 0.14)

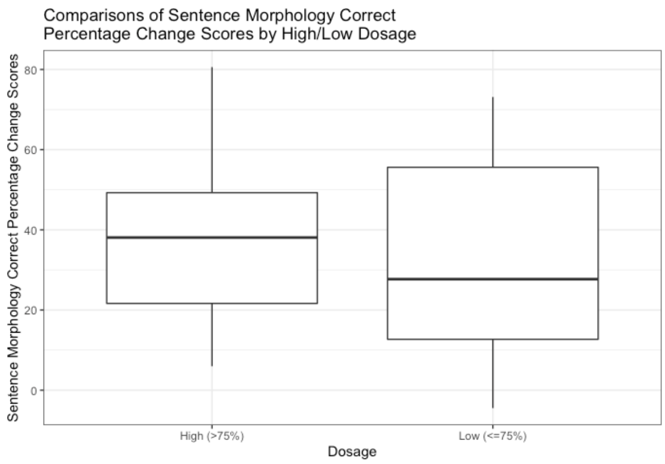
By Dosage and Intervention Arm

For high dosage (>75%) BEST with sign arm there was not a significant correlation between dosage and content correct percentage change score from baseline to outcome. Pearson's correlation coefficient -0.23, (p= 0.48)

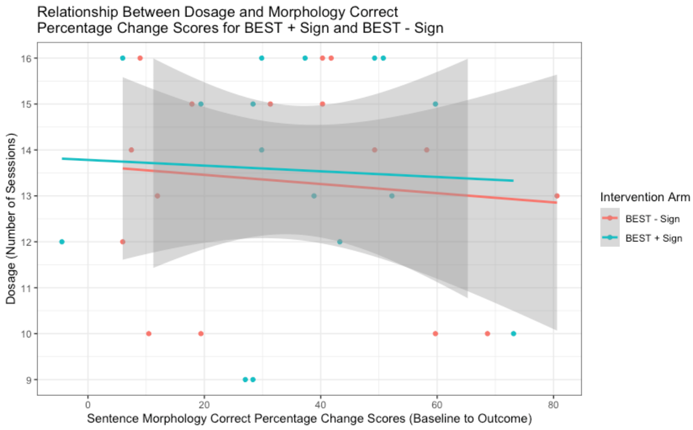
For high dosage (>75%) BEST without sign arm there was not a significant correlation between dosage and content correct percentage change score from baseline to outcome. Pearson's correlation coefficient -0.41, (p= 0.21)

For low dosage (<=75%) BEST with sign arm there was not a significant correlation between dosage and content correct percentage change score from baseline to outcome. Pearson's correlation coefficient 0.74, (p= 0.15)

For low dosage (<=75%) BEST without sign arm there was not a significant correlation between dosage and content correct percentage change score from baseline to outcome. Pearson's correlation coefficient 0.56, (p= 0.24)



Morphology correct percentage scores by high and low dosage



The relationship between dosage (<=75% />75%) and sentence morphology correct percentage change scores from baseline to outcome was non-significantly correlated. Pearson's correlation coefficient was -0.08 (p = 0.68).

By Dosage

For high dosage (>75%) there was not a significant correlation between dosage and morphology correct percentage change scores from baseline to outcome. Pearson's correlation coefficient -0.23, (p= 0.31)

For low dosage (<=75%) there was not a significant correlation between dosage and morphology correct percentage change scores from baseline to outcome. Pearson's correlation coefficient -0.34, (p= 0.33)

By Dosage and Intervention Arm

For high dosage (>75%) BEST with sign arm there was not a significant correlation between dosage and morphology correct percentage change score from baseline to outcome. Pearson's correlation coefficient -0.21, (p= 0.54)

For high dosage (>75%) BEST without sign arm there was not a significant correlation between dosage and morphology correct percentage change score from baseline to outcome. Pearson's correlation coefficient -0.27, (p= 0.43)

For low dosage (<=75%) BEST with sign arm there was not a significant correlation between dosage and morphology correct percentage change score from baseline to outcome. Pearson's correlation coefficient -0.26, (p= 0.67)

For low dosage (<=75%) BEST without sign arm there was not a significant correlation between dosage and morphology correct percentage change score from baseline to outcome. Pearson's correlation coefficient -0.51, (p= 0.38).

Appendix 24 Thematic analysis of Stakeholder Consultation carried out with school staff

Practicalities: TA as a resource	Practical impact	Key messages: level of need	key messages: desired CPD	Interactions: rapport and social capital	Interactions: Communication with parents	Interactions: Parental engagement	Perceptions: positive	Perceptio negative
staff to child ratio is compromised	The sessions (assessments) didn't impact on us greatly	provide what children don't get at home	Would like feedback on the individual children's assessments to know what progress children have made	accommodating, professional	They preferred to speak to the teachers	useful if the researcher could have met parents	the complexity built up quickly	it was ver repetitive
TAs were underused, just moving objects. Not a good use of adult's time	lots of form filling for staff	key messages and how do you get them across to parents?	Would like to be able to observe sessions again before delivering them	friendly	parents knew children had been working on the project and could be proud	parents didn't give a high level of support	Liked the homework booklets, they were handy and provided a focus	Repetitive
project drained resources	questionnaire was very lengthy with too many options	EAL is increasing	Would like to know how to deliver the intervention,	fantastic.	The form had phone numbers	fun, interactive app would be good to attract interest and attention	Mixed benefits, beneficial to some children	The proje took som children backward
	time that the researcher was available was tricky	language barrier	resources and practical ideas	always available and happy to be contacted	phoned parents	helped when parents already had a child in the school, and trusted the teacher	Some children were going forwards, it was at the right level for children	repetitive
		[don't know if] issues are 'genuine' or due to bilingualism.	Provide training to anyone who would be interested, unlikely to leave own	great	It's a lot of information for parents, we need to give them the information but also need to consider the	Parents know that children are all at different levels and were keen for their	Therapy delivery was great	No benefi to staff, difficult t see the benefits when in t control group

		Local Authority		types of children	child to receive input		
	Speech and language problems generally are getting worse	Already have a lot of training	took time to talk to staff and communication was easy	Some parents had difficulty reading the information.	researcher came to talk to parents but didn't have much response		FOCUS seemed more appropria to ask abo adults
	skills have declined, could technology be a possible reason	Networking would be just another thing to do	took their time explaining	would have been better emailing parents directly for forms	Children had already been identified with SLCN so there were no problems with parents	Some children were not always paying attention but everything was repeated so it didn't matter	they alrea have year of experienc using visu and other materials
	Lots of children with SLCN	understand where the intervention comes from	clear	make the information sheets more engaging	The majority of parents were fine. Parents weren't worried about why their child was chosen and just wanted access to anything that would help them		Would it work bett 1-1?
	accessing Local Authority intensive provision	like to know how to do more in school.	spoke to teachers	be aware of parents own level of interest and understanding	meeting may have been useful		
	a high level of need	get on board with the programme and to help parents do the activities	school information sheets were fine	Homework was not explained to parents	Parents were amenable, not concerned that their children had been selected		

	chance to work alongside in a way that is not detrimental to the setting	We didn’t know about the classroom observation	a video to show parents what to do at home	parents saw it as an opportunity to access support
	like someone coming in to school to monitor progress		Some parents gave positive feedback	Many parents did not return the forms
	activities that would work for all children			lot of chasing parents to try and get forms returned,
	OFSTED have a new focus on vocabulary and language			It (homework) seemed boring and not sure how it could be made more interesting
	Input regarding the strategies			
	TAs would like to know more			