Improving Working Memory in Children with Language Difficulties

Emma Christopher

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Division of Language and Communication Sciences
School of Health Sciences
City, University of London.

October 2020

Declaration

Declaration I grant powers of discretion to City University's Librarian to allow this thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

Children with language impairments show considerable difficulties with working memory, the underpinning abilities that allow us to carry out every day thinking and reasoning tasks. Most working memory interventions, usually carried out through computer-based approaches, have provided limited support for benefits to other areas of academic ability and cognitive skill. However, previous research that has adopted a non-computerised approach has found effects on working memory and other cognitive skills such as reading comprehension in typically developing children (Henry, Messer, & Nash, 2014). The current research aimed to assess whether the same intervention would be effective for children with language impairments.

Forty-seven children with language difficulties as a primary need, aged 6-11 years, were randomly allocated to either a working memory intervention group or an active control group. Both groups of children were visited by a researcher three times a week for six weeks, totalling 18 sessions, for approximately 10 minutes at a time. Children in the working memory intervention group were administered 11 trials of an Odd One Out span task (Henry, 2001) and 11 trials of a Listening Recall task (Gathercole & Pickering, 2001). Both tasks required executive level working memory skills. All trials in both tasks were adaptively titrated to the child's ability levels and adjusted continuously to ensure that the tasks remained appropriately challenging in order to facilitate improvement. Children in the active control group were administered the same number of trials but with the working memory requirement removed. All participants were given a battery of assessments before the intervention, immediately after the intervention was completed, and also at a 9 month

follow-up, to see if there were any near and far transfer effects from the intervention to other working memory and language skills, and if so, how long these were maintained.

Regression analyses controlling for pre-intervention scores (and age) showed that group membership (trained vs active control) was a significant predictor of performance on both of the directly trained tasks, Listening Recall and Odd One Out, immediately post-intervention and at a 9-month follow-up. In similar regression analyses, group was also a significant predictor of performance in all six near transfer measures (working memory tasks that had not been trained), digit recall, word list recall, block recall, counting recall, backward digit recall and pattern span, at both post-intervention and at a 9 month follow-up. In further regression analyses, group was also a significant predictor of performance in one far transfer effect (a cognitive skill not directly related to working memory), sentence comprehension, both immediately post-intervention and at a 9 month follow-up. This suggests that improving the ability to divide attention between processing and storage may have resulted in benefit to sentence comprehension ability. However, no group differences following intervention were found for reading accuracy, reading comprehension or receptive grammar.

The current study's findings reveal that following a short, adaptive, face-to-face working memory intervention, group membership (trained vs active control) is a significant predictor of immediate and longer-term performance in directly trained, near and far transfer tasks. Further research is needed to better understand why many computer-based working memory interventions have been unsuccessful in obtaining far transfer effects, when the current intervention produced some significant effects.

Acknowledgements

I would firstly like to express my sincere gratitude to my PhD supervisors, Professor Lucy Henry, Professor Shula Chiat and Professor David Messer. Your guidance, support, and many words of wisdom throughout this PhD have been invaluable to me. I have thoroughly enjoyed working with you all, you have truly been an inspiration and I thank you for everything you have taught me at every stage of this research project.

I would also like to offer my special thanks to all the primary schools, teachers and especially the children who participated in this study. Thank you for being so keen and enthusiastic and for always bringing a smile to our training sessions!

I would also like to personally thank my family and my husband Michael, who have been a constant source of support and encouragement. I dedicate this work to my two beautiful children, Grace and Charlie, who have always been my driving force to succeed in this journey.

List of Contents

List of Tables

List of Figures

Chapter One

Language Disorder, Heterogeneity and the Working Memory Model

This chapter will begin by giving an overview of language disorder, discussing how discrepancies in diagnostic criteria have been argued to be underlying factors in the heterogeneity of this population. This chapter will then move on to discuss each component of 'The Multi Component Model of Working Memory' (Baddeley & Hitch, 1974; Baddeley, 2000) as the central executive component of the working memory system is central to this research. This chapter will conclude with an exploration of the relationship between working memory and language.

1.1 Language Disorder and Heterogeneity

This section will begin by discussing the variation of severity and linguistic patterns in the language disorder population, and how there is often little agreement concerning the criteria to identify and classify these problems. Critique regarding a 'test score' only approach to diagnosis will be discussed alongside the need to include functional impact of language difficulties as part of the diagnosis process. Discrepancies of standardised language test score 'cut off' points and non-verbal intelligence tests will be addressed before moving on to discuss appropriate terminology for those with language difficulties, following the removal of the term Specific Language Impairment (SLI) from the latest publication of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013). The impact of a lack of consensus in diagnostic criteria, and appropriate and consistent terminology to clinical practice and research will be discussed, as will the need for these matters to be resolved in order for a clearer understanding of language difficulties to develop, prevalence determined and effective treatments and preventions to be adopted. This section will conclude with an overview of the participant inclusion criteria used for the current study, which has been based on the literature discussed throughout this section. In terms of the current thesis, having an awareness of the heterogeneity of the language disorder population is vital when comparing previous research undertaken, and will ensure that assumptions are not made that clinical groups are always compared on a like for like basis.

Language is an integral part of daily life, in terms of learning, communication, social interaction and expression of thoughts and feelings. Language development allows children to become better at engaging in socialization and learning from the environment around them. The terms language, speech and communication are often used interchangeably, however they have very distinct meanings (Bishop et al., 2016). Language is a form of communication which relies on meaning being conveyed by its structure. It can be used in different modalities such as speech, written text or sign and involves recognition and comprehension of words and sentences as a foundation for expression of ideas and information (Bishop et al., 2016). Speech however, is the ability to produce vocal sounds and is a process that incorporates both motor and linguistic skills; an individual can, therefore, possess an impaired speech function but an intact language ability (Bishop et al., 2016).

Most children are able to develop the fundamental skills of language and communication relatively effortlessly (Roulstone et al., 2011). However, for approximately 7% of the population, a significant deficit is shown in language ability that cannot be attributed to a hearing loss, low nonverbal ability or neurological damage (Leonard, 2014). As diagnostic criteria for language disorders are largely exclusionary, there is considerable heterogeneity among individuals who have been classified as affected (Fisher et al., 2003). Individuals within a language disorder population show considerable variation in both severity and linguistic patterns, with language deficits commonly diagnosed in syntax, morphology, phonology, the lexicon and pragmatics, and in receptive and expressive language (Mason et al., 2010).

Although unexplained language impairment is common in children, there is often little agreement concerning the criteria to identify and classify these problems. This has been suggested to be an underlying factor for the heterogeneity found among individuals within this population (Leonard, 2014). When a diagnosis of a language disorder is made, the first step is to establish whether an individual's limitations in language ability constitute an impairment (Leonard, 2014). Low scores on standardised language tests have primarily played a role in deciphering whether a child exhibits a language disorder. Diagnosis is often seen as justifiable when a child's language score falls at the very low end of the normal distribution on a standardised test (Leonard, 2014). The 'test score only' approach has come under scrutiny, as although a child may present with a relatively weak language ability, this may not on its own be sufficient reason to conclude that a language impairment exists (Leonard, 2014). Reilly et al. (2014b) suggest that there is an urgent need to develop a reliable and valid measure of functional impacts of language problems for an individual that can then be used in clinical decision making rather than a sole focus on language test scores.

The heterogeneity found among those with language impairments has also been suggested to be due to a discrepancy in the 'cut off' points used for standardised language test scores. Prevalence of language disorders is 2.3% if a cut off score of two standard deviations below the mean is selected, whereas a prevalence of language disorders is at a much higher level of 15.9% if a cut off score of one standard deviation below the mean is used (Leonard, 2014). Language is a crucial aspect of development, and longitudinal studies of children with language impairment have found a detrimental effect on later outcomes in terms of social, employment and educational achievements (Reilly et al., 2014a). These risk factors have been graded by the degree of language impairment experienced in early life, and a score of one standard deviation below the mean for language ability measures has been associated with longer term risk (Reilly et al., 2014a). It has been suggested, therefore, that those children within this mild/below average category of language impairment are monitored and supported (Reilly et al., 2014a).

A further key measure used in language disorder diagnosis is nonverbal intelligence testing, with the intent to show that an individual's language impairment cannot be due to a deficit in general intelligence (Leonard, 2014). As with standardised language test scores, there are also discrepancies over the appropriate 'cut off points' to use in nonverbal intelligence tests. Although most use a score of 85 (one standard deviation below the mean) as a lower boundary for language disorders, the fact that the cut off point for an intellectual disability is a score of 70 leads many researchers such as Spaulding et al. (2008) to employ a lower cut off point of 75. The difference in the nonverbal intelligence test scores used to determine a diagnosis of language disorders could therefore be argued as another underlying factor in the heterogeneity among this population. Children with a nonverbal intelligence score greater than 85 have been shown to not significantly differ in their language skills or response to intervention compared to children who score in the borderline range of 70-84 in nonverbal IQ (Friel-Patti, 1999). This indicates that it could be beneficial to include children with a nonverbal intelligence score of 70 or above in intervention studies.

‘Specific language impairment’ (SLI) has been used to describe children with language impairments which are greater than would be expected from non-verbal abilities in most clinical practice and research over the past 30 years (Leonard, 2014; Reilly et al., 2014b). Following the removal of the term SLI from the most recent publication of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013) discussions regarding the most appropriate terminology to employ for language difficulties have been at the forefront of an international debate (Ebbels, 2014). Following the recommendation from The American Speech and Hearing Association (ASHA), the DSM-5 now uses the term ‘Language Disorder’ within its manual (Ebbels, 2014).

Debate has surrounded the use of the term ‘specific’ within SLI, as it has been argued that this could imply that there are no other areas of clinical concern (Haines, 2015). In clinical practice however, it is argued that many children with language difficulties present with other co-existing problems ranging from learning difficulties, attention deficit hyperactivity disorder, autism spectrum disorder, speech-sound disorders or developmental coordination disorders, therefore making the term ‘specific’ inaccurate (Bellair et al., 2014; Bishop, 2014). This also relates to the issues already discussed concerning heterogeneity within this population, portraying the differences among those with a language disorder. The term ‘language impairment’ has often been suggested by a number of authors as a more appropriate term to use in order to avoid the controversy surrounding the use of the term ‘specific’. However, as previously discussed, problems have arisen due to a difficulty in categorising children into clinical groups on the basis of language impairment versus non language impairment due to a discrepancy of agreed standardised test scores (Haines, 2015). The term ‘primary language impairment’ has been debated in a similar manner to the term ‘specific’, as for many children, language need is not always the ‘primary’ difficulty, especially when children have co-existing conditions (Haines, 2015). The use of the word 'primary' has also been suggested to imply a specific age range and, therefore, led to the assumption that unexplained language difficulties could only occur in primary aged children, taking the focus away from identification and intervention procedures that are crucial in the early pre-school years (Bishop, 2014).

In this debate to choose the most appropriate terminology, even 'language disorder', which is now used in the DSM-5, has come under scrutiny by Bishop (2014) as the acronym LD can be easily confused with learning difficulty. 'Developmental language disorder' (DLD) however, does seems to have attracted more support with the view that this may also align language disorders more closely to other medical developmental disorders such as ASD, which in turn may offer it a higher status and attract a greater level of support (Haines, 2015).

A lack of consensus and use of appropriate and consistent terminology of language disorders has, unfortunately, resulted in miscommunication across different professionals both in clinical practice and research (Reilly et al., 2014a). Although other developmental conditions such as autism spectrum disorder, attention deficit hyperactivity disorder or developmental dyslexia have similar challenges in terms of their diagnostic criteria, the terminology used for these conditions does seem to remain fairly stable (Reilly et al., 2014a). Following the international debate around appropriate terminology, it has been questioned whether a single term can in fact be adopted for language disorders due to the diverse nature of language, language difficulties, and needs of those with language difficulties. The diverse range of needs surrounding individuals within this population have been argued not to be 'specific', and this may explain why it has been so difficult for those within this field to reach an agreement in terms of both diagnostic criteria and terminology (Reilly et al., 2014a).

The need for consistency, however, in both criteria and terminology, is essential in order for effective communication to exist between families, those with language disorders, professionals across a wide range of health and educational sectors, medical and policy makers, service providers and researchers (Reilly et al., 2014a). It is vital that those working with individuals with language disorders understand the cause of the problem, determine prevalence and adopt a consistent approach to delivering effective treatments and preventions. Language difficulties have no clear-cut boundaries, and a focus on the functional impact of language difficulties over eligibility criteria (with clear moves to develop a reliable and valid measurement of functional impact which can be used within clinical decision making) could be seen a positive step forward (Bishop, 2014; Reilly et al., 2014a).

An awareness of the heterogeneity of the language disorder population is required to be at the forefront of consideration in terms of the current thesis, and when comparing previous research undertaken. It could be assumed that clinical groups have been compared on a like for like basis, whereas this has been shown not always to be the case. There is clear discrepancy between diagnostic criteria in terms of language ability and nonverbal intelligence levels for those with language difficulties, and co-existing developmental disorders commonly exhibited in those with language difficulties which also contribute to variation in the profiles of this group (Reilly et al.,2014b; Tomblin et al.,1996; Wetherell et al., 2007; Williams & Lind, 2013). For the purpose of the current research, the term 'language disorder' will be used in line with recommendations made by The American

Speech and Hearing Association and the publication of this terminology in the latest Diagnostic and Statistical Manual of Mental Disorders (DSM-5).

The current thesis also selected participants based on the following criteria:

i) Participants were required to score one standard deviation below the mean in two core areas of language, assessed via the Clinical Evaluation of Language Fundamentals 4 (CELF 4) (Semel et al., 2006).

ii) Participants required a score of 70 or above in non-verbal intelligence testing.

iii) Participants with co-occurring conditions (with the exception of autism spectrum disorder) were included in the study, provided language difficulty was identified as a primary need.

Risk factors for detrimental effects, in terms of social, employment and educational achievements, have been shown to be higher for those with language ability just one standard deviation below the mean (Reilly et al., 2014a). Therefore, the current thesis included participants who had a score of one standard deviation below the mean in two core areas of language, assessed via the CELF 4 (Semel et al., 2006). Although non-verbal intelligence cut off points in previous research are often set at one standard deviation below the mean, giving a score of 85, as intellectual disability has a cut-off point of 70, the current thesis will include participants who score 70 or above in non-verbal intelligence testing, indicating their language disorder is not due to an intellectual disability. Language impairments are rarely 'specific' so participants have been included with co-occurring conditions (as long as language difficulty is identified as a primary need), with the exception of autism spectrum disorder. This section has highlighted the vast heterogeneity of the language disorder population. Through being as inclusive as possible in participant inclusion criteria, the current thesis can therefore be considered representative of the language disorder population.

Having discussed the considerable heterogeneity within the language disorder population, which has influenced the participant inclusion criteria used within the current study, this chapter will move on to discuss working memory. A main focus of the remainder of this chapter will be the 'Multi Component Model of Working Memory' (Baddeley & Hitch, 1974; Baddeley, 2000), as the central executive component of the working memory system is central to this research.

1.2 Working Memory

Working memory allows individuals to store information for short periods of time and use it for current thinking (Holmes, 2012). Working memory can be seen as a mental workspace which encompasses a number of skills used in daily life (Henry, 2012; Holmes, 2012). It is fundamental for tasks such as reading comprehension, mathematics/arithmetic, planning of thoughts and actions and maintaining focused behaviour (Holmes, 2012). A more demanding task that requires increased concentration and effort will draw upon working memory resources more strongly than a less demanding task; and it has been demonstrated that working memory is closely linked to educational achievement (Henry, 2012). Although crucial to particular tasks, working memory does have a limited capacity in terms of the amount of information that it can store and manipulate. This limited capacity can change over a lifetime, and although it steadily increases in typically developing children until the age of 14 or 15 years to reach adult levels, the difference between individuals, even of similar ages, can vary greatly (Holmes, 2012).

There have been a number of approaches proposed to understand working memory, i.e. the 'Embedded Processes Model of Working Memory' devised by Cowan (1988), that postulates activated memory, focus of attention and long-term memory all contribute to working memory function. Alternatively, approaches such as Ericcson and Kintsch (1995) 'Long term Working Memory', suggest that experts and skilled performers have an expanded working memory capacity in activities that they have practiced and acquired knowledge and specialist memory skills. Various approaches to understand working memory have been proposed, however one of the most influential models in working memory research over the past several decades, which undoubtedly will continue to influence research in this field, is Baddeley and Hitch's (1974) 'Multi Component Model of Working Memory', and its subsequent revision by Baddeley (2000). Due to its considerable influence and as one of the components of this system, the central executive, is central to the research described in this thesis, the current research has adopted the Baddeley and Hitch (1974), and later revised Baddeley (2000), working memory model.

Although several approaches to working memory have been mentioned here, it is important to highlight a similarity between them all. These working memory models all assume working memory involves storage alongside attentional control. This is of particular relevance to this thesis as the current study focuses on attentional control and as discussed above, the central executive component of working memory is central to this research.

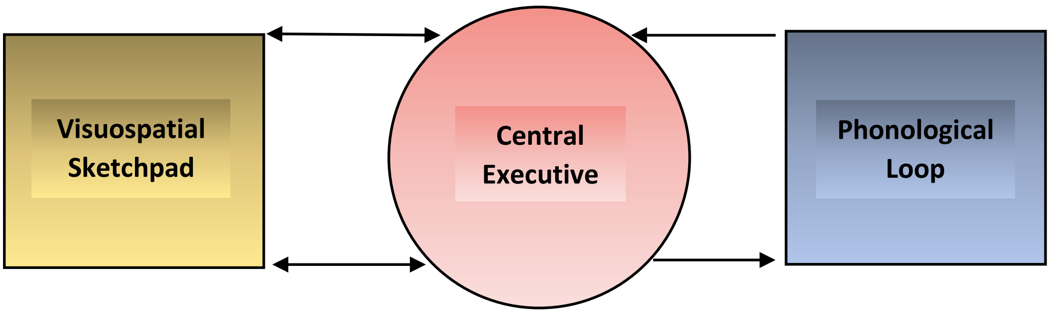
The following section will therefore give an overview of the 'Multi Component Model of Working Memory' (Baddeley & Hitch, 1974; Baddeley, 2000), giving a brief description of its four components, the central executive, the phonological loop, the visuospatial sketchpad and the episodic buffer, before moving forward to discuss how these components relate to language disorder.

1.2.1 The 'Multi Component Model of Working Memory' (Baddeley and Hitch, 1974)

The understanding of human memory has been a topic of discussion and research for decades. In 1949, Hebb proposed that there was a distinction between short-term memory, based on a temporary electrical activation, and long-term memory, based on a neuronal growth (Baddeley, 2003a). For the next decade, further support for this proposed distinction became apparent through research that showed that small amounts of information could be easily forgotten if they were not actively rehearsed (Baddeley, 2003a). It was proposed that this may be due to a unitary long-term memory system. However, by the late 1960s further models had emerged in an attempt to understand short-term memory. Atkinson and Shiffrin's (1968) 'Multi Store Model of Memory' described short-term memory as a unitary system with a limited store, holding information for relatively short periods of time with little processing. The 'Multi Store Model of Memory', also referred to as 'The Modal Model' consisted of 3 stores: sensory memory, short-term memory and long-term memory. According to this memory model, information enters the sensory memory via the sensory organs and if attended to then enters the short-term memory. If this is then rehearsed through repetition, information from the short-term memory then enters the long-term memory. If the information is not maintained through rehearsal, the information is forgotten and lost from the short-term memory through decay (McLeod, 2017). The 'Multi Store Model of Memory' described memory as an information processing model whereby information passes from store to store in a serial manner.

Assumptions were that patients with short-term memory impairment would have a very limited capacity for long-term learning and everyday cognitive tasks (Baddeley, 2003a). Through a series of tasks that individuals undertook in order to deplete the availability of their short-term memory, it was found that although a clear decrement occurred for long-term memory, it was far from catastrophic. The linear approach of the 'Multi Store Model of Memory' (Atkinson & Shiffrin, 1968) was therefore deemed too simplistic by Baddeley and Hitch (1974). This led to their proposal of the original 'Multi Component Model of Working Memory' (see figure 1.1 below), which consisted of three main components, the central executive, the phonological loop and the visuospatial sketchpad.

Figure 1.1 Baddeley and Hitch's (1974) Original 'Multi Component Model of Working Memory'



Each component of Baddeley and Hitch's (1974) 'Multi Component Model of Working Memory' will be discussed in turn, starting with the central executive, the area of the model responsible for attentional control and the allocation of resources within the working memory system.

1.2.2 The Central Executive

Baddeley and Hitch's (1974) original working memory model consisted of a limited attentional capacity control system, termed the central executive, which was assisted by two further components, the phonological loop, for holding speech-based information, and the visuospatial sketchpad, for holding visual and spatial information (Baddeley, 2003a; Henry, 2012). The central executive component is seen as the most important part of the working memory model and relies heavily but not exclusively on the frontal lobes (Baddeley, 2003b). The central executive is responsible for attentional control; it directs and allocates resources appropriately to ensure that goal-directed behaviour is achieved (Baddeley, 2003b; Henry, 2012). The central executive is seen as the 'brains' behind the working memory system, whereas the phonological loop and visuospatial sketchpad act as slave subsystems that only hold information passively (Henry, 2012).

Although viewed as the most important component of the working memory model for cognition, the central executive component has historically been the least understood part of this model (Baddeley, 2003a). It is known to provide overall regulation and to control the allocation of resources within the working memory system, which is agreed to be achieved by focusing, dividing and switching attention as required (Henry, 2012).

In contrast to the central executive component of the working memory model, there is more detailed knowledge about the phonological loop (Rudner & Rönnberg, 2007), a 'slave' subsystem of the working memory model that holds speech-based information. The following section will discuss the phonological loop in further detail.

1.2.3 The Phonological Loop

The phonological loop is a specialised storage system for speech-based information and, in relation to language processing, the phonological loop has received the greatest research focus. The phonological loop acts as a slave subsystem within the working memory model; it has no capacity to control attention or influence our decision making abilities, and has a small temporary storage capacity for information that has been heard, particularly speech (Henry, 2012; Rudner & Rönnberg, 2007). The phonological loop comprises two subsystems, the phonological store and the articulatory rehearsal system. When speech enters the phonological loop, auditory memory traces are held within the phonological store, which decay after about two seconds, unless rehearsed via the articulatory rehearsal subsystem (Rudner & Rönnberg, 2007). Although memory traces can be refreshed via the retrieval and re-articulation of speech sounds within the articulatory rehearsal mechanism, phonological short-term memory has a limited span: as the number of items rehearsed increases, the first item/s will tend to drop off and fade before being rehearsed again (Baddeley, 2003a).

Many studies provide support for the phonological loop component of the working memory model. The acoustic similarity effect was demonstrated by Conrad (1964), who showed that errors in the recall of a string of visually represented letters were made more frequently if the letters sounded similar than if not (Baddeley & Larsden, 2007). Baddeley and Larsden (2007) state that using the term 'acoustic' may be misleading, as errors in this study occurred using visually presented stimuli. A new term, the 'phonological similarity effect', was suggested. This finding demonstrates that phonological short-term memory is affected by sound similarity and supports the existence of a phonological store subsystem within the phonological loop, as memory for speech material has been shown to utilise some sort of sound-based storage system (Henry, 2012).

Evidence for the articulatory rehearsal subsystem of the phonological loop has also been shown via 'the word length effect', which refers to the finding that when adults are asked to recall a list of words, their recall is better for shorter rather than longer words (Baddeley etal., 1975). Experiments on word length effects have found that memory span is sensitive to both the number of letters within a word and also to the number of syllables or phonemes within a word (Baddeley et al., 1975). It is argued that word recall is better for shorter words because the verbal rehearsal mechanism operates in real time; hence the articulatory rehearsal subsystem can rehearse shorter items more quickly than longer items, keeping more of them active in the phonological loop.

The following section will discuss the second 'slave' subsystem of the working memory model, the visuospatial sketchpad. In comparison to the phonological loop, which holds speech-based information, the visuospatial sketchpad stores visual and spatial information.

1.2.4 Visuospatial Sketchpad

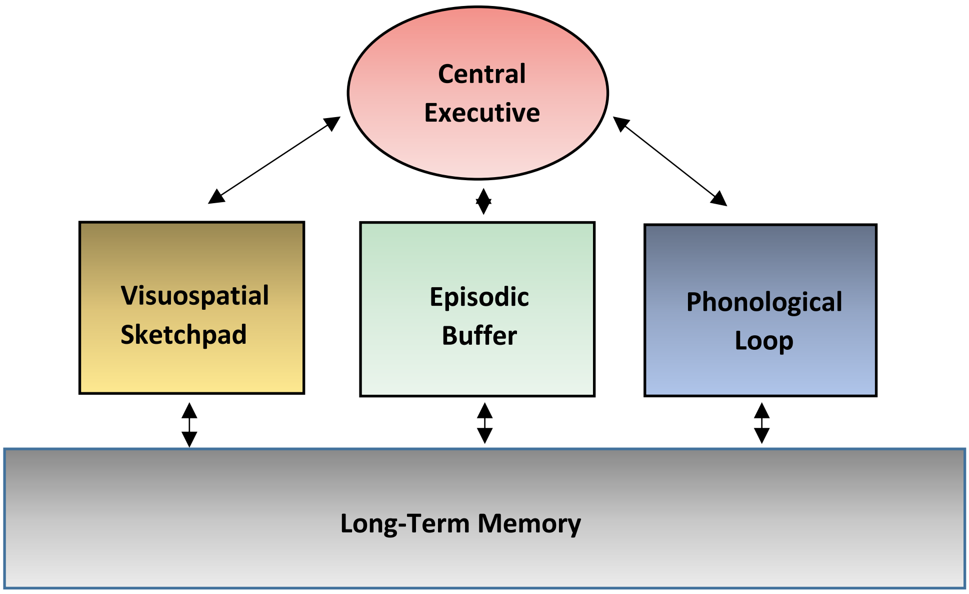
The second 'slave' subsystem of the working memory model is known as the visuospatial sketchpad, and this component stores visual and spatial information for short periods of time. Similar to its verbal equivalent, the phonological loop, the visuospatial sketchpad is limited in capacity and can typically hold about three or four items (Baddeley, 2003a). Again, this component is purely a temporary storage system with no control over attention. Visuospatial short-term memory and phonological short-term memory are seen as completely distinct from each other, with the visuospatial sketchpad relying on structures in the right hemisphere of the brain, and phonological short-term memory relying on structures in the left hemisphere, the area often associated with language (Henry, 2012). In parallel with the phonological loop, it has been proposed that the visuospatial sketchpad has a rehearsal process which, without use, leads to very rapid visuospatial memory decay (Baddeley, 2003a). Logie (1995) proposed the sketchpad had a visual storage component called a 'visual cache' and a retrieval and rehearsal process termed 'the inner scribe' (Baddeley, 2003a). The visual cache is seen by Logie (1995) as responsible for holding information about colour and form, whereas the inner scribe holds information regarding movement sequences. The inner scribe also carries out rehearsal in order to refresh information stored in the visuospatial sketchpad and reduce memory trace decay over time (Baddeley, 2003a; Henry, 2012). Similar to the central executive component of the working memory model, the visuospatial sketchpad would benefit from further research and improved understanding of its functioning (Baddeley, 2003a).

The central executive, phonological loop and visuospatial sketchpad are all components of the original 'Multi Component Model of Working Memory' (see figure 1.1), devised by Baddeley and Hitch (1974). Baddeley (2000) subsequently revised this working memory model through the addition of the episodic buffer, which was added to account for the impact of long-term memory and also the integration of all subsystems within the memory system. The following section will discuss the episodic buffer in further detail and the impact its addition has had to the working memory model.

1.2.5 The Episodic Buffer

Although Baddeley and Hitch's (1974) original working memory model became one of the most influential models and was very well received, there were areas of difficulty, particularly surrounding the impact of long-term knowledge on working memory, that led to criticism (Rudner & Rönnberg, 2007). Baddeley, therefore, revised his original working memory model by adding a fourth component, the episodic buffer (see figure 1.2 below), which took into account the function of long-term knowledge within the system (Baddeley, 2000).

Figure 1.2 Baddeley's (2000) revised 'Multi Component Model of Working Memory'



The episodic buffer has also been included to account for the integration of all subsystems within the working memory model, and also multidimensional representations from visual, phonological or semantic codes, visual, auditory or tactile perceptions and episodic or semantic sources (Rudner & Rönnberg, 2007). The episodic buffer has a limited amount of multi-modal storage, therefore, and acts as the link between other systems and long-term knowledge. This integration role has also been described as the 'binding' together of information from different subsystems, and it contributes to experiences being perceived as coherent. Therefore, the function of the episodic buffer is related to processing, storage, using long-term knowledge within the system, and the integration of information across working memory components. However, there remains relatively little research on this most recent addition to the working memory model (Henry, 2012).

Following the introduction of each component of the 'Multi Component Model of Working Memory' (Baddeley & Hitch, 1974; Baddeley, 2000), this chapter will continue by exploring the relationship between working memory and language.

1.3 Working Memory and Language Disorder

Working memory capacity has been shown to change over the lifespan and vary greatly between individuals. In terms of the current thesis, there is considerable evidence to suggest that difficulties with working memory are present in those with language disorders (Farquharson & Franzluebbers, 2014; Henry & Botting, 2017). Children with language impairments commonly demonstrate not only a marked difficulty in expressive and/or receptive language but also a deficit in working memory ability, despite having a normal hearing range and normal nonverbal intelligence (Montgomery et al., 2010). As the impairments in working memory commonly found in those with language difficulties are often seen as barriers to learning and educational achievement, this has led to widespread interest in this field (Holmes, 2012).

The current section will continue to explore working memory and language, firstly discussing phonological short-term memory and language disorder, mainly focusing on performance on non-word repetition tasks, which are often used as a clinical marker as part of the diagnosis of language disorder. This section will then explore the inconsistent findings regarding visuospatial impairments in those with language disorder and discuss whether visuospatial difficulties are in fact related to the central executive component of the working memory model rather than a language difficulty, therefore being more of an attentional control deficit. Focussing on the central executive, this section will then discuss the importance of executive-loaded working memory trained tasks, as implemented in the current study, and will explore impaired executive functions in those with language disorder and discuss the debate surrounding the causality of these impairments. This section will conclude by looking at the episodic buffer and its relationship to language disorder, discussing the limited research available, due to its relatively late addition by Baddeley (2000).

1.3.1 Phonological Short-Term Memory and Language Disorder

'Mapping Theory' proposes that for a child to be able to learn a language, they must be able to segment heard speech into individual words and morphemes, and attach this to a meaning / context (that allows a 'map' to occur) to each sound segment (Chiat, 2001; Claessen et al.,2013). It has been proposed that language difficulties may be explained by an impaired phonological processing ability in mapping phonological form onto meaning (Chiat, 2001). This highlights the importance of phonological processing skills and phonological memory in the segmentation and development of accurate phonological representations in the lexicon when learning new words (Claessen et al., 2013). A poor phonological sensitivity and a difficulty in perceiving and distinguishing speech sounds and representations may therefore explain deficits in those with language disorders (Henry, 2012).

Children with language disorder have repeatedly been shown to have difficulties in phonological short-term memory (Graf Estes et al., 2007; Stavrakaki, 2015). These findings have led many researchers to claim that deficits in tasks such as non-word repetition, a well-known measure of phonological short-term memory, are in fact the root cause of language disorders and a clinical marker for diagnosis (Henry, 2012; Stavrakaki, 2015). A key influential study in this area was undertaken by Gathercole and Baddeley (1990), which looked at the phonological short-term memory abilities of children with language disorders.

Six children aged eight years were assessed on a range of cognitive, language and phonological short-term memory assessments. Children with a language disorder were matched to two comparison subgroups: one matched for language ability (a younger age group to match on language age) and one matched for non-verbal ability (a similar chronological age group as non-verbal abilities remain relatively age-appropriate). Gathercole and Baddeley (1990) found, in relation to phonological short-term memory ability, that the language disorder group performed more poorly on memory span measures compared to both comparison groups. Non-word repetition task performance in the language disorder group was also found to be about four years behind chronological age level, with particular difficulties found for three or four syllable non-words. This suggests a possible difficulty for those with a language disorder when the capacity of the phonological store is over-stretched (Henry, 2012).

Replication and extension of Gathercole and Baddeley's (1990) study has been undertaken by many subsequent researchers in children from a range of ages (Dollaghan & Campbell, 1998; Laws & Bishop, 2003; Pickering & Gathercole, 2004.), and a meta-analysis of the findings from 23 separate studies indicated substantial impairment (mean effect size 1.27) in children with language disorder compared to age-matched controls (Graf Estes et al., 2007) Bishop, North and Donlan (1996) provided evidence that phonological short-term memory contributed to language disorder by including three comparison groups in their study: children with persistent language disorder; children with resolved language disorder; and children with a matched non-verbal IQ who had no language disorder (Bishop et al., 1996). By expanding Gathercole and Baddeley's (1990) original study design via including those with resolved language difficulties, Bishop et al. (1996) not only replicated the original findings, but reported that alongside those with existing language disorder, the resolved group also showed a significantly poorer performance on non-word repetition tasks compared to those with no language impairments. This indicates that although language difficulties can resolve, the original deficit in phonological short-term memory may still be apparent. This suggests that phonological short-term memory deficit is not purely a consequence of language disorder. An individual may be able to develop strategies to compensate for their language difficulties, however the original phonological short-term memory deficit can still remain (Henry, 2012).

In recent years non-word repetition ability has continued to be a focus within research to understand phonological short-term memory deficits in those with language disorder (Jackson et al., 2016; Jackson et al., 2019). In support of previous research, an initial study by Jackson et al. (2016) found that pre-school children with language disorder performed significantly less well on non-word repetition tasks than their typically developing peers. The length of non-words has been found to play a pivotal role in repetition rates for those with language disorder, with a greater rate of accuracy found in shorter non-words, which has even been comparable to the rate of accuracy found in typically developing children (Coady & Evans, 2008; Jackson et al., 2019). In contrast, as mentioned briefly above, a meta-analysis study conducted by Graf Estes et al., (2007), found children with language disorder performed on average 1.27 standard deviations below typically developing children in non-word repetition tasks. Across 23 studies, children with language disorder portrayed difficulty on even short non-words, however greater difficulty was apparent with longer non-words (Graf Estes et al., 2007). Recent research by Jackson et al. (2019) has also explored the relationship between phonological short-term memory, word learning and word length in five year old children with and without language disorder. Twenty-three children with language disorder and 26 typically developing five-year-olds undertook a fast mapping task, test of non-word repetition and receptive vocabulary. As expected from previous research, the children with language disorder performed significantly less well than typically developing peers in non-word repetition tasks. It was also hypothesised, to complement previous research, that those with language disorder would perform better and show a greater accuracy in repetition of non-words that contained fewer syllables (Jackson et al., 2019). Children with language disorder, however, were found to be significantly less accurate in repetition for all non-word lengths compared to their typically developing peers. Jackson et al. (2019) highlight that although a group significant difference was present across all word lengths, there was a greater effect size for words of three or four syllables, showing that the deficit was greater in those with language disorder when non-word length increased. The difference in recent findings of length effects in non-word repetition was suggested to be due to the age of the children used in Jackson et al.'s (2019) study. The participants were five years of age which could reflect a developmental characteristic of phonological short-term memory capacity size. A further suggestion for this difference was that phonological short-term memory capacity may not be the underlying reason for non-word repetition deficits in those with language disorder (Jackson et al., 2019). Jackson et al. (2019) question whether a broader phonological processing difficulty that presents general deficits in phonological encoding, phonological representation and perception may be apparent. Linking back to section 1.1 of this chapter, which discussed language disorder and heterogeneity, Jackson et al. (2019) also discuss whether the differences found between their study and others in non-word repetition accuracy may be solely due to the heterogeneity of the language disorder population and the sample differences of various research studies. It is also important to note at this point, that not all children who meet the criteria for language disorder have deficits in non-word repetition tasks, and that not all children who have non-word repetition deficits meet the criteria for language disorder. Gathercole (2006) relates non-word repetition ability to word learning ability in both typically developing children and those with language disorders. In terms of non-word repetition as a measure of phonological short-term memory, Gathercole (2006) focuses on phonological storage as a key aspect in non-word repetition ability and vocabulary acquisition. Therefore, if not all children with language disorder present with non-word repetition deficits, and vice versa, this indicates that low language performance could arise from other deficits apart from a purely limited phonological short-term memory.

Current research in this field shows that there are still many areas to be explored in phonological short-term memory and language disorder. Overall, the evidence for difficulties in phonological short-term memory, for the majority of children with language disorder, is robust, extensive and reliable and suggests phonological short-term memory could provide a useful assessment of a language related ability that might be affected by interventions that target the central executive.

The following section will discuss the visuospatial sketchpad and its relationship to language disorder. In contrast to our understanding of the role of phonological short-term memory within working memory and its links to language disorder, our understanding of the visuospatial domain is much less extensive.

1.3.2 Visuospatial Sketchpad and Language Disorder

There is growing evidence that non-linguistic factors may play an important role in contributing to the working memory difficulties of those with a language disorder (Vugs et al., 2013).However, research findings are inconsistent about whether visuospatial ability is impaired, and the relationship between the visuospatial sketchpad and language disorder is still debated. For example, no significant difference was found in visuospatial short-term memory in a group of 15 nine-year-olds with language disorder when compared to either age- or language-matched controls (age approximately six years) on a dot matrix task (Archibald & Gathercole, 2006a). Interestingly, however, this study did find that the language-matched controls, who were approximately three years younger than children with a language disorder, performed significantly less well on the dot matrix task. This suggests that the visuospatial domain may not be related to language at all and differences in ability may be due to developmental factors (Henry, 2012).

Likewise, a study by Archibald and Gathercole (2006b) found no significant group differences in performance on a visuospatial storage task between those with a language disorder and controls. However, when the scores of individuals with language disorder were considered separately, over 50% performed outside of the average range, suggesting that a visuospatial storage deficit may be present in some children with language difficulties (Archibald & Gathercole, 2006b).

A longitudinal study conducted over a year looking at verbal short-term memory, visuospatial short-term memory and visuospatial processing ability at three time points in children with language disorder, compared to typically developing children, also found that those with a language disorder had slower visuospatial short-term memory development (Hick et al., 2005).

A meta-analysis of the literature on visuospatial working memory and those with language disorder, conducted by Vugs et al. (2013), provides a useful overview of relevant findings. The meta-analysis examined 21 studies and concluded that those with language disorder performed on average half a standard deviation below their typically developing peers on visuospatial working memory tasks, suggesting that a difficulty in this area may exist for those with language impairments (Vugs et al., 2013). However, there were large effect size differences between studies within the meta-analysis, which again could be linked to heterogeneity among those with language disorder, with studies comparing groups that differed in diagnostic criteria or co-existing difficulties. It was suggested by the authors that participants included in studies with larger effect sizes were those with a more pervasive language impairment across two or more language domains. However, the small sample sizes of many of the studies in the meta-analysis means that this hypothesis needs to be investigated further (Vugs et al., 2013).

A recent study explored visuospatial working memory in 78 children with language disorder and 39 typically developing children aged 5-8 years over three time points, each at one-year intervals (Blom & Boerma, 2020). Children performed a dot matrix task where they were required to identify the location of a series of dots in reverse order. Results indicated that across the three time points, children with language disorder performed similarly to their typically developing peers. (Blom & Boerma, 2020). Initially, these findings seem to contradict Vugs et al's. (2013) meta-analysis where those with language disorder were shown to display some difficulty with visuospatial working memory tasks. However, in addition to exploring visuospatial working memory performance, Blom and Boerma (2020) considered whether differences between those with language disorder and typically developing children were influenced by severity and persistence of language disorder. They found that language ability within the language disorder group significantly predicted the variance in working memory function, indicating that children with a less severe form of language disorder did not experience visuospatial working memory difficulties (Blom & Boerma, 2020). Children with less persistent language disorder were also shown to perform better in visuospatial working memory tasks than those children in the more persistent language disorder group. According to Blom and Boerma (2020), these findings indicate that children with a more severe and persistent form of language disorder do experience difficulties in their visuospatial working memory but those with a milder and less persistent form of language disorder do not. The conclusions drawn from Blom and Boerma's (2020) recent study could therefore help to explain the inconsistencies found in visuospatial working memory and language disorder research. This again highlights the importance of understanding the heterogeneity among those with language disorder and how such differences between individuals from the language disorder population can alter results between research studies.

Visuospatial domain deficits have also been suggested to relate to the central executive component of the working memory model, rather than to a language disorder, arising from an attentional control deficit which affects an individual's capacity in choosing what to pay attention to (Vugs et al., 2013). A study by Marton (2008) compared children with a language disorder who had either good or poor attentional control. Children with poor attentional control were shown to have poorer performance on visuospatial working memory tasks. If this is the case, an intervention that targets central executive functioning and improves attentional control should lead to improvements in visuospatial performance. Findings may therefore be informative about the relationship between attentional control, visuospatial abilities and language disorder.

The next section of this chapter will explore the central executive and language disorder, as the research intervention in this thesis targets central executive functioning this component of the working memory system is of particular relevance.

1.3.3 Central Executive and Language Disorder

The central executive decides which information an individual will attend and which parts of the working memory to send information to, in order for it to be dealt with (McLeod, 2012). A common feature found amongst effective working memory intervention studies across various subgroups of individuals, is the inclusion of executive-loaded working memory trained tasks, i.e. training that taps into both attentional and processing resources under executive control and not just the storage of information (Rowe et al.,2019b). It has been evidenced that frequent rehearsal of executive-loaded working memory tasks, instead of just practicing storage only short-term memory tasks, could help improve both efficiency of processing and also facilitate the storage of information within working memory (Rowe et al., 2019b).

Executive-loaded working memory tasks have been examined in both the verbal and visuospatial domains. The Competing Language Processing Task (CLPT: Gaulin & Campbell, 1994), involves children saying whether a sentence is true or false and then remembering the final word from each sentence. Ellis Weismer, Evans and Hesketh (1999) used the CLPT and reported that, although similar results were found between groups in terms of the true/false component of the task, i.e. the processing aspect, children with language disorders recalled significantly fewer words than typically developing age-matched controls. However, as the majority of the children managed both of the processing and recall aspects of the task, if the demand was low, it has been suggested that those with language disorders have a reduced attentional capacity compared to their typically developing peers (Montgomery et al., 2010). A very similar measure, termed the 'listening span' task, requires the same combination of processing and storage skills that are the hallmark of executive-loaded working memory tasks (Henry, 2012). According to Henry (2012) the 'listening span' task is probably one of the most commonly used measures of complex memory span available within developmental literature, with standardised versions available. Henry and Botting (2017) discuss that although the majority of existing evidence relates to those with language disorders having a poorer verbal short-term memory and verbal central executive, there is emerging evidence that indicates wider difficulties may be present with visuospatial short-term memory and visual central executive. An example of a key task that measures complex span in the visuospatial domain, is the 'Odd One Out' task, where participants are shown three items in a 3x1 array and asked to point to the item which is seen as slightly different from the other two, i.e. the 'odd one out' (Henry, 2001; Henry, 2012). In addition to the processing aspect of this task, the child is asked to remember the spatial location of the item selected and recall this on an empty grid, with trials increasing in length as task difficulty increases (Henry, 2012).

Language disorder research has increasingly been looking at the central executive component of working memory, and, in particular, aspects such as attentional capacity, allocation, inhibitory control and working memory updating (Montgomery et al., 2010).

The central executive is responsible for the control and regulation of cognitive processes, i.e. executive functions, which are often linked to the function of the frontal lobes (Miyake et al., 2000). The relationship between three executive functions: shifting, updating and inhibition were explored in an individual difference study, that questioned to what extend different functioning to the central executive could be considered unitary and to what degree different executive functions reflected the same underlying mechanism or ability (Miyake et al., 2000). Shifting relates to the ability to devote attention between two different tasks or two different levels of a task, working memory updating is a cognitive function that adds new content to the focus of a task, and inhibition relates to the ability to prevent incoming irrelevant information entering the current focus of attention (Ecker et al., 2010; Montgomery, 2012). A total of 137 undergraduate students completed nine tasks that were hypothesised to tap into one of the three target executive functions, alongside five complex executive tasks: Wisconsin Card Sorting Test (WCST), Tower of Hanoi (TOH), random number generation (RNG), operation span, and dual tasking. Confirmatory factor analysis (CFA) indicated that all three executive functions, (shifting, updating and inhibition) were distinct but not completely independent of each other, sharing underlying commonalities, therefore indicating both unity and diversity of executive functions (Miyake et al., 2000). Further exploration indicated that different executive functions contributed to the performance of different tasks, i.e. shifting ability contributed to WCST performance and inhibition ability contributed to the proficiency in solving the TOH puzzle (Miyake et al., 2000). Overall, Miyake et al.'s (2000) study indicated that the three executive functions, shifting, updating and inhibition, contributed differentially to performance on commonly used executive tasks even though they were moderately correlated to one another.

There are indications that there could be impaired shifting, working memory updating and inhibitory control in children with language disorders (Im-Bolter et al., 2006; Pauls & Archibald, 2016). Im-Bolter et al. (2006) found that those with a language disorder presented comparable shifting ability to their peers, which suggests that there is not a deficit in allocation ability. The language disorder group did, however, show a poorer updating of working memory and poorer inhibitory control compared to a typically developing group, therefore implying that there is a difference between attentional capacity and this supports the reduced attentional capacity in those with language disorders as proposed by Ellis et al. (1999). Similar results were found by Henry et al., (2012): they reported several executive difficulties in a sample of children with language disorder (executive-loaded working memory, inhibition, planning and fluency), which were not restricted to verbal-based tasks.

A meta-analysis of 46 studies examined cognitive flexibility and inhibition control in 4-14-year-old children with and without language disorder (Pauls & Archibald, 2016). Children with language disorder were found to perform more poorly than typically developing peers in both cognitive flexibility and inhibition tasks. It has been suggested that poor cognitive flexibility may be accounted for by inhibitory control deficit: if a former task cannot be completely cleared from the focus of attention, then there could be difficulty in switching to the next task (Pauls & Archibald, 2016). Pauls and Archibald (2016) also discuss that poor inhibitory control in those with language disorder may be the result of a difficulty in the ability to suppress irrelevant acoustic information within the environment resulting in an interference with the perception of language and the ability to form accurate phonological representations. This perspective therefore indicates that a deficit within inhibitory control could also impact language acquisition. Alternatively, phonological short-term memory deficits in those with language disorder have been suggested to be the reason for poor inhibitory control (Pauls & Archibald, 2016). Having a clearer understanding of task demands, enables typically developing children to enhance their inhibitory control task performance (Pauls & Archibald, 2016). Alternatively, a child with language disorder may only be able to maintain a weaker representation of the task demands, negatively impacting their performance level.

Raising further issues for this debate, tasks of both verbal and non-verbal inhibition have produced inconsistent results with some studies finding deficits in inhibition regardless of verbal demand (Henry et al.,2012), whereas other studies have found inhibition deficits only in verbal based tasks (Lukacs et al., 2015). A recent study looking at interference control in children with language disorder found that, compared to typically developing peers, those with language disorder again displayed deficits (Blom & Boerma, 2020). However, those with a milder form of language disorder performed better than those with a more severe and persistent language difficulty, again questioning whether impairments to the central executive are the cause of language disorders, or the language disorders cause impairments to the verbal central executive (Blom & Boerma, 2020; Pauls & Archibald, 2016; Pimperton & Nation, 2012).

Overall, there is clearly increasing evidence that children with language disorder are at a higher risk of domain general difficulties with executive functioning that go beyond their working memory difficulties. With discussion surrounding whether impairments to the central executive cause language disorders or vice versa, the current research intervention, that targets the central executive in children with language disorders, is one way to explore causal relations and gain new insights into this debate.

This chapter will conclude by discussing the episodic buffer and its relationship to language disorder. As the episodic buffer is the latest component to be added to the working memory model, we may expect there to be less research on its role in language disorders in comparison to the three original components already discussed in this chapter.

1.3.4 The Episodic Buffer and Language Disorder

Research within the episodic buffer and language disorder has mainly concerned sentence recall tasks to measure episodic buffer functioning, where individuals are required to repeat a sentence immediately after hearing it (Henry & Botting, 2017). Although limited research is available, findings have indicated impaired performance in those with language disorder (e.g. Archibald & Joanisse, 2009). It has however, been argued that the ability to repeat sentences is strongly dependent on having a familiarity with lexical phonology and morphosyntax, while familiarity with semantics and prosody played a small role (Polišenská et al., 2015). Therefore, further research is needed to confirm which aspects of language knowledge are in fact involved in sentence recall tasks and whether other components in the working memory system may affect task performance (Henry & Botting, 2017).

1.4 Chapter Summary

Overall, working memory difficulties, particularly with phonological short-term memory, seem to have an association with language disorders. Evidence is emerging that verbal and in some cases visuospatial short-term memory difficulties may be present in some individuals with language disorder as well as a reduced capacity on executive-loaded working memory, inhibition, planning and fluency tasks. This suggests that there are clear working memory difficulties in children with language disorder in several of the components of the working memory system. The working memory model proposed by Baddeley and Hitch (1974), and its subsequent revision (Baddeley, 2000), is the focus of this thesis as not only does the current intervention target the central executive component of the working memory model, the model itself has been extremely influential in working memory research over several decades.

Phonological short-term memory in particular seems to have an association with language disorders. Due to the robust evidence linking phonological short-term memory and language impairments, it has even been suggested that phonological short-term memory deficits (usually assessed using non-word repetition tasks) are a marker for language disorders. Current research continues to use non-word repetition tasks as a measure of phonological short-term memory, with a recent focus upon non-word length disproportionately affecting repetition accuracy in those with language disorder (Coady & Evans, 2008; Jackson et al., 2019). In those with resolved language disorders, a clear deficit in phonological short-term memory remains apparent (Bishop et al., 1996). This suggests that for some, compensatory strategies or strengths in other areas of language processing can be developed and implemented to deal with language difficulties but the original working memory deficit may continue to be ongoing (Bishop et al., 1996). This provides support for Reilly et al. (2014a)'s argument that even those with language abilities one standard deviation below the mean, therefore with a mild/below average language impairment, should be monitored and supported. It may be the case that these children have developed compensatory strategies to improve their language ability, but working memory difficulties may still be present in this subgroup. These findings also suggest that this component of the working memory system may be resistant to change by interventions.

Visuospatial short-term memory deficits in those with language disorders still remain under debate, although emerging evidence suggests milder difficulties may be present in at least some of those with language disorder. Recent research has explored the heterogeneity amongst those with language disorder (Blom & Boerma, 2020). Although at group level those with language disorder have been found not to display any visuospatial working memory difficulties when compared to their typically developing peers, when exploring within-group performance, severity and persistence of language disorder does seem to effect visuospatial difficulties (Blom & Boerma, 2020). This has been suggested to account for the conflicting research findings within visuospatial performance and language disorder literature (Blom & Boerma, 2020). When considering executive functioning and the role of the central executive component, allocation of attention does seem to be relatively comparable to typically developing peers. However, those with language disorders show a reduced capacity on executive-loaded working memory, inhibition, planning and fluency tasks, and these difficulties are not restricted to tasks that are language heavy – i.e. they appear to be domain general.

This suggests that there are clear working memory difficulties in children with language disorder in several of the components of the working memory system. The causal relationships of these difficulties however have often been debated. Implementing a working memory intervention to children with language disorder, which targets the central executive and focuses on training executive-loaded working memory, may help to investigate whether executive function can be improved and if so, whether any such improvements impact on other aspects of working memory and language outcome measures, such as language comprehension. If improvements do occur following executive-loaded working memory training, this could be due to executive-loaded working memory being the source of language deficits and therefore the intervention itself has addressed this, or alternatively, improvements in executive-loaded working memory could enable children to make more effective use of their reduced language processing and knowledge. An intervention therefore, that focuses on executive-loaded working memory tasks in both the verbal and visuospatial domains of the central executive, could be considered beneficial to children with language disorder.

A main theme throughout this introductory chapter has surrounded the heterogeneity among those affected with language disorder and the discrepancy surrounding its diagnostic criteria. Variable ‘cut off points’ in both standardised language tests and nonverbal intelligence tests, as well as variations in the assessments used to make language disorder diagnosis, are a major underlying factor in the heterogeneity and differing sample profiles often found within research studies.

As longitudinal studies have shown, risk factors for detrimental effects in later outcomes in terms of social, employment and educational achievements, are higher for those who have a language ability one standard deviation below the mean during their early life (Reilly et al., 2014a). As explained previously, participants within this current study therefore, have been selected on the basis of having a score of one standard deviation below the mean in two core areas of language, assessed via the CELF 4 (Semel et al., 2006). The inclusion of a wider population of children with language difficulties within this thesis will allow exploration of whether language impairment severity has an impact on improvements made through working memory intervention.

A further source of discrepancy that has been discussed concerns the nonverbal intelligence test ‘cut off points’ that are used to determine that a language impairment cannot be due to a deficit in general intelligence. As previously discussed, the most common nonverbal intelligence score used in diagnostic criteria is 85 (one standard deviation below the mean). However, with a score of 70 or below used in the diagnosis of an intellectual disability, other researchers commonly use a lower cut off point (Spaulding et al., 2008). Friel-Patti (1999) discusses that the performance of language skills or response to intervention generally does not differ significantly between children with a nonverbal IQ score above 85 or for those who score between 70 and 84. For the purpose of this research, children who scored above 70 on their nonverbal intelligence testing were included in this intervention study. This score ensures that no participants have an intellectual disability while enabling investigation of the effects of nonverbal intelligence as well as language level on children's ability to benefit from working memory training.

Recent international debate regarding appropriate terminology for language disorder has highlighted that language impairments are rarely 'specific'. In light of this, although participants were recruited on the basis of having a language difficulty as a primary need, individuals who had any other co-occurring conditions, with the exception of autistic spectrum disorder, were included in this sample. Given the well-established heterogeneity among those with language disorder, an inclusive sample will be more representative of the language disorder population.

In conclusion, based on modelling of working memory and previous research, an intervention that focuses on executive-loaded working memory tasks in both the verbal and visuospatial domains of the central executive would provide evidence of whether and how training in each affects working memory and language ability, which would in turn inform our understanding of the relationship between verbal executive-loaded working memory, visuospatial executive-loaded working memory, and language.

Chapter 2

Working Memory Training, Critiques, Pessimism and Moving Forward

2.1 Introduction

Due to the association between working memory and higher level cognition (chapter 1), a growing body of literature has investigated targeted training that aims to improve working memory capacity and functioning (Morrison & Chein, 2011). The idea of expanding available central 'workspace' has generated enormous interest. The general view is that limitations in working memory capacity have a wide reaching effect on other aspects of cognition, and working memory training could benefit different aspects of learning and cognitive skills (Melby-Lervag et al., 2016; Morrison & Chien, 2011). Working memory capacity is seen as a strong predictor of a range of cognitive abilities, including reading, comprehension, language acquisition, nonverbal problem solving and mathematical skills (Morrison & Chien, 2011). The rationale is that if these areas of higher level cognitive functioning are constrained via working memory limitations, increasing working memory capacity should produce transfer effects to diverse untrained tasks (Melby-Lervag et al., 2016).

This chapter will firstly focus on two ongoing debates within working memory research, whether working memory should be considered domain-general or domain-specific, and whether working memory training with strategy instruction is more beneficial than training without strategy instruction. The benefits of near and far transfer effects to evaluate the effectiveness of working memory training will be considered, as will methodological critiques, arising from previous working memory intervention research, and the need to address these factors in future working memory intervention study designs. Pessimism regarding outcomes of traditional computer-based working memory interventions will be discussed before focusing on recent research which has looked at the effectiveness of non-computerised working memory interventions in children's real world contexts. This chapter will review the limited literature available regarding working memory training with children with language disorder and will conclude with how the current research design attempts to address concerns raised in previous literature in terms of methodological critiques, training approaches and understudied target groups.

The following section of this chapter will explore whether working memory is domain general, domain specific, or a combination of the two by discussing previous research studies that have administered dual tasks and explored individual differences.

2.1.1 Domain Specific vs Domain General Working Memory Storage

Research over many decades has debated how information is maintained within the working memory system. One perspective is that items from different sensory domains are retained within separate working memory stores and are maintained under distinct mechanisms (Li et al., 2014). This approach is the domain-specific view of working memory storage and is compatible with Baddeley and Hitch's (1974) 'Multi Component Model of Working Memory', discussed previously in section 1.2.1. This model consists of a series of specialized domain-specific buffers which are driven by an attentional system, the central executive, which controls processing (Uittenhove et al., 2019).

In contrast, the domain-general view of working memory storage argues for the existence of a unitary working memory maintenance system. This postulates a single central system driven by an attentional resource that is in charge of the maintenance of any type of information (Li et al., 2014; Uittenhove et al., 2019).

The question of whether working memory operates via a domain-general or domain-specific approach has traditionally been explored through administering dual tasks and looking at individual differences. Children with various kinds of learning disabilities often display difficulties across both verbal and visuospatial working memory domains. However when looking more closely, those who have reading difficulties seem to show more difficulties in verbal working memory (Swanson et al., 2009), whereas visuospatial difficulties have been postulated to relate more closely to mathematical difficulties (Swanson & Jerman, 2006), supporting a domain-specific approach. Children with language disorder also seem to show larger difficulties in verbal working memory, although there is emerging evidence that indicates wider difficulties may also be present with visuospatial working memory, see section 1.3.2. (Blom & Boerma, 2020; Henry & Botting, 2017). Dual tasks examine the ability to simultaneously retain dissimilar material presented in different modalities and from different domains, such as verbal and visuospatial information (Uittenhove et al., 2019). A domain-specific perspective would expect little interference during dual tasks due to the separate subsystems within working memory, whereas a domain-general model would expect a strong interference if both processes are attempted simultaneously.

Conflicting research outcomes have led to a continued debate in this area. Dual tasks which examined both verbal and visuospatial task functioning have revealed that when performing an ongoing verbal task, visuospatial ability was not affected (Cocchini et al., 2002). However, others suggest that visual and verbal domains do compete for a central pool of domain-general resources after performing dual tasks which manipulate cognitive load of concurrent verbal and visuospatial activities (Vergauwe et al., 2010).

It has also been suggested that working memory is constrained by both domain-specific and domain-general sources (Fougnie & Zughni, 2015). When auditory and visual information has been required to be maintained simultaneously, the costs appear to be much less if the information is from the same modality. However, combining high auditory and visual loads together also appears to have a cost, possibly suggesting that both domain-specific and domain-general resources are working collaboratively (Fougnie & Zughni, 2015).

Conducting an executive-loaded working memory intervention that focuses on two different domains, verbal and visuospatial, will ensure that we are targeting both domains, whilst keeping neutral about whether working memory is domain general, domain specific, or a combination of the two. This was regarded as the best option, given that there is no general agreement in the literature on this issue.

In addition to whether working memory should be considered domain-general or domain-specific, a further issue regarding working memory training interventions is whether training with strategy instruction is more beneficial than training without strategy instruction (Peng & Fuchs, 2015). The following section will consider this, discussing previous research studies that have investigated the effects of working memory training with and without strategy instruction. Training and encouraging the use of memory strategies with an aim to improve working memory ability will also be discussed as will evidence of the natural use of strategy mechanisms, which have been suggested to enhance task performance.

2.1.2 Working Memory Training - With or Without Strategy Instruction?

Strategy instruction training involves teaching trainees effective approaches to encode, maintain and retrieve information from their working memory. The primary aim of strategy instruction training is to increase the performance level within tasks that require retention of information over a delay (Morrison & Chien, 2011). Trainees are shown a particular task strategy and then given a number of practice sessions in order to encourage the strategy of interest, ranging from articulatory rehearsal to encoding (Morrison & Chien, 2011).

In contrast, training without strategy, involves trainees undertaking working memory trained tasks without being given any method or strategy instruction to do so. Many researchers who train working memory without strategy instruction base their approach on Engle and Kane's (2004) 'Capacity Theory' (Pengs & Fuchs, 2015). Capacity theory views working memory as a mental space that can be expanded, with the purpose of training to increase working memory size rather than increase efficiency (Peng & Fuchs, 2015). Alternatively, Strategy Mediation Theory views working memory as having a relatively fixed cognitive capacity with performance determined by its efficiency (Peng & Fuchs, 2015). Strategy Mediation Theory suggests that working memory training generates the adoption of task specific strategies that can be used to facilitate performance on trained tasks and untrained variants (Fellman et al., 2020). A recent study has analysed Strategy Mediation Theory by performing a 6 week 'n back' working memory training intervention with adult participants aged 18-50 years (Fellman et al., 2020). All participants were randomly allocated to either a strategy training group, traditional training group (no strategy) or a passive control group (no training) (Fellman et al., 2020). Results indicated that the strategy group exhibited a steeper improvement curve than the traditional training group following the first training session. However, this advantage was not maintained by the strategy group as their initial advantage was found to be very short-term and not present in either intermediate or post-test time points (Fellman et al., 2020). Although incorporating strategy training into working memory training was found to be beneficial initially, Fellman et al.

(2020) suggest that the advantage for those given a strategy during training is relatively short lived. It should be clarified, that Fellman et al.'s (2020) study focused on adult participants rather than children, therefore limiting its relevance to the current study.

Further research has also explored the effects of teaching memory strategies, without the training of specific working memory tasks, by encouraging the use of rehearsal, visual imagery, creating stories and grouping (St Clair-Thompson et al., 2010). As young children do not generally spontaneously employ memory strategies, as rehearsal strategies for example do not emerge until a child is approximately seven years of age, training younger children memory strategy techniques has been suggested to be beneficial (St Clair-Thompson et al., 2010). The study recruited 254 participants, aged 5-8 years old, who were tested on measures of the phonological loop, visuospatial sketchpad and central executive components of the working memory model. Subgroups of children also completed tasks in following instructions, mental arithmetic and reading (St Clair-Thompson et al., 2010). Memory Booster, a computer game that teaches memory strategies, was used for 6-8 weeks with half of the study's participants within their own school environments. For those who completed memory strategy training, a significant improvement was evident for tasks which assessed both the phonological loop and central executive component of working memory. Memory strategy trained children also showed significant improvements in their ability to follow instructions and in mental arithmetic in the classroom (St Clair-Thompson et al., 2010). Improvements were not evident however in standardised tests for reading, mathematics or arithmetic. Therefore, although results from this study indicated that memory strategy training benefitted aspects of working memory ability and performance in the classroom for trained children, the benefit to performance on standardised tests appears minimal (St Clair-Thompson et al., 2010).

Further studies have also investigated the effects of working memory training with or without strategy instruction (Peng & Fuchs, 2015). A study conducted with 58 school aged children aged approximately 6-7 years, randomly allocated each child to one of three training groups, training with a rehearsal strategy instruction, training without strategy instruction and a control group. Children within the two trained groups were given one to one training sessions focussing on verbal working memory. Sessions lasted 35 minutes and were administered for 10 consecutive school days (Peng & Fuchs, 2015). Results indicated that those within the rehearsal strategy group significantly improved on all verbal working memory tasks administered across the 10 sessions. Those without strategy instruction significantly improved in a calculation span and puzzle task (Peng & Fuchs, 2015). Although children within the rehearsal strategy group used strategies for 99% of the training, children who were shown no strategies were also found to adopt them in 28% of their training tasks. Of the children who adopted strategies without instruction, 59% performed a rehearsal strategy. As children in the training without strategy group particularly used rehearsal strategy without any explicit instruction to do so, this suggests that rehearsal may be a more natural form of strategy to employ than other types (Peng & Fuchs, 2015).

Adopting strategy without specific instruction to do so could also relate to the ‘Overlapping Waves Model’ (Siegler, 1996), According to this model, children have a number of different strategies in their repertoires at any given time, however, they transfer reliance to different strategies with age and experience (Kwong & Varnhagen, 2005). As children learn a task, they show increasing reliance to more effective strategies and decrease their reliance on less effective strategies (Kwong & Varnhagen, 2005).

Overall, although strategy training has been shown to initially increase performance of working memory tasks, the benefits seem to only be short-term and not sustainable (Fellman et al., 2020). In addition to this, further research has also shown that some children naturally use strategy mechanisms, such as rehearsal, in order to enhance their performance even without instruction to do so (Peng & Fuchs, 2015). It has been suggested that improvement found to other areas of working memory following training, a near transfer effect, can reflect an improvement in the working memory task itself rather than an increase in working memory capacity (Shipstead et al., 2012). It has been recommended that training should not teach specific strategies in order for individuals to just be able to remember more information. A strategy may improve working memory scores on certain tasks but this is very different to altering underlying ability (Shipstead et al., 2012). Therefore implementing a working memory intervention without explicit strategy instruction and observing participants throughout training to see if any particular strategies evolve naturally would be beneficial to enhance understanding in this area of working memory training. This is, therefore, the approach taken in the current study.

The efficiency of working memory training, either with strategy instruction or without, is often determined by transfer effects, the benefits gained following an intervention to other areas of working memory (near transfer) or other cognitive skills (far transfer). The following section will consider how the methodology of previous working memory training studies may account for contradictory results concerning the presence of near and far transfer effects following interventions.

2.2 Methodological Critique of Previous Working Memory Training Research

According to Morrison and Chien (2011), domain-general training studies often and unsurprisingly produce significantly improved performance on the trained working memory tasks. Although 'near transfer' effects such as these show benefits on working memory tasks which have either been trained or untrained, the strongest evidence for the efficiency of working memory training interventions is from 'far transfer' effects, which demonstrate a gain in ability that is not directly related to the working memory task in structure or content (Henry et al., 2014).

A meta-analysis by Melby-Lervag et al. (2016) has looked at the methodological differences between working memory training studies in an attempt to account for contradictory results concerning gains relating to near and far transfer effects. The age of participants has been suggested as one difference across studies that may account for variations in findings, as younger participants may be more receptive to gains from working memory intervention due to increased brain plasticity.

The duration of administered training doses has also been highlighted as a possible reason for conflicting results, as someone that has undergone a longer duration of working memory intervention would be expected to show an increased benefit (Melby-Lervag et al., 2016). Jaeggi et al. (2008) found evidence of transfer from a demanding working memory task to fluid intelligence. The extent of gain in fluid intelligence was found to be dependent on the amount of training given, i.e fluid intelligence improved more with increased amounts of working memory training. Due to these promising findings, this study was replicated and extended by Thompson et al. (2013), who administered more training sessions in an effort to see more of a dosage effect and further gains. Thompson et al's (2013) findings however did not support Jaeggi et al.'s (2008) original results and although improvements were found in the trained working memory task, no significant improvements were evident in other areas of working memory capacity, fluid intelligence or cognitive ability measures.

Differences between study design types, in particular the inclusion of an active control group and the randomisation of participants, were also proposed as further reasons for conflicting findings. Control groups are vital to eliminate test-retest effects and other experimental confounds according to Shipstead et al. (2012). The cost of having a 'no contact' control group however, with engagement only at pre and post-intervention time points has also raised concerns (Shipstead et al., 2012). The 'Hawthorne Effect' (French, 1953) links to how individuals behave in relation to their involvement in an experiment (Shipstead et al., 2012). Conflicting outcomes may arise in working memory training effects as trained participants may see themselves as being more personally invested in improving outcomes compared to those who have experienced no contact. Melby-Lervag et al. (2016) even went as far to recommend that investigators stop conducting working memory training studies using untreated control groups. Non-adaptive versions of working memory tasks, where difficulty remains at a fixed low level, are therefore recommended for active control groups alongside the trained participants, in order to maintain active involvement in the working memory training intervention (Shipstead et al., 2012). Many of these issues have been addressed in the current study, including the inclusion of an active control group who participated in non-adaptive versions of the trained working memory tasks.

It has also been suggested that those with disabilities such as language disorder, dyslexia, attention deficit hyperactivity disorder, autism spectrum disorder or a generally lower cognitive ability, may have more room for improvement and therefore gain more from a working memory intervention compared to typically developing individuals (Melby-Lervag et al., 2016).

Inappropriate baseline measures, inappropriate pre and post testing, inconsistent testing time points, very small sample sizes resulting in inadequate statistical power, the need for mediator analysis in order to understand the relationship between far transfer effect improvements and working memory gains, and the use of untreated controls have all been critiqued in previous working memory training studies (Melby-Lervag et al., 2016). In the current study, many of these issues have been addressed and the study design shows an awareness of the methodological issues discussed throughout this section (see chapter 3, methodology chapter).

Alongside methodological critique, limited evidence of transfer effects to other areas of cognition following working memory training has led to pessimism concerning training effectiveness (Melby-Lervag & Hulme, 2013; Schwaighofer, Fischer & Buhner,2015; Sala & Gobet 2017; Melby-Lervag et al., 2016). The next section of this chapter will explore various meta-analysis studies of working memory training interventions, discussing the evidence of apparent near and far transfer effects.

2.2.1 Pessimism Regarding Previous Working Memory Training Studies

Whilst researchers acknowledge the benefits of working memory training for working memory related cognitive skills, i.e. near transfer effects, there is more negativity regarding the ability to achieve far transfer effects in other academic and cognitive skills (Sala & Gobet, 2017).

Melby-Lervag & Hulme (2013) conducted a meta-analysis of 23 different computerised based working memory training studies with both clinical and typically developing children and adult samples. Reliable short-term improvements were found in working memory skills, with verbal working memory showing an initial near transfer effect which was not sustained at follow-up. Limited evidence indicated gains in visuospatial effects, and these were maintained for a longer period and evident at follow-up (Melby-Lervag & Hulme, 2013). There was no evidence of a generalisation of working memory to any other skills: far transfer effects to non-verbal or verbal ability, inhibitory processes in attention, word decoding or arithmetic were not present (Melby-Lervag & Hulme, 2013). Overall, Melby-Lervag and Hulme (2013) expressed considerable doubt about the clinical relevance of working memory training interventions and their usefulness in improving cognitive functioning in typically developing children and adults.

Similar results were found in a further meta-analysis conducted by Schwaighofer et al., (2015) which included 47 working memory training studies. The aim of this meta-analysis was to examine near and far transfer effects following a working memory training intervention and to consider potential moderators. Near transfer effects were again apparent with improvements found in both short-term and working memory skills, all of which were maintained at follow-up (Schwaighofer et al., 2015). Far transfer effects however, were small and limited, with gains only found in non-verbal and verbal ability, which were not maintained at follow-up. Schwaighofer et al. (2015) reported no far transfer effects of working memory training to word decoding or mathematical abilities immediately post- intervention or at follow-up. Training dose, session duration, training supervision and location were all considered significant moderators to transfer effects (Schwaighofer et al., 2015). According to Schwaighofer et al. (2015) the fundamental purpose of working memory training is to tap into a domain-general attention capacity which can then enhance cognitive abilities. As no sustained far transfer effects were identified during this meta-analysis, Schwaighofer et al. (2015) stated that working memory training does not have any practical benefit for education and learning.

An additional meta-analysis has also considered the gains of working memory training in typically developing children aged 3-16 years (Sala & Gobet, 2017). Findings indicated a significant near transfer effect across studies in working memory and short-term memory measures, however little evidence was apparent for any far transfer effects (Sala & Gobet, 2017). In support of methodological critique discussed in section 2.2, Sala and Gobet (2017) considered that the level of effect size was related to the quality of the study design, i.e. having a better quality of design such as the inclusion of an active control group or implementation of random allocation between groups, resulted in a lower effect size. In agreement with the other meta-analyses discussed so far, Sala and Gobet (2017) stated that far transfer effects rarely occur in working memory intervention research and such effects are often minimal. As other researchers such as Shipstead et al. (2012) have argued, Sala and Gobet (2017) questioned whether near transfer effects represent just an improvement in working memory performance rather than a genuine enhancement in working memory capacity, hence the reason for a general lack in far transfer effects within working memory training research.

Following on from Melby-Lervag and Hulme's (2013) meta-analysis, a much larger review was later conducted looking at 87 publications with 145 experimental comparisons (Melby-Lervag et al., 2016). Although this later meta-analysis aimed to be as inclusive as possible, in relation to methodological critique and an awareness of study design, working memory training studies were only included in the analysis if the study involved computerised training, included pre and post-test measures, and a control group. Similar to previous findings, Melby-Lervag et al.'s (2016) meta-analysis found evidence of short-term improvements in both visuospatial and verbal working memory tasks, with visuospatial gains remaining evident at follow-up. Two areas showed weak evidence of far transfer effect, one from working memory to reading comprehension immediately post-intervention and one to arithmetic at follow-up (Melby-Lervag et al., 2016). However, as there was a decrease between the pre-test and post-test scores in the control group for both of these measures, Melby-Lervag et al. (2016) dismissed these findings and stated that there was once again no evidence of working memory training producing far transfer effects.

Although dismissed by Melby-Lervag et al. (2016) as a credible far transfer effect, variations in reading comprehension ability have been suggested to be predicted by working memory capacity and executive processing has been deemed valuable to comprehension abilities (Swanson et al., 2012). The following section will therefore explore previous research and although limited, will discuss studies that have found a positive link between verbal and visuospatial domain working memory training and comprehension. As traditional working memory studies have often been computer based, and with repeated evidence of limited transfer effects resulting from this method, this section will also discuss the need to adopt alternative approaches to working memory training interventions in future research.

2.2.2 The Special case of Far Transfer Effects to Reading Comprehension Ability

Although dismissed by Melby-Lervag et al. (2016) as a credible far transfer effect, variations in reading comprehension ability have been suggested to be predicted by working memory capacity and executive processing has been deemed valuable to comprehension abilities

(Swanson et al., 2006; McVay & Kane, 2012). Theories of discourse understanding initially presumed that the main aim of comprehension was the gradual formation of a textual representation in our episodic memory (Van Dijk, 1985). This has since evolved and now includes a model of what the text is about and the 'situation' of the text. Reading comprehension is seen as dependent on the reader's ability to construct a coherent meaning-based mental representation of a situation described in a piece of text, often referred to as a 'Situation Model' (De Koning & Van der Schoot, 2013). By not exclusively relying on words being read to extract meaning, and through using senses, particularly vision, to create non-verbal representations of the concepts within a piece of text, this will increase the likelihood of adequate understanding (De Koning & Van der Schoot, 2013).

According to De Koning and Van der Schoot (2013), achieving 'Situation Model' level of comprehension is a complex process and requires a reader to be able to move beyond word, phrase or sentence comprehension, which is only achievable once an individual becomes a proficient and confident reader. 'Situation Models' are steadily formed by continuously updating information regarding various text elements and integrating this information with the reader's background knowledge and understanding of the world. Through doing so this results in a 'coherent and richly connected visuospatial representation of the situation and events that are described in the text, enabling readers to draw interferences' (De Koning & Van der Schoot, 2013, p.262).

A 'Structure Building Model' devised by Gernsbacher et al., (1990) suggests that the modality of initial information (i.e. reading or listening comprehension) does not influence the creation of a 'Situation Model'. Therefore, comprehension has been suggested to be a domain-general skill which is not tied to the modality of the initial input. Both reading and listening comprehension are therefore seen as two versions of the same comprehension skill (Gernsbacher et al., 1990).

The current study, which targets both the verbal and visuospatial domain could, therefore, be beneficial because training visuospatial ability may have a direct impact on the ability to form ‘situation models’ and therefore improve comprehension ability.

Previous studies have shown a positive link between verbal and visuospatial domain working memory training and comprehension. A 5 week computerised working memory intervention that focused on verbal and visuospatial working memory training with 57 children aged 9-12 years with special educational needs found that skilled readers outperformed less skilled readers in all measures relating to working memory, processing speed and updating (Dahlin, 2010). Significant differences in working memory tasks amongst subgroups of less skilled readers, i.e. those with a comprehension-only deficit, reading deficit or poor readers, was also found to be due to variations in working memory skill (Dahlin, 2010). As working memory measures were found to relate to children’s word reading and reading comprehension ability, Dahlin (2010) suggests that working memory can be seen as an essential factor in reading development among children with special educational needs and, therefore, interventions to improve working memory may help children improve in reading comprehension ability.

Typically developing children have also been shown to improve in reading comprehension following an adaptive face to face working memory intervention that focused on executive-loaded working memory tasks in verbal and visuospatial domains (Henry et al., 2014). 36 children aged between five and eight years of age were randomly allocated to either a working memory intervention group or an active control group (Henry et al., 2014). Children received 18 working memory training sessions overall, which were conducted three times a week for a six week period. Immediately post-intervention, trained participants showed a significant difference in both trained tasks and near transfer effects were evident for word recall and counting recall. Far transfer effects were less evident, and were not present for reading accuracy or mathematics ability. However, the trained group of participants showed a significant difference in reading comprehension ability compared to the active controls at a 12 month follow-up (Henry et al., 2014). This suggests that improving ability to divide attention between processing and storage could have a specific benefit to reading comprehension (Henry et al., 2014).

Although the previously discussed meta-analytic studies have contributed to a rather pessimistic view regarding the effectiveness of working memory training, in particular for far transfer gains in other cognitive abilities and academic skills (Melby-Lervag & Hulme, 2013; Schwaighofer et al., 2015; Sala & Gobet, 2017; Melby-Lervag et al., 2016), the evidence of gains in comprehension ability, although limited, does bring some positivity (Dahin, 2010; Henry et al. 2014). For this reason, it cannot be concluded that working memory training interventions cannot ever produce positive near and far transfer effects. Melby-Lervag and Hulme (2013) note that although the current evidence is not encouraging, it is not impossible to rule out that future studies will demonstrate far-transfer effects through using different working memory training methods. Traditionally, working memory training has widely adopted a computer-based training approach (Rowe et al., 2019a). This approach has repeatedly evidenced gains on trained working memory tasks and untrained working memory tasks, which share common features to the trained tasks, but minimal and un-sustainable gains in far transfer effects to other areas (Rowe et al., 2019a).

Henry et al. (2014) adopted a face to face working memory intervention, which reported positive gains in reading comprehension for typically developing children at a 12 month follow-up, although some limitations in the study design limit the conclusions from this study.

In summary, as variations in reading comprehension have been suggested to be predicted by working memory capacity and as executive processing is deemed valuable to comprehension abilities, this section has explored reading comprehension as a far transfer effect to working memory training interventions. As previous studies, both with children with special educational needs (Dahlin, 2010) and typically developing children (Henry et al., 2014) have found a positive link between verbal and visuospatial domain working memory training and comprehension, it cannot be definitively claimed that working memory interventions do not produce far transfer effects. As traditional working memory training has often been conducted via computerised methods, exploring alternative approaches to training has been deemed pivotal (Melby-Lervag & Hulmes., 2013). The following section will therefore discuss the need for change in working memory intervention studies, specifically considering how adopting a different approach using non-computerised methods of training may enhance gains for both near and far transfer effects.

2.2.3 Non-Computerised Working Memory Training

A recent systematic review has assessed the effectiveness of non-computerised working memory interventions in children aged 4-11 years within their everyday environments (Rowe et al., 2019a). The authors included 18 studies consisting of a range of non-computerised working memory intervention approaches such as adapting the child's environment to reduce working memory loads, direct working memory training with and without strategy instruction, and training skills that may indirectly influence working memory ability (Rowe et al., 2019a).

In all effective non-computerised working memory intervention studies it was evident that the trained tasks were executive-loaded, meaning tasks which tap into attentional and processing resources rather than just storage (Rowe et al., 2019a). Evidence of near and far transfer effects from the working memory trained tasks was limited to just three studies. One was the study discussed in section 2.2.2: Henry et al. (2014) adopted a face to face direct working memory training intervention without strategy instruction, training tasks in both the verbal and visuospatial domains. Also discussed previously, significant improvements were found in both trained tasks, with near transfer effects to counting recall and word recall and evidence of a far transfer effect to reading comprehension.

A further non-computerised direct working memory training intervention, but with strategy instruction was conducted by Witt (2011). 38 children aged nine-ten years were part of a 6 week working memory intervention that focused on tasks relating to the central executive. Pre and post-intervention measures of backward digit recall, visual patterns and mathematical addition were collected. Participants were separated into two groups, trained and control, on a matched pair basis. Significant differences were found between groups at post-intervention in visual patterns and on the mathematical addition task (Witt, 2011). It has been suggested by Witt (2011) that the use of digit stimuli in the backward digit recall task, may have helped the trained group to improve on their addition performance.

Volckaert and Noel (2015) observed whether a cognitive intervention aimed at improving inhibition abilities would have an influence on executive functioning and in turn lead to behavioural changes and a decrease in external behaviour problems. Although not a direct working memory training intervention study, Volckaert and Noel's (2015) research is informative regarding the use of non-computerised methods in improving executive functioning. A sample group of 47 children with a mean age of 60 months undertook pre and post-tests in attention, motor and cognitive inhibition, flexibility and working memory. Behaviour was also assessed using questionnaires and observational tools (Volckaert & Noel, 2015). Children were randomly allocated to either a trained or control group with the trained group receiving two 45 minute inhibition training sessions per week for an eight week duration. Following the intervention, significant group differences were found in inhibition, attention and working memory measures in short-term memory, visuospatial short-term memory and the central executive (Volckaert & Noel, 2015). Inattention was also found to significantly change between groups, indicating that through the improvement of executive functioning, external behaviour could be impacted (Volchkaert & Noel, 2015).

Rowe et al. (2019a) noted that further non-computerised studies have also reported far transfer effects (Peng & Fuchs, 2015; Cornoldi et al., 2015; Passolunghi & Costa, 2016), however as no measures of near transfer effects were taken in any of these studies, this has weakened the evidence for far transfer to other areas of cognition.

Of the 18 non-computerised working memory intervention studies included in Rowe et al.'s (2019a) systematic review, only four included outcome measures at a follow-up period in order to assess transfer effect durability. Although the four studies all adopted different approaches in delivering working memory interventions, and their follow-up time points also differed in duration from the end of the working memory interventions, all studies demonstrated that gains in working memory were achieved and maintained, suggesting durability in transfer effects from non-computerised working memory training.

In summary, a range of approaches to training working memory via non-computerised methods, within children's everyday contexts, have been shown to produce improvements in working memory ability (Rowe et al., 2019a). Furthermore, these intervention studies have shown there is the potential to produce both near and far transfer effects via a non-computerised working memory training intervention. All effective interventions discussed by Rowe et al. (2019a) trained executive-loaded working memory, therefore, future studies need to consider this in their design. Methodological critiques of previous research, both computerised and non-computerised, also need to be addressed in future working memory intervention studies, with the inclusion of active control groups and randomised group allocation (Melby-Lervag et al., 2016). Finally, Rowe et al. (2019a) evaluated who would benefit from non–computerised working memory interventions and concluded that in future research children from understudied groups should be included in samples.

With this in mind, the current study conducted a non-computerised working memory intervention that focused on executive-loaded working memory tasks for children with language disorder, a subgroup of children who are understudied. The following section will explore working memory interventions for children with language disorder, who as the main focus of the current research, commonly display deficits in working memory ability (Farquharson & Franzluebbers, 2014).

2.3 Working Memory Training and Language Disorder

Although those with language disorder have been shown to display considerable difficulty with working memory ability, little research regarding executive-loaded working memory intervention for children with language disorder has been undertaken (Henry & Botting, 2017; Farquharson & Franzluebbers, 2014). Attempting to improve working memory ability in typically developing individuals has received extensive interest, as working memory is seen as fundamental in tasks such as reading comprehension, mathematics/arithmetic, planning of thoughts and actions and maintaining focused behaviour (Holmes, 2012). However, the importance of exploring working memory training interventions with subgroups of children who have been relatively understudied has been recently highlighted (Rowe et al., 2019a).

Although limited research regarding working memory interventions and those with language disorder is available, there are some promising findings emerging. After following a working memory training intervention using the 'Cogmed Working Memory Training' program, a group of 12 children, aged 8-11 years, with low language ability were found to have made significant gains in visuospatial short-term memory, verbal short-term memory and central executive (Holmes et al., 2015). The 'Cogmed Working Memory Training' program adopts a computerised training approach, where trainees undertake verbal and visuospatial tasks over a five to ten week period. The standard protocol for trainees aged seven years and over is to undertake eight training tasks, which in total last 35-45 minutes, three times a week for five consecutive weeks (Roche & Johnson, 2014). Holmes et al.'s (2015) findings following a 'Cogmed Working Memory Training' intervention seemed very promising for the language disorder group, however once corrections were made for multiple comparisons, all measures except visuospatial short-term memory became non-significant despite large effect sizes.

A further small scale study has also shown a good response to working memory intervention in a group of nine 7-9 year old children with language difficulties (Wener & Archibald, 2011).

Significant gains were present in the performance of the 'word structure' task of the Clinical Evaluation of Language Fundamentals 4 (CELF-4) (Semel et al., 2006) at a 4 month follow-up, showing a retained improvement in grammatical (word) structure (Wener & Archibald, 2011).

A recent study conducted with 32 children with language disorder and a chronological age matched control group explored the effectiveness of a working memory intervention for children with language disorder (Acosta et al., 2019). All children were given pre and post-intervention assessments in short-term verbal memory, verbal working memory, short-term visuospatial memory, visuospatial working memory, attention and processing speed and lexical-semantic processing ability. Pre-intervention assessment scores showed that those with language disorder performed significantly below age matched controls on all measures, indicating that children with language disorder have a neuropsychological profile characterised by severe memory difficulties, in both short and long-term memory, and across both verbal and visual domains (Acosta et al., 2019). Training sessions were conducted twice a week for a full academic year, with the first session each week conducted on an individual basis, and the second session in group of two to five children, one having language disorder and the rest being typically developing. Following the working memory intervention that targeted immediate verbal memory, visuospatial working memory and verbal working memory, children with language disorder showed significant improvements in all measured functions, and children in the language group made significant gains compared to the control group in all measures. This recent working memory intervention study, therefore, suggests that children with language disorder do respond to working memory training and that improvements can be made in short-term memory, working memory and language skills. What remains uncertain is whether or not there can be far transfer effects (section 7.3).

In summary, although working memory training with typically developing children has received extensive research, training children with language disorder, a group who show considerable difficulties in working memory ability, has received little attention. Although limited evidence is available, promising results show evidence that working memory training interventions can produce gains for those with language disorder, particularly in primary aged children. With this in mind, the current study recruited primary aged children aged 6-11 years who have completed their Foundation Stage curriculum and are now either in Key Stage 1 or 2. Through recruiting a slightly older child from 6 years of age and above, this also helps to ensure that any language difficulties experienced by the child cannot be purely due to their younger age.

The number of studies that have addressed working memory training and language disorder are also limited and often have very small sample sizes. Considering this, the current study uses a larger sample size of 47 participants (24 trained and 23 active controls) and will therefore increase power to detect near and far transfer effects.

The following section will give an overview of the areas discussed throughout this chapter. The ongoing debate between domain specific versus domain general working memory storage, and working memory training with and without strategy training will be summarised. The methodological critique of previous research, the pessimism regarding outcomes of traditional computer-based working memory interventions, and the need to alter training methods to non-computerised working memory interventions in children's real world contexts will also be overviewed. This section will conclude with a summary of how the current research attempts to address concerns raised in previous literature in terms of study design, training approaches and understudied target groups.

2.3.1 Moving forward with Working Memory Training – A Chapter Overview

Working memory has been repeatedly found to link to higher order cognition (chapter 1 section 1.2), and previous research studies have repeatedly attempted to train working memory in order to improve functionality and capacity. However, there remains considerable uncertainty in this area, including debates over whether information is maintained in the working memory system via a domain specific or domain general approach, whether training delivery is better with strategy instruction or without (section 2.1.2), and the nature of the training itself.

In response to conflicting outcomes following working memory interventions (section 2.2), the design of these studies has been carefully scrutinised. The need for future research to include appropriate baseline measurements, sufficient pre and post-intervention testing measures, consistent testing time points, adequate sample sizes, inclusion of active control groups and random allocation of participants has been argued as pivotal in successfully moving forward with working memory training research (Melby-Lervag et al., 2016).

Although the benefits of working memory training are commonly acknowledged for working memory related skills, i.e. near transfer, there still remains pessimism regarding gains in other academic and cognitive skills, i.e. far transfer (section 2.2.1). A number of meta-analyses have been conducted, mainly reporting relatively short-term improvements to working memory skills, with no or limited evidence of generalisation to other areas of learning (Melby-Lervag & Hulme, 2013; Schwaighofer et al., 2015; Sala & Gobet 2017; Melby-Lervag et al., 2016; section 2.2 & 2.1.1). Although evidence from these meta-analyses has not been encouraging, it has been suggested that it would be possible for future research to demonstrate far transfer effects via the implementation of different training methods (Melby-Lervag & Hulme, 2013).

Traditionally working memory research has involved computerised training programs, however after a general lack of transfer effect evidence (see above), it has been suggested the effectiveness of non-computerised working memory training interventions within children's everyday context should be considered (Rowe et al., 2019a). Furthermore, evidence, although relatively limited, suggests that a non-computerised, varied approach to working memory training within the child's everyday context, can lead to improvements in working memory ability and, furthermore, potential to produce near and far transfer effects (Rowe et al., 2019a).

Typically developing children have received extensive research attention in respect of gains and benefits of working memory interventions, yet moving forward it has been suggested that subgroups of understudied children should become the focus (Rowe et al., 2019a). Although children with language disorder have been shown repeatedly to display working memory difficulties, working memory interventions with this subgroup remain limited (Henry & Botting, 2017; Farquharson & Franzluebbers, 2014; section 2.3).

Emerging evidence that children with language disorder respond to working memory interventions, with some training gains, is a promising start, however studies in this field are limited and often have been conducted with very few participants (Acosta et al., 2019; Holmes et al., 2015; Wener & Archibald, 2011).

Given all these arguments, the current thesis aims to assess whether we can improve working memory and language task performance in children with language disorder.

Previous literature has shown the need to move away from computerised working memory training and employ an executive-loaded face to face intervention with a well-constructed design that satisfies the methodological critique discussed in this chapter (Rowe et al., 2019a, Melby-Lervag et al., 2016). Therefore, in the current thesis, a 6 week non-computerised working memory intervention is conducted targeting both verbal and visuospatial domains using executive-loaded working memory training tasks. In response to previous methodological critiques, all children are randomly allocated to either a trained or active control group. The need to understand understudied subgroups of children will be addressed through the focus on children with language disorder. Furthermore, as previous research has shown evidence of far transfer gains from executive-loaded working memory training to reading comprehension at a 12 month follow-up (Henry et al., 2014), the current study will focus on language outcome measures.

The current study will address several key research questions:

i) Whether children with language disorder respond to a face to face adaptive working memory intervention that targets verbal and visuospatial executive-loaded task domains. (see chapter 4)

ii) If so, whether training gains will transfer to other areas of working memory skill, i.e. do near transfer effects occur? (see chapter 6)

iii) More importantly, whether training gains also transfer to other areas of cognition and academic skills, specifically a range of language outcome measures, i.e. do far transfer effects occur? (see chapter 7)

Through addressing previous methodological limitations, focussing on non-computerised training, and including an understudied subgroup of children with language disorder, this thesis makes a novel contribution to the literature. The current study could also substantially impact our understanding of the relationship between working memory capacity, reading comprehension, language measures and the types of far transfer effects that may be most amenable to working memory interventions within a language disorder group.

Chapter 3 Methodology

3.1 Methodology Overview

This research project concerned a sample of 51 children with language difficulties as a primary need, aged six to eleven years, across six schools in Dorset and Hampshire. Each child was randomly allocated to either a working memory training group or an active control group by an independent colleague. Both groups of children, whether in the trained or active control condition, were visited by a researcher three times a week for six weeks, totalling 18 sessions, for approximately ten minutes at a time. During each session, children in the working memory intervention group were administered 11 trials of an Odd One Out span task (Henry, 2001) and 11 trials of a Listening Recall task (Gathercole & Pickering, 2001). All trials in both tasks were adaptively titrated to the child's ability levels and adjusted continuously to ensure that the task remained appropriately challenging in order to facilitate improvement. Children in the active control group were also administered these trials, however any working memory requirement was removed. The intervention study was designed to look at direct and near transfer effects (benefits in the directly trained areas and other related areas of working memory) and also far transfer effects (untrained areas such as reading accuracy, reading comprehension, sentence comprehension and receptive grammar). All participants were given a battery of assessments before the intervention process took place, immediately after and also at a 9 month follow-up to see if there were any near and far transfer effects from the working memory intervention and if so, how long these were maintained.

3.2 Participants

A total of 51 participants, aged between six to eleven years with language difficulty as a primary need, were recruited from six different schools across Dorset and Hampshire. The six schools involved in the research project were mainstream primary schools with four of these including a specialist language provision onsite. Before participant recruitment commenced, full ethical approval was sought and obtained. Schools were initially contacted via email with a brief introduction and overview of the research project. It was requested that if interested in taking part in the intervention study, the schools should contact the researcher personally to make a face to face appointment. Following this first contact, the schools who responded were visited; often the researcher would meet with the head teacher, special educational needs coordinator (SENCO) and the specialist language provision manager to discuss the study in more detail and the required involvement of both the staff and pupils. During these meetings the study inclusion criteria of participants were discussed: having a language difficulty as a primary need; a nonverbal IQ of 70 or above; no formal diagnosis of autism spectrum disorder; and English as a first spoken language at home. School professionals' highlighted particular children that were felt would be potentially suitable to participate in the intervention. Parent information and parental consent procedures were then explained and all parents of the children highlighted as potential participants were given a parent information sheet (see Appendix B) and parental consent form (see Appendix C) through the school. Once signed consent forms were returned to the school, the researcher was then informed of the total numbers for each setting, and arrangements were made to commence screening tests for each participant to ensure that inclusion criteria were met.

3.2.1 Inclusion Criteria

3.2.1.1 Language Difficulty as a Primary Need

Clinical Evaluation of Language Fundamentals 4 (CELF 4) (Semel, Wiig & Secord, 2006)

The CELF 4 was administered in the current study to determine whether language difficulty was a primary need during participant selection, as this test is well renowned and commonly used amongst speech and language professionals. The core language subtests 'Concepts and Following Instructions', 'Word Structure', 'Recalling Sentences' and 'Formulated Sentences' were administered to all participants. An additional subtest of 'Word Classes – Total' was also administered to those aged nine to eleven years as this is a requirement in order to calculate a core language score for children over eight years. In order to participate in this study children were required to attain a scaled score of seven (one SD below the mean) or below in at least two of the core language subtests. The subtests are outlined below.

The Concepts and Following Directions subtest was used to assess the child's ability to interpret, recall, and execute oral commands. The child was asked to listen to a set of instructions and to then point to the relevant pictures which were presented on a separate page varying amounts of rows. For example a child may have been asked to 'point to a fish next to a house' or 'point to all the shoes on the top row'. The commands increased in length and complexity and the task stopped once a child had scored seven consecutive scores of zero.

The Word Structure subtest of the CELF 4 assessed a child's knowledge of grammatical rules in a sentence completion task. Various grammatical rules were presented such as irregular plurals, possessive pronouns, future tenses and irregular past tenses. The child was asked to complete an orally presented sentence that was related to an illustration that was shown to the child such as 'Here is one book. Here are two _____'. There was no discontinuation point for this subtest and all children were administered all 32 sentences.

The Recalling Sentences subtest evaluated a child's ability to recall and produce sentences that varied in length and syntactic complexity. A child was asked to imitate the sentences that had been orally presented by the researcher and when doing so, any errors were recorded, which impacted the score attained. If a sentence was fully recalled with no errors, a score of three was given, one error gave a score of two, two to three errors gave a score of one and more than four errors during recall gave the child a score of zero. If a child obtained five consecutive scores of zero, the subtest was discontinued.

The Formulated Sentences subtest of the CELF 4 required a child to formulate semantically and grammatically correct sentences using a particular target word to describe a given picture. Depending on the grammatical and semantic correctness of a sentence, a child was given a score between zero and two for each item. If a child obtained five consecutive scores of zero the subtest was discontinued.

The 'Word Classes – Total' subtest was only administered to children aged 9 and over (as part of the core language score for this age group). The 'Total' is the sum of scores on Receptive and Expressive subtests. In the Receptive subtest, the child was presented with four words such as 'pillow', 'door', 'blanket' and 'lamp'. The child was then asked to identify a word pair from the four presented words, such as 'pillow' and 'blanket', and if identified correctly was awarded one point towards their receptive score. Following this the child was then asked how the two words from the chosen word pair are related, giving the child a point towards their expressive score if answered correctly. It was important to note that if the child gave an incorrect word pair during the receptive questioning, the child was asked the expressive part of the item using the word pair that the child had chosen. Each item of the subtest was administered to the child, unless a child obtained five consecutive scores of zero and then the subtest was discontinued. The final 'Word Classes – Total' raw score was calculated through the sum of both the receptive and expressive scores. Once a raw score had been calculated, the researcher was then able to produce a 'Word Classes – Total' scaled score which contributed to the older child's core language profile.

3.2.1.2 Non Verbal IQ above 70

Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI –II) (Wechsler, 2011)

The block design and matrix reasoning subtests of the WASI – II were administered to check if the child had a nonverbal IQ score above 70. Through administering both the block design and matrix reasoning subtests of the WASI-II a perceptual reasoning index (PRI) score could be calculated for each child. If the PRI score was at 70 or above the child met the inclusion criteria for this aspect of the screening tasks.

The block design subtest of the WASI-II required a child to analyse and re-produce abstract visual stimuli. There were a total of 13 items, the first four required the design to be re-produced from a picture stimulus and a previously constructed model, which the child observed being made by the researcher, and the remaining items from only a picture stimulus. Depending on completion time, the child was allocated a score from zero to six and if two consecutive scores of zero were recorded, the subtest was discontinued.

The matrix reasoning subtest required a child to view an incomplete series of matrices and choose one from five options to complete the set. Children were scored either one point for a correct answer or zero for an incorrect answer and the subtest was discontinued after three consecutive scores of zero.

3.2.1.3 No Formal Diagnosis of Autism Spectrum Disorder (ASD)

Any history of ASD was discussed with the child's school and if a formal diagnosis had previously been made, the child was not included in this research project. This was to ensure that any language difficulty identified could not be equated to the child having a diagnosis of ASD.

3.2.1.4 English as the First Spoken Language at Home

A child's first language spoken at home and school was established with each school. When selecting children suitable for this research project, schools were asked to only put forward children who spoke English as a first language, therefore any language difficulty identified could not be due to the child speaking English as an additional language.

3.2.2 Random Allocation of Children

Once recruitment and screening tests had taken place, children who fulfilled the study inclusion criteria were randomly assigned by an independent colleague to one of two groups, a working memory intervention group (n=24) or an active control group (n=23). Excluding the researcher, all other individuals including staff at the schools were blind to this random allocation. The researcher however was aware of the participant's group membership as unfortunately for a PhD with one researcher, it would have been impossible to conduct a double blind experiment.

3.2.3 Screening Data

51 children were recruited in the current study from six primary schools. 47 of these children fulfilled the inclusion criteria, with four being excluded after screening tests were administrated. This was due to perceptual reasoning index scores falling below 70 on the WASI II (Wechsler, 2011) for two children, therefore not meeting the nonverbal ability criteria for the study, one child was undergoing an ASD assessment at the time of the intervention and one child did not speak English as a first language.

For the remaining 47 participants, once random group allocation and pre-intervention testing had taken place, independent t-tests or Mann Whitney U tests (depending if data was parametric or non-parametric) were conducted to ensure that there were no significant differences between working memory ability or language outcome measures between the two groups prior to the intervention commencing. For a detailed overview, please refer to chapter 4, section 4.2.1 (directly trained tasks), chapter 6, section 6.2.1 (near transfer tasks) and chapter 7, 7.2.1 (far transfer tasks).

3.3 Working Memory Training Intervention

Those assigned into the working memory intervention group undertook the full intervention protocol over a six week period. Children were visited three times a week for six weeks and spent approximately 10 minutes with the researcher at each visit. In each of the 18 sessions, the children were administered 11 trials of the Odd One Out Span task (Henry, 2001), and 11 trials of the Listening Recall task (Gathercole & Pickering, 2001).

Those assigned to the active control group also had 18 sessions with the researcher, and received the same materials and activities except that all working memory aspects of the tasks were removed as detailed below. One of the main limitations of t that it was not a double blind experiment and the researcher was aware of the participant's group membership. Unfortunately for a PhD with one researcher, it would have been impossible to conduct a double blind experiment.

3.3.1 Procedure

All children were visited three times a week within their school environment at the same time each week. Upon arrival, the researcher would go to the allocated place within the school grounds set aside for the working memory intervention to take place. This was normally a quiet area either in a separate classroom, library area or intervention room which the school staff had booked for use each time. Each child was collected from their classroom during each visit and spent 10 minutes with the researcher being administered 11 trials of both the Odd One Out task and the Listening Recall task. Children took great delight in remembering their previous memory span levels and trying to improve their scores, so the beginning of each intervention session commonly evolved around this and talking to the child about their week. Record forms were firstly completely, looking at the session number, previous span levels on each task and ensuring that the task was administered interchangeably so that the child received a different training task first each time. The researcher would then check that the child was aware of the task instructions and would then administer 11 trials of each task separately as explained below. Following the training session being implemented, the researcher would check if the child had any questions and then would let the child know when their next training session would be. The child was then returned to their classroom and the next child was located. The same procedure occurred for the active control group, however the working memory element was removed from the task. The children within the active control still seemed to enjoy the sessions however and were happy to attend each time.

3.3.2 Working Memory Training Materials

3.3.2.1 Odd One Out Span Task (Henry, 2001)

The Odd One Out Span Task (Henry, 2001) assessed a child's executive-loaded working memory in the nonverbal, visuospatial domain. The task required a child to identify one pictorial item (these are abstract and not easily nameable) out of 3 presented on a strip of card that was the 'odd one out'. The child was then required to remember the spatial location of the 'odd one out' item, and when presented with a subsequent blank strip of card, recall its previous location. The task increased in difficulty as the child's memory span increased, i.e., requiring the child to make a greater number of odd one out judgements and remember an increased number of odd one out spatial locations. Dependent on the child's own memory span level, the child could be asked to make a minimum of 1 odd one out judgement per trial or increase to a maximum level of difficulty, determined by the child's own ability to recall odd one out spatial locations. The maximum level of difficulty therefore, was completely determined by each child's own memory span level.

The training materials required for this task were the 'odd one out' card strips (see Appendix D), which are 3 pictorial items (one different from the other two identical pictures) presented in a horizontal line on a strip of card, the odd one out response sheets for the child to recall the spatial locations (see Appendix E), and a working memory training record form (see Appendix F) for the researcher to record each trial span level.

Figure 3.1 Example of 'Odd One Out' Card Strip and Response Sheet

The very first training trial begun at the child's span level, established from a pre-intervention test previously administered (see below for full details). The researcher started at this span level, unless the span level was previously zero, in which case the researcher started at span level one. Holding the odd one out strip cards (but concealing them from the child's view) the researcher placed the first card in front of the child asking them to point to the 'odd one out'. Depending on the start span level the researcher then presented the requisite number of cards required in sequence (e.g., for span levels of higher than one, two cards or more as appropriate). After the requisite number of odd one out cards had been presented, the researcher placed the empty grid response sheet in front of the child and asked them to recall the spatial locations of the 'odd one out' cards just shown. At the beginning of the very first training session, the span level established from a pre-intervention test previously administered was administered for the first two trials. If both trials had been answered correctly, the researcher moved the child up one span level, thus, increasing the task difficulty by one item. If both trials, however, had been answered incorrectly the researcher lowered the span level by one (or if already at one would continue with this level). If the child had answered one trial correctly and one trial incorrectly, the span level remained the same for a further two trials. This was then reassessed, continuing in this format until the child had answered 11 trials. At the end of the 11 trials the researcher calculated the most commonly administered span level and recorded this on the record form so that this would be the 'new' start level at the next session. The process was then repeated for the next 17 training sessions of the working memory intervention.

Children who had been assigned to the active control group were also administered 11 Odd One Out trials, but made the odd one out judgment only, i.e. they were not required to recall the spatial locations. For the active control, therefore, the same materials were used and the same interaction was included with the researcher as for the intervention, but removing any working memory requirement to the task.

3.3.2.2 Listening Recall Task

The Listening Recall task required the child to listen to a sentence that the researcher orally presented, saying whether the sentence was true or false, and then recalling the final word of the sentence. The task increased in difficulty as the child's memory span increased requiring the child to remember final words of an increased number of sentences that had been orally presented by the researcher.

The training materials required for this task were the 'listening recall sentences' (see Appendix G) and a working memory training record form (see Appendix F) for the researcher to record each trial span level.

The very first training trial began at the child's span level, established from the pre-intervention test previously administered as part of The Working Memory Battery (WMTB-C) (Gathercole & Pickering, 2001), see section 3.4.1.2 below for details. The researcher started at this span level; if the span level was previously zero, the researcher started at span level one. The researcher orally presented the first sentence and depending on the start span level the researcher presented the required number of sentences. After the requisite number of sentences had been presented and judged as true or false by the child, the researcher asked the child to recall the final word from each sentence. Two trials at the initial span level of the child were presented. If both trials were answered correctly the researcher moved the child up one span level, increasing the task difficulty. If both trials were answered incorrectly the researcher lowered the span level by one (or if already at one continued with this level). If the child answered one trial correctly and one trial incorrectly the span level remained the same for a further two trials. This then continued in this format until the child answered 11 trials. At the end of the 11 trials the researcher calculated the most commonly administered span level and recorded this on the record form so that this would be the start level for the next training session. The process was repeated for the next 17 training sessions of the working memory intervention. The children assigned to the active control group were also administered 11 trials per session but were not required to recall the final word of the sentence. Thus, the active control group just needed to identify whether the sentence was true or false, removing any working memory requirement to the task.

3.4 Outcome Measures, Pre-Intervention, Post-Intervention and 9 month Follow-Up

All children in both the working memory intervention and active control group were administered a variety of assessments before the working memory intervention commenced, immediately after the 18 intervention sessions were completed and at a 9 month follow-up. These tests were used to evaluate whether there were any direct or near transfer effects due to the intervention, with benefits to the trained areas of working memory and also other areas of working memory, and any far transfer effects relating to other aspects of the child's learning such as reading accuracy, reading comprehension, sentence comprehension and receptive grammar. See table 3.1 below for an overview.

Table 3.1 Screening tests, intervention methods and pre-intervention, post-intervention, and 9 month follow-up tests administered to each group

Groups	Screening Tests	Pre, Post and 9-month follow-up tests	Intervention
Working memory training group:	• CELF 4: Core Language subtests. • WASI: Matrix reasoning and block design.	• WMTB-C: digit recall, word list recall, block recall, counting recall, backward digit recall, listening recall. • Odd One Out Task • Pattern Span • ACE – Sentence Comprehension subtest • TROG 2 • NARA	18 working memory training sessions of Odd One Out Task and Listening Recall
Active control group:	Both groups had the same set of assessments.	Both groups had the same set of assessments.	18 sessions of Odd One Out Task and Listening Recall (working memory element removed from the task)

3.4.1 Outcome Measures (Direct Transfer)

3.4.1.1 Odd One Out Span Task (Henry, 2001)

The Odd One Out Span task (Henry, 2001) assessed a child's executive-loaded working memory in the nonverbal, visuospatial domain. The principle behind this outcome measure was exactly the same as for the training materials described within the working memory intervention, however this task was administered on a tablet using a PowerPoint presentation. A different set of stimuli to the working memory training materials was used to eliminate any practice effects. The PowerPoint presentation begun with the instructions of the task, which the researcher read through with the child. Once the child was happy to start the task the researcher presented practice trials for the child to become familiar with the format and how to respond to each trial using the tablet. Each trial showed 3 pictures on the screen of which the child deciphered which was the 'odd one out'. The child was still required to select the 'odd one out' by pointing to the chosen picture on the screen. They then remembered the spatial location of the shape and recalled it by pointing to the correct location (left, middle or right) on the next slide which displayed an empty grid. As with the training materials, the child was presented with six trials per span length, starting with span level one. If the child recalled four or more spatial locations from the six trials correctly, the researcher moved the child up to the next span level. If however, the child did not recall this many spatial locations, the task was discontinued.

3.4.1.2 Listening Recall, Working Memory Test Battery (WMTB - C) (Gathercole & Pickering, 2001)

This task within the WMTB-C was administered in exactly the same way as the training intervention task (see section 3.3.2.2) however the sentences used were different from those used in the working memory intervention.

3.4.2 Outcome Measures: (Near Transfer)

3.4.2.1 Working Memory Test Battery (WMTB - C) (Gathercole & Pickering, 2001)

- ***Digit Recall***

This task involved the experimenter reading out a list of numbers which the child repeated back in the correct order. The experimenter started with span level one and presented six trials at this span length. If the child completed four or more out of six trials correctly the experimenter moved up to the next span level, now presenting two numbers at a time. If after this set of six trials, the child completed four or more out of six correctly, the experimenter again moved up a span level. If the child scored less than this, the researcher stopped the task and recorded the span level last administered.

- ***Word List Recall***

This task involved the experimenter reading out a list of one syllable words which the child repeated back in the correct order. The experimenter started with span level one and read out six trials at this span length. If the child completed four or more out of the six trials correctly the experimenter moved up to the next span level, now reading out two words at a time. If after this set of six trials, the child completed four or more out of six correctly, the experimenter again moved up a span level. If the child scored less than this, the researcher stopped the task and recorded the span level last administered.

- ***Block Recall***

A block board with nine randomly spread out identical raised cubes (all 3cm x 3cm coloured grey – with identifying digits one –nine on one side of each cube) was placed between the child and the experimenter ensuring that the numbers on the cubes faced the researcher and were not visible to the child. Following the number sequences detailed on the record form, the researcher tapped the spatial positions on the blocks and asked the child to recall the positions by tapping the same sequence immediately afterwards. The spatial sequence started with a span of one and after six trials at this level, if the child completed four or more out of six correctly the span level then increased to two. Six further trials are then administered at this level and if the child again completed four or more trials out of the six administered at span level two, the level increase. If not, the task was discontinued and the last administered span level was recorded.

- ***Counting Recall***

This task required the child to count, out loud, the number of dots that were presented to them in a stimulus book. After counting the dots, the researcher turned the page of the stimulus book and the child was then asked to recall the number of dots they previously counted. Trials started with a single page of dots and were given in list lengths of six. At the end of each list length, if a child scored four or more correctly out of the six presented trials, the researcher then moved the child up a level so that they now needed to recall two pages of dots per trial. The task discontinued once the child reached their span level and scored less than four on any given list length.

- ***Backward Digit Recall***

This task involved the experimenter reading out a list of numbers which the child had to repeat backwards. The experimenter started with span level two, reading out two numbers per trial, with six trials at this span length. If the child recalled four or more out of six correctly the experimenter moved up to the next span level, now reading out three numbers at a time. If after this set of six trials, the child recalled four or more out of six correctly, the experimenter again moved up a span level. If the child scored less than this, the researcher stopped the task and recorded the span level last administered.

3.4.2.2 Pattern Span: (based on that of Della Sala et al., 1997)

Pattern span was a task that required the child to identify the position of a red square on a 2x2 grid, and then recall that position on an identical but black and white grid. This task again had six trials per span length and increased in difficulty as the task progressed. The first set of trials required the child to recall one red square on a 2x2 grid and as the span level increased, the size of the grid and number of red squares increased, until the last set of trials required the child to recall seven red squares across a 4x5 grid.

The researcher discontinued the task if the child did not score four or more correct answers for any set of six trials. This then indicated the child's span level for this task and this was recorded on the record form.

3.4.3 Outcome Measures (Far Transfer)

3.4.3.1 Assessment of Comprehension and Expression 6-11 (ACE) (Adams et al., 2001)

Sentence Comprehension Subtest

This subtest of the ACE was administered to all children to evaluate their ability to retain a sentence and be able to process both the words and the concepts within it. The child was shown four pictures on a page which were all similar, but not identical. Each picture had a slight variation to the theme relating to a sentence which was read aloud to the child such as 'The helicopter flew above the clouds'. After hearing the sentence, the child was then asked to pick which picture fit best. The subtest had 35 items which were all administered to the child and there was no discontinuation point for this test.

3.4.3.2 Test of Receptive Grammar Version 2: (TROG 2) (Bishop, 2003)

The TROG-2 measured a child's understanding of grammatical constructs. The test covered 20 different constructs, each represented by a block of four items, totalling 80 items. For each item, the child was shown four pictures on one page. One of these pictures represented a target sentence that was read to the child whereas the other three pictures acted as foils that depicted a sentence that was altered by a grammatical or lexical element. The child's responses were recorded by the researcher and a discontinuation rule applied if the child answered five consecutive items incorrectly.

3.4.3.3 Neale Analysis of Reading Ability – Second Revised British Edition (NARA II) (Neale, 1997)

The NARA II measured a child's reading accuracy, comprehension and reading rate. The child was asked to read a practice passage relevant to their age group aloud to the researcher.

During the practice passage the researcher timed the reading duration with a stopwatch and recorded the errors made placing them into error categories listed on the NARA II record form. Following the practice passage the researcher decided on a test starting point for the child. The general rule is that children aged seven and under began at level one, those aged eight to nine years began at level two and those aged ten + began at level three, but this was assessed using judgement and from evaluating the child's reading ability in the practice passages. Once a starting level was decided upon, the researcher administered the test passages. The child was asked to read each passage, again with the researcher timing the reading duration and recording the errors in error categories. Passages increased in difficulty and the test was discontinued if a child makes more than 16 errors (20 errors for level six) whilst reading the passage. If a child made fewer than 16 errors, the researcher administered the comprehension questions that related to the passage. Researchers decided to discontinue the test if the child made 12-16 errors and it was felt that they would find it difficult to master the next level of difficulty. Thus, the researcher made an informed judgement as to whether the child would be able to read the next passage of text with meaning, or whether the number of errors had reached a level where this was not achievable. Once the test had been discontinued the researcher finished off the session with a positive interaction with the child, asking questions relating to their preferred passage.

Chapter 4 Direct Transfer

The Effects of the Intervention on Targeted Working Memory Abilities.

4.1 Introduction

It is generally accepted that working memory training interventions produce gains in directly trained working memory tasks (Melby-Lervag et al., 2016; chapter 2). A meta-analysis of 87 publications in this field questioned whether working memory training was effective in improving the performance on measurements that were either identical or closely related to the working memory trained tasks administered (Melby-Lervag et al., 2016). Large effects were found for trained task improvement and these gains were also maintained at a 5.5-month follow-up.

Gains from a directly trained task have been widely reported in studies which have targeted working memory in various groups of individuals. As discussed in previous chapters, in a similar intervention to the current study, typically developing children showed an immediate direct transfer effect following a 6-week working memory intervention on two executive-loaded working memory tasks, one in the visuospatial and one in the verbal domain. Furthermore, these direct gains were maintained at a 6-month follow-up (Henry et al., 2014).

Further studies have also found direct transfer gains in working memory intervention studies with children with developmental differences. In a meta-analysis concerning ten different studies, the effectiveness of working memory training with individuals with learning difficulties and Down Syndrome was examined (Danielsson et al., 2015). Danielsson et al. (2015) suggested that in order to gain significant training effects for this group, a mixed working memory training approach was required which used both a verbal and visuospatial component.

One particular group who have cumulative evidence to suggest difficulties with working memory tasks are those with language difficulties (Farquharson & Franzluebbers, 2014; chapter 1). Children with language disorder commonly demonstrate not only a marked difficulty in expressive and/or receptive language but also reductions in working memory ability, even when they have a normal hearing range and normal nonverbal intelligence (Montgomery et al., 2010). Impairments in verbal measures of executive-loaded working memory in children with language disorder are commonly found, with particular difficulties present in sentence and listening span tasks which draw on concurrent processing and storage skills in the verbal domain (Henry et al., 2012). In contrast, executive-loaded working memory within the non-verbal domain has mixed evidence, however visuospatial executive-loaded working memory performance has been shown to be weaker for children with language disorder (Henry et al., 2012, Im-Bolter et al., 2006 & Marton, 2008).

Previous research in language disorder and working memory training is extremely limited. As pointed out above, it is well documented that children with language disorder have difficulties in all areas of working memory to at least some extent (Blom & Boerma, 2020; Vugs et al., 2013). Referring back to Danielsson et al.'s (2015) meta-analysis with children with developmental differences, it was found that in order to gain significant training effects, a mixed working memory training approach that used both verbal and visuospatial components was necessary. Furthermore, the importance of using executive-loaded working memory trained tasks within an intervention, i.e. those that tap into both attentional and processing resources rather than just storage, has also been recently discussed (Rowe et al., 2019a). By investigating the effectiveness of non-computerised interventions that target working memory in children aged 4-11 years within their everyday contexts, Rowe et al. (2019a) found all effective studies included executive-loaded working memory trained tasks. Therefore, by training Listening Recall and Odd One Out span tasks, the current study focuses on two executive-loaded working memory trained tasks within the verbal and visuospatial domains, two components suggested to be contributing factor in effective working memory training interventions.

This chapter will address the following research questions:

i) Do children with language disorder show immediate post-intervention improvements in two areas of executive-loaded working memory after direct training in these tasks?
ii) Do children with language disorder maintain improvements in two areas of executive-loaded working memory after direct training in these tasks, for a period of 9 months?

As Henry et al.'s (2014) study with typically developing children found an immediate training effect and also maintenance of gains at a six month follow-up on two executive-loaded working memory tasks, and as it is also generally accepted that working memory training interventions produce reliable gains in trained working memory tasks which are commonly found to be maintained at a 5.5 month follow-up (Melby-Lervag et al., 2016), it is hypothesised that children in this sample group will benefit from the working memory intervention implemented and will produce gains in both direct trained tasks immediately post-intervention and at the 9 month follow-up period.

4.2 Methods

4.2.1 Participants

As outlined in section 3.2, 51 children aged 6-11 years with language difficulty as a primary need were recruited for this intervention study. The same participants and procedures were adopted as described in chapter 3. Please see section 3.2.1 for inclusion criteria, section 3.2.2 for random allocation procedures and section 3.2.3 for participant screening procedures.

As outlined in section 3.2.2, once all children had been recruited and all parental consents given, children who fulfilled the study's inclusion criteria were randomly allocated to either a working memory intervention group or an active control group. Once randomly allocated, each group's pre-intervention scores for both trained tasks, Listening Recall and Odd One Out were inspected for normality. To investigate whether there were significant group differences on either directly trained working memory task, independent t-tests or Mann-Whitney U tests (depending on whether data were parametric or non-parametric) were conducted between trained and active control children at pre-intervention time point (see table 4.1 below).

Table 4.1 Mean raw scores, standard deviation, normal distribution, independent t-test or Mann-Whitney U results for both Listening Recall and Odd One Out trained tasks at pre-intervention

	Mean	***SD***	***Normal Distribution***	***T-test Result***	***Mann-Whitney U***
Listening Recall					
Trained Group (n=24)	9.21	3.62	*w*(24)=.97, *p*=.58	*t*(45)= -.23, *p*=.82	
Active Control (n=23)	9.52	5.53	*w*(23)=.94, *p*=.20		
Odd One Out					
Trained Group (n=24)	9.08	4.10	*w*(24)=.86, *p*=.**003**		*u*=259.50, *p*=.72
Active Control (n=23)	8.78	3.25	*w*(23)=.96, *p*=.49		

4.2.2 Measures and Procedures

Measure and procedures of the current study are outlined in section 3.2.2. For detailed procedural information of both the Listening Recall and Odd One Out trained tasks, please refer to section 3.3.1, with section 3.3.2 detailing the training materials and procedures required to administer both working memory trained tasks.

4.2.3 Statistical Analysis

Hierarchical multiple regressions were conducted at two time points, one immediately post-working memory intervention and one at a 9 month follow-up. The regression analyses were used to detect whether group, trained vs active control, was a significant predictor of performance for each trained task. As two children within the sample did not score highly enough to attain a standardised score for the Listening Recall task, and as standardised scores were not available for the Odd One Out task, raw scores were used with age controlled. For descriptive statistics of standardised Listening Recall scores at pre-intervention time point, please see Appendix H. There were no significant group differences between the two dependent variables at pre-intervention (see Table 4.1), so if the intervention was effective, group would be a significant predictor of Listening Recall and Odd One Out scores at post-intervention and 9 month follow-up.

Hierarchical multiple regression was chosen as a method of statistical analysis so that the influence of pre-intervention performance could be taken account of in the analysis. As pointed out above, raw scores were used in all analyses of trained tasks. However, as these scores were uncorrected for age, in each regression age was used as an additional 'control' predictor. Therefore, in the regression analyses the pre-intervention score of the dependent variable was entered at Step 1 (e.g. Listening Recall pre-intervention score) alongside age in months at the pre-intervention time point, and at step 2 group was entered as a dummy variable.

The regression analyses were carried out separately for the two dependent variables, and separate sets of regression analyses were conducted to analyse the effects of the intervention at post-intervention and follow-up. This meant that four separate regression analyses were conducted. Bonferroni corrections were made to the significance levels from these analyses with a *p*-value of 0.01 required to identify a significant effect (0.05 ÷4). For all regression analyses, key statistical checks were carried out (Durbin-Watson, tolerance and VIF statistics, Cook's and Mahalanobis distances, standardised DFBeta, leverage values, plots of standardised residuals, predicted standardised values, standardised residuals and partial plots).

4.3 Results

Results will be reported in separate sections relating to each trained task, Listening Recall and Odd One Out. Each section contains descriptive statistics (mean, standard deviation, a range at all three time points, and distribution of changes in scores across time points), and hierarchical regression analysis in relation to direct trained task effects at immediate post-intervention and at a 9 month follow-up.

4.3.1 Listening Recall

4.3.1.1 Listening Recall, Descriptive Statistics

Table 4.2 shows the descriptive statistics for raw scores, calculated from trials answered correctly, for the Listening Recall task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at a 9 month follow-up.

Table 4.2 Mean, standard deviation and range of Listening Recall raw scores for both trained and active control groups

		Mean	*SD*	*Range*
Listening Recall	Pre (n=24)	9.21	3.62	1-17
Trained Group	Post (n=24)	16.33	3.77	8-25
	Follow Up (n=24)	16.38	4.41	8-28
Listening Recall	Pre (n=23)	9.52	5.53	0-19
Active Control	Post (n=23)	10.00	5.71	0-19
	Follow Up (n=23)	9.78	5.13	1-18

Examination of Table 4.2 shows that the trained group of children had a mean Listening Recall score increase from pre-intervention to 9 month follow-up of 7.17 compared to a mean increase of just 0.26 in the active control group.

The percentage of improvements, decreases or no change in Listening Recall scores for both the trained group and active control group between pre-intervention and 9 month follow-up, pre-intervention and post-intervention, and post-intervention and 9 month follow-up time points, are shown below in table 4.3.

Table 4.3 Percentage of participants who made improvement, a decrease or no change in Listening Recall raw scores across all time points

		Pre-Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	100%	100%	41.7%
	Decreased	0%	0%	41.7%
	No Change	0%	0%	16.6%
Active Control (n=23)	**Improved**	47.8%	52.2%	43.5%
	Decreased	30.4%	26%	52.2%
	No Change	21.7%	21.7%	4.3%

Examination of table 4.3. shows that all trained children improved in Listening Recall scores between pre-intervention and the 9 month follow-up time point compared to 47.8% of active controls. All trained children also improved their Listening Recall scores between pre and post-intervention in comparison to 52.2% of the active controls. Further improvement between post-intervention and 9 month follow-up was relatively similar, with only a 1.8% difference between both trained and active control groups.

4.3.1.2 Listening Recall (Post-Intervention), Hierarchical Regression Analysis

As shown in Table 4.4 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention Listening Recall scores after controlling for pre-intervention score and age. Step one of the model involved entering pre-intervention scores and age in months as 'control' variables. These two variables accounted for 44.3% of the variance in post-intervention Listening

Recall scores, however, examination of the standardised beta-values indicated that only the pre-intervention score of Listening Recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 33.8% of the variance in post-intervention Listening Recall scores. The final model at step two was significant ($F(3,43) = 51.06$, $p<.001$) and predicted 78.1% of the Listening Recall score variance (Adj. R^2 .766).

Table 4.4 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, Listening Recall

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	6.62 (-2.28-15.51)	4.41		p=.14
Pre-Intervention	.85 (.51-1.19)	.17	.69	p<.001
Age	-.01 (-.12-.09)	.05	-.04	p=.79
Step 2				
Constant	16.72 (10.55-22.90)	3.06		p<.001
Pre-Intervention	.89 (.67-1.11)	.11	.71	p<.001
Age	-.02 (-.09-.05)	.03	-.05	p=.54
Group	-6.61 (-8.25 - -4.98)	.81	-.58	p<.001

Step 1 R^2=.44, p=<.001 **Step 2 ΔR^2= .34, p<.001**

4.3.1.3 Listening Recall (9 Month Follow-Up), Hierarchical Regression Analysis

As shown in Table 4.5 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9-month follow-up Listening Recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age in months accounted for 44.9% of the variance in 9 month follow-up Listening Recall scores, however, only the pre-intervention score of Listening

Recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing this variable to be a significant predictor accounting for a further 35.8% of the variance in 9 month follow-up Listening Recall scores. The final model ($F(3,43) = 59.89$, $p<.001$) at step two predicted 80.7% of the Listening Recall score variance (Adj. R^2 .793).

Table 4.5 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, Listening Recall

Pre- Intervention – 9 Month Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	3.23 (-5.67-12.13)	4.42		p=.47
Pre-Intervention	.79 (.51-1.19)	.17	.63	p<.001
Age	.03 (-.12-.09)	.05	.07	p=.62
Step 2				
Constant	13.70 (7.87-19.53)	2.90		p<.001
Pre-Intervention	.83 (.62-1.03)	.10	.66	p<.001
Age	.02 (-.04-.08)	.03	.05	p=.53
Group	-6.85 (-8.40 - -5.31)	.77	-.60	p<.001

Step 1 R^2=.45, p=<.001 **Step 2 ΔR^2= .36, p<.001**

4.3.2 Odd One Out

4.3.2.1 Odd One Out, Descriptive Statistics

Table 4.6 below shows the descriptive statistics for the raw scores, calculated from trials answered correctly, on the Odd One Out task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at 9 month follow-up.

Table 4.6 Mean, standard deviation and range of Odd One Out raw scores for both trained and active control groups

		Mean	*SD*	*Range*
Odd One Out	Pre (n=24)	9.08	4.10	4-23
Trained Group	Post (n=24)	18.46	6.16	9-33
	Follow Up (n=24)	18.75	5.98	7-31
Odd One Out	Pre (n=23)	8.78	3.25	1-14
Active Control	Post (n=23)	8.52	3.37	1-16
	Follow Up (n=23)	9.48	2.98	2-15

Examination of Table 4.6 shows that the trained group of children had a mean Odd One Out score increase of 9.38 from pre-intervention to 9 month follow-up compared to a mean increase of just 0.7 in the active control group.

The percentage of improvements, decreases or no change in Odd One Out scores for both the trained group (n=24) and active control group (n=23) between pre-intervention and

9-month follow-up, pre-intervention and post-intervention, and post-intervention and 9 month follow-up time points, are shown below in table 4.7.

Table 4.7 Percentage of participants who made improvement, a decrease or no change in Odd One Out Scores across all time points

		Pre to Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	100%	100%	50%
	Decreased	0%	0%	41.7%
	No Change	0%	0%	8.3%
Active Control (n=23)	**Improved**	56.5%	26.1%	73.9%
	Decreased	21.7%	47.8%	17.4%
	No Change	21.7%	26.1%	8.7%

Examination of Table 4.7 shows that all trained children improved in Odd One Out scores from pre-intervention to 9 month follow-up compared to 56.5% of active controls. All trained children improved their Odd One Out scores between pre and post-intervention in comparison to 26.1% of the active controls. Improvement between post-intervention and 9 month follow-up was higher for active controls with 73.9%, compared to 50% of trained children.

4.3.2.2 Odd One Out (Post-Intervention), Hierarchical Regression Analysis

As shown in Table 4.8 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention Odd One Out scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 30% of the variance in post-intervention Odd One Out scores, however, only the pre-intervention score of Odd One Out was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing this variable to be a significant predictor accounting for a further 48.3% of the variance in post-intervention Odd One Out scores. The final model ($F(3,43) = 51.58$, $p<.001$) at step two predicted 78.3% of the Odd One Out score variance (Adj. R^2 .767).

Table 4.8: Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, Odd One Out

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	-3.39 (-15.15-8.38)	5.84		*p*=.57
Pre-Intervention	.85 (.34-1.37)	.26	.45	*p*=.002
Age	.10 (-.03-.23)	.06	.21	*p*=.13
Step 2				
Constant	11.12 (3.84-18.40)	3.61		*p*=.004
Pre-Intervention	.79 (.50-1.08)	.15	.41	*p*<.001
Age	.10 (.03-.18)	.04	.22	*p*=.007
Group	-9.7 (-11.7 - -7.7)	.99	-.70	*p*<.001
Step 1 R^2=.30, *p*=<.001	**Step 2 ΔR^2= .48, *p*<.001**			

4.3.2.3 Odd One Out (9 Month Follow-Up), Hierarchical Regression Analysis

As shown in Table 4.9 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up Odd One Out scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 31% of the variance in 9 month follow-up Odd One Out scores, however, only the pre-intervention score of Odd One Out was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing this variable to be a significant predictor accounting for a further 47.2% of the variance in 9 month follow-up Odd One Out scores. The final model (F (3,43) = 51.47, p<.001) at step two predicted 78.2% of the Odd One Out score variance (Adj. R^2 .767).

Table 4.9 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, Odd One Out

Pre-Intervention–9 month follow-up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	-1.32 (-12.32-9.69)	5.46		*p*=.81
Pre-Intervention	.85 (.36-1.33)	.24	.47	*p*=.001
Age	.08 (-.04-.20)	.06	.19	*p*=.17
Step 2				
Constant	12.20 (5.33-19.06)	3.41		*p*=.001
Pre-Intervention	.79 (.51-1.07)	.14	.44	*p*<.001
Age	.09 (.02-.16)	.03	.20	*p*=.013
Group	-9.04 (-10.92 - -7.15)	.94	-.69	*p*<.001

Step 1 R^2=.31, *p*=<.001 **Step 2 ΔR^2= .47, *p*<.001**

4.4 Discussion

The results reported from the hierarchical multiple regressions show that group membership (trained vs active control) was a significant predictor of both Listening Recall scores (p<.001, β =-.58) and Odd One Out scores (p<.001, β=-.67) immediately post-intervention. At a 9 month follow-up, group membership (trained vs active control) continued to be a significant predictor of both Listening Recall scores (p<.001, β =-.60) and Odd One Out scores (p<.001, β=-.69). Therefore, as predicted, following the working memory training implemented with children with language difficulties, group (trained vs active control) was a significant predictor of directly trained task performance immediately post-intervention and at a 9 month follow-up. These results show that not only did the trained group of participants respond to the working memory training intervention by improving in directly trained tasks, but these gains were maintained for a further nine months. Findings support previous research which has shown evidence of gains, immediately following an intervention and at a follow-up several months later, in either identical or closely related working memory trained tasks (Henry et al., 2014; & Melby-Lervag et al., 2016).

In the Listening Recall task, 100% of trained participants improved between both pre-intervention and 9 month follow-up, and between pre-intervention and post-intervention. Between post-intervention and 9 month follow-up, 41.7% of trained participants continued to improve in Listening Recall scores and 16.6% remained at the same level. Therefore, between post-intervention and the 9 month follow-up period, 58.3% of trained participants either maintained their initial improvement made during training or improved this level further. Although this also indicates that 41.7% of trained participants decreased in Listening Recall scores during the same time period, it is important to note that, as 100% of participants improved their scores between pre-intervention and 9 month follow-up, all participants achieved a higher Listening Recall score at the 9 month follow-up than they had prior to the training intervention.

Similar results were also found for trained participants in the Odd One Out span task, with 100% of trained participants again improving from both pre-intervention to 9 month follow-up, and between pre-intervention and post-intervention. From post-intervention to the 9 month follow-up period, 50% of trained participants continued to improve their Odd One Out scores and 8.3% retained their original gain. Therefore, from post-intervention to the 9 month follow-up, 58.3% of trained participants either maintained their original gain in Odd One Out scores or further improved. Similar to Listening Recall, although this indicates that 41.7% of trained participants decreased their Odd One Out scores from post-intervention to 9 month follow-up, as 100% improved from pre-intervention to follow up, this again shows that all trained participants still scored a higher Odd One Out score on completion of the study compared to the start.

As 41.7% of trained participants continued to improve their Listening Recall scores and 50% continued to improve Odd One Out scores between post-intervention and the 9 month follow-up, it could be considered whether this 9 month time period allowed the trained participants time to consolidate and process any learning achieved during the working memory training intervention. This time of consolidation therefore allowed these trained participants to achieve further working memory gains without any additional training being implemented.

Interestingly, 73.9% of active controls improved in the Odd One Out task from post-intervention to the 9 month follow-up, compared to 43.5% in Listening Recall. Not only does this comparison emphasise the potential difficulties that those in the language disorder sample may have had in verbal working memory compared to visuospatial ability, it also highlights the robust improvements made by the trained participants in the Odd One Out span task. Even though 73.9% of active controls improved in Odd One Out scores between post-intervention and the 9 month follow-up time point, group was still a significant predictor of Odd One Out task performance at the 9 month follow-up, indicating the high level of improvement made by the trained group of children.

Overall this chapter has shown that those with a language disorder engage positively to a face to face adaptive working memory intervention that targets both the verbal and visuospatial executive-loaded task domains. Not only was group membership (trained vs active control) a significant predictor of performance in both Listening Recall and Odd One Out tasks immediately post-intervention, this was also maintained for a further nine month period.

Chapter 5 will explore individual differences within the trained group of children in more depth. This detailed analysis may help us better understand the processes and reasons for improvements in directly trained tasks that have been discussed in the current chapter. Chapter 1 highlighted the heterogeneity of children with language disorders, and this is a key issue. Further exploration of individual differences in the current sample of trained children will add a novel contribution of the current research.

Chapter 5

Progress by Trained Participants during the Intervention

5.1 Introduction

As discussed in chapter one, section 1.1, heterogeneity amongst those with language disorder is considerable. An awareness of heterogeneity within the language disorder population is essential when evaluating previous research. It could be assumed that clinical groups have been compared on a like for like basis, whereas this is not always the case. There is clear discrepancy between diagnostic criteria in terms of language ability and nonverbal intelligence levels for those with language difficulties (Leonard, 2014). Co-existing developmental disorders, commonly exhibited in those with language difficulties, display variation in the profiles of this group (Bellair, 2014; Bishop, 2014). Across all analysis in the current study, interestingly age does not predict working memory performance. When considering profiles of this group, this suggests the participants in the current study may display a range of working memory impairments that mean they are unlikely to show typical patterns of development and follow age expected ability. Differences in findings between participant groups within working memory training studies have raised questions about the types of children who may benefit from working memory intervention (Rowe et al., 2019a), about individual differences within groups of children, and about the gains made during the intervention. Individual differences have been shown to play a pivotal role in working memory difficulties, and as discussed previously severity and persistence of language disorder are evidenced as factors in visuospatial difficulty (Blom & Boerma., 2020). The current study, therefore, explored individual differences in the trained group of participants to gain a clearer understanding of these children and their heterogeneity, and explore factors that might account for different outcomes.

This chapter will address the following research questions:

i) Does looking at high versus low improvers provide any insight into the variability in trained task performance?

ii) Do participants who have a higher increase in span levels in Listening Recall or Odd One Out tasks over the 18 training sessions show a different trajectory than those who improve less?

iii) Can improvement levels be linked to those who display a ‘spikey’ profile, (a fluctuation in span levels throughout training sessions), compared to those who display a ‘smooth’ profile (a relatively steady increase from sessions 1-18)?

iv) Does the current intervention require all 18 training sessions to be administered for training gains to be realised, or can dosage be adjusted for future interventions?

The additional and novel analyses presented in this chapter will focus solely on the trained group of participants (n=24). Exploration of individual differences in response to intervention among trained participants in this manner is rare, and will provide additional novel research findings that could be useful for clinicians. These data will allow for further insight and enhance our understanding of what is happening during a working memory training intervention.

5.2 Methodology

5.2.1 Participants

This chapter and its analysis focus solely on the trained participants in the current study, active control participants are not included in this chapter's analysis. The participants include the 24 children who were randomly allocated from a group of 47 participants, to a trained intervention group. For further inclusion criteria and a detailed overview of the current study's participants please refer to chapter 3, section 3.2.

5.2.2 Measures and Procedures

The focus of this chapter is an exploration of whether individual differences of trained participants could account for variability in performance in the directly trained tasks, i.e., Listening Recall and Odd One Out. For detailed procedural information on both these tasks, please refer to chapter 3, section 3.3.1, with section 3.3.2 detailing the training materials and procedures required to administer both working memory trained tasks.

5.2.3 Statistical Analysis

All 24 trained participants were initially divided into two subgroups – low and high improvers (for a detailed account of this procedure please go to section 5.3.2 of this chapter). Following this division independent *t*-tests were conducted to explore whether there was a significant difference between the low and high improvers' starting span at session 1, and their final span at session 18.

Further exploration of profile types was conducted through the division of trained participants into two subgroups – 'spikey versus smooth' profile type. 'Spikey' profile types were represented by those who showed fluctuations between span levels across training sessions whereas ‘smooth’ profile types were represented by those who steadily increased their span levels from sessions 1-18 (for a detailed procedure of how participants were grouped into ‘spikey’ and ‘smooth’ profile types please refer to section 5.4.1 in this chapter). A repeated measures two-way ANOVA was conducted to assess the effect of session (sessions one and 18), profile type (spikey versus smooth) and the interaction between these.

To explore dosage effects and whether the number of training sessions in the current intervention could be adjusted for future research, a repeated measures one-way ANOVA was conducted which looked at the difference between mean span levels in both trained tasks across training sessions 1, 6, 12 and 18. Post hoc Bonferroni analysis identified any significant differences between sessions 1 and 6, sessions 1 and 12 and sessions 1 and 18.

5.3 Progress by Trained Participants during the Intervention

The following sections will focus on the overall individual improvement made by trained participants in both trained tasks (section 5.3.1), before exploring improvement made by low and high improvers in Listening Recall and Odd One Out (section 5.3.2).

5.3.1 Overall Individual Improvements made in Listening Recall and Odd One Out Tasks

Figures 5.1 and 5.2 show an overview of the mean Listening Recall (figure 5.1) and Odd One Out (figure 5.2) span levels for all trained participants (n=24) across all 18 working memory training sessions. Each line represents one trained participant’s mean span level at each of the 18 working memory training sessions in each task.

Figure 5.1 Mean Listening Recall Span Levels for Each Trained Participant at each of the 18 Working Memory Training Sessions

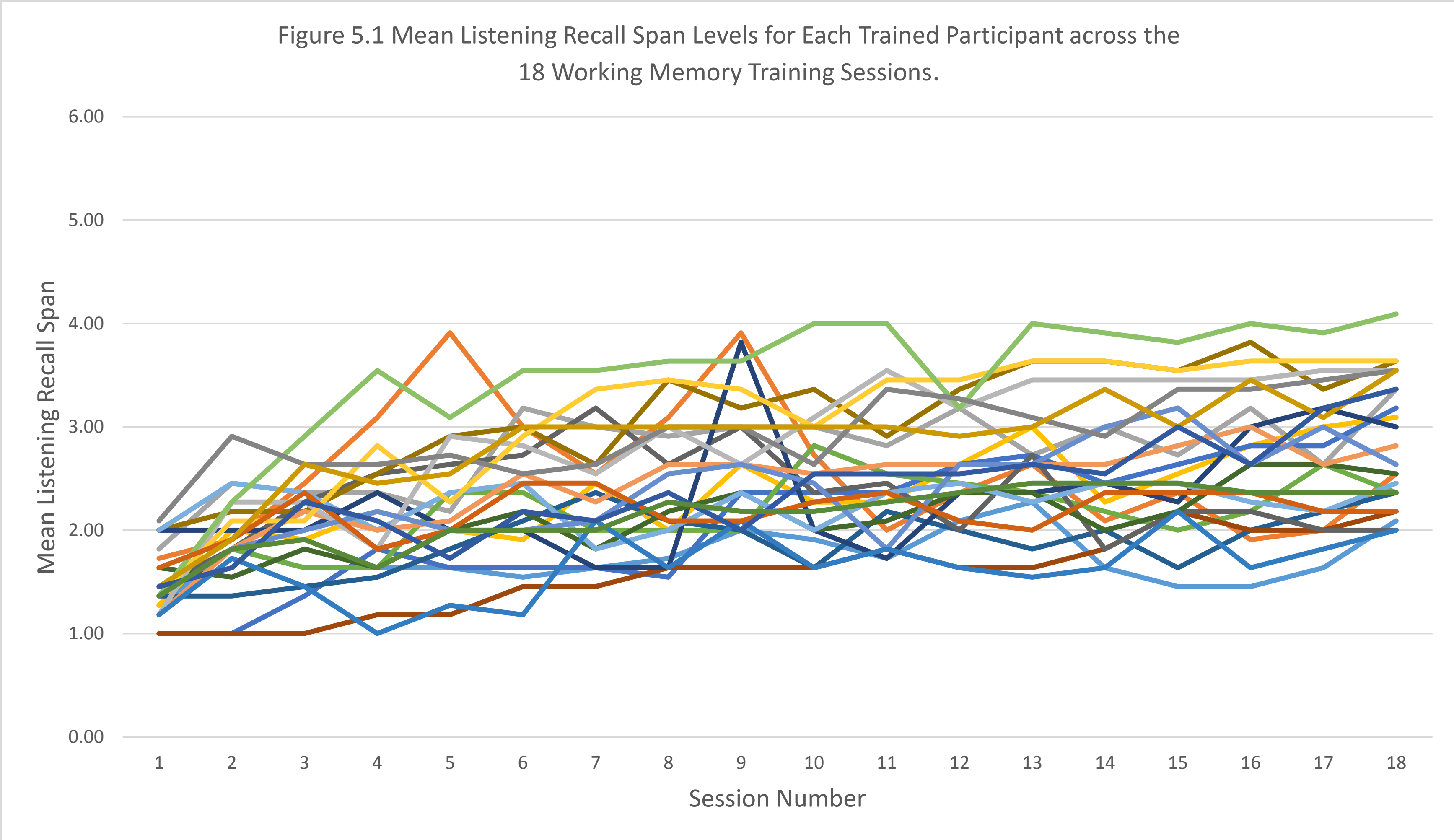
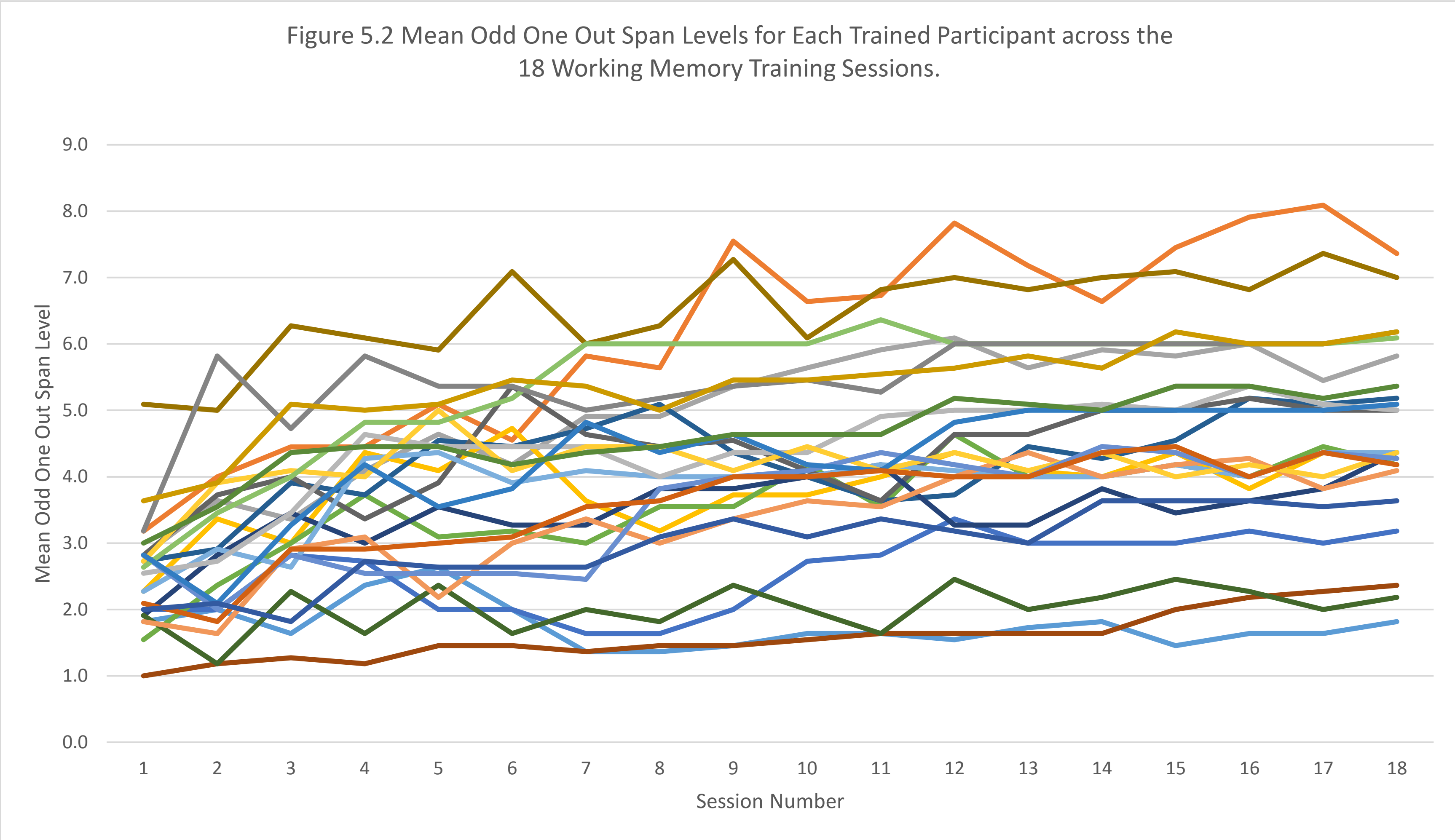


Figure 5.2 Mean Odd One Out Span Levels for Each Trained Participant at each of the 18 Working Memory Training Sessions



Figures 5.1 and 5.2 show that there are distinct differences in span levels achieved across all participants in Listening Recall and Odd One Out tasks. Although the lowest mean span level achieved in both trained tasks was one, the highest mean span level achieved was 4.09 for Listening Recall and 8.1 for Odd One Out. As figures 5.1 and 5.2 show, there were a range of trajectories, some relatively smooth and others with distinct peaks and troughs. The next section will explore the high and low improvers in both trained tasks to investigate any variability in trained task performance.

5.3.2 Low versus High Span Level Improvement in Listening Recall Task

Trained participants who had the lowest and highest Listening Recall span level difference between session 1 and session 18 were identified. The mean span level of all 11 trials for each training session was calculated, giving 18 mean span level scores per participant. The trained participants were then ranked in order from 1-24 depending on the difference between their mean span score in session one and the mean span score at session 18. This, therefore, showed the total improvement from start to finish of the training intervention, and the span level that the participants were working at upon completion of the training sessions. As 24 trained participants took part, 12 participants were assigned to each of the two subgroups.

5.3.2.1 Initial and Final Span Levels of Low vs High Improvers in Listening Recall

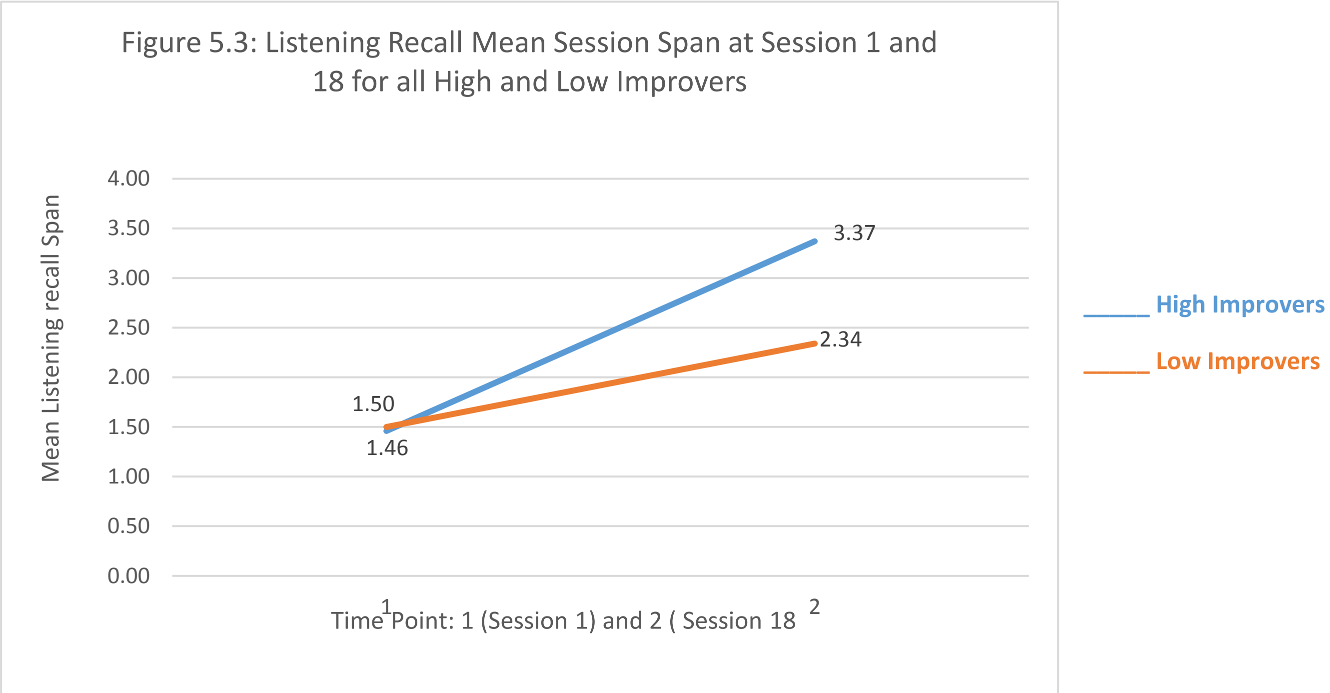
Table 5.1 presents the descriptive statistics (mean, standard deviation and range) at session 1 and session 18 for the low vs high improvers in Listening Recall.

Table 5.1 Mean, standard deviation and range of mean Listening Recall span levels at session 1 and 18 for the low and high improvers

	Mean Listening Recall Span	*SD*	*Range*
Session 1			
Low Improvers (n=12)	1.50	.31	1:00-2.00
High Improvers (n=12)	1.46	.34	1.00-2.09
Session 18			
Low Improvers (n=12)	2.34	.28	2.00-3.00
High Improvers (n=12)	3.37	.39	2.64-4.09

As is evident in table 5.1, both low and high improvers had very similar Listening Recall span levels at session 1, but a notable difference of 1.03 at session 18. Using an independent samples *t*-test, it was confirmed that there was no significant difference in Listening Recall span between the low and high improvers at session 1 of the training, $t(22) = .000$, p=.50 (one tailed). A further independent *t*-test showed that by session 18, there was a significant difference between the mean span scores in Listening Recall of the low and high improvers, $t(22)= -6.92$, $p<.001$ (one tailed). Based on these results, illustrated in Figure 5.3, it can be assumed that any variability across performance from session 1-18 is due to training effects.

Figure 5.3 Listening Recall mean session span at session 1 and 18 for high and low improvers



5.3.2.2 Trajectory of Low vs High Improvers in Listening Recall

Figures 5.4 and 5.5 show the trajectory of the 12 trained participants in the low and high improver groups respectively. Each line represents one trained participant's mean Listening Recall span level across all 18 working memory training sessions.

Figure 5.4 Listening Recall span levels from training sessions 1-18 in the low improver group

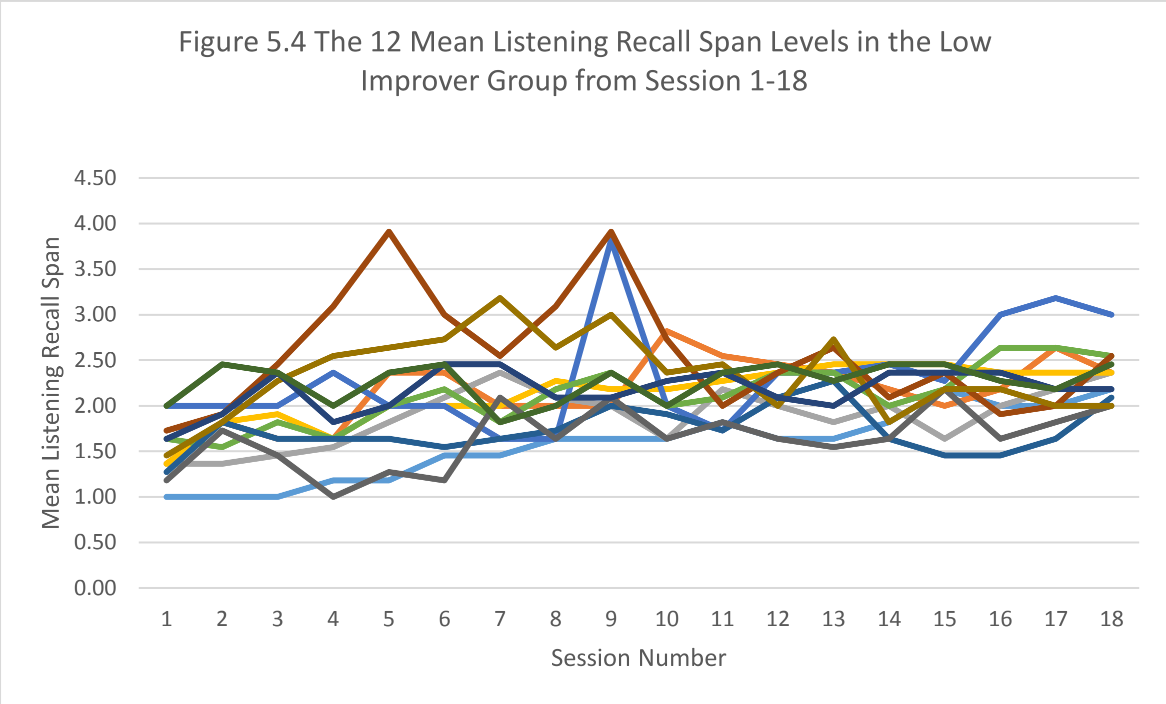
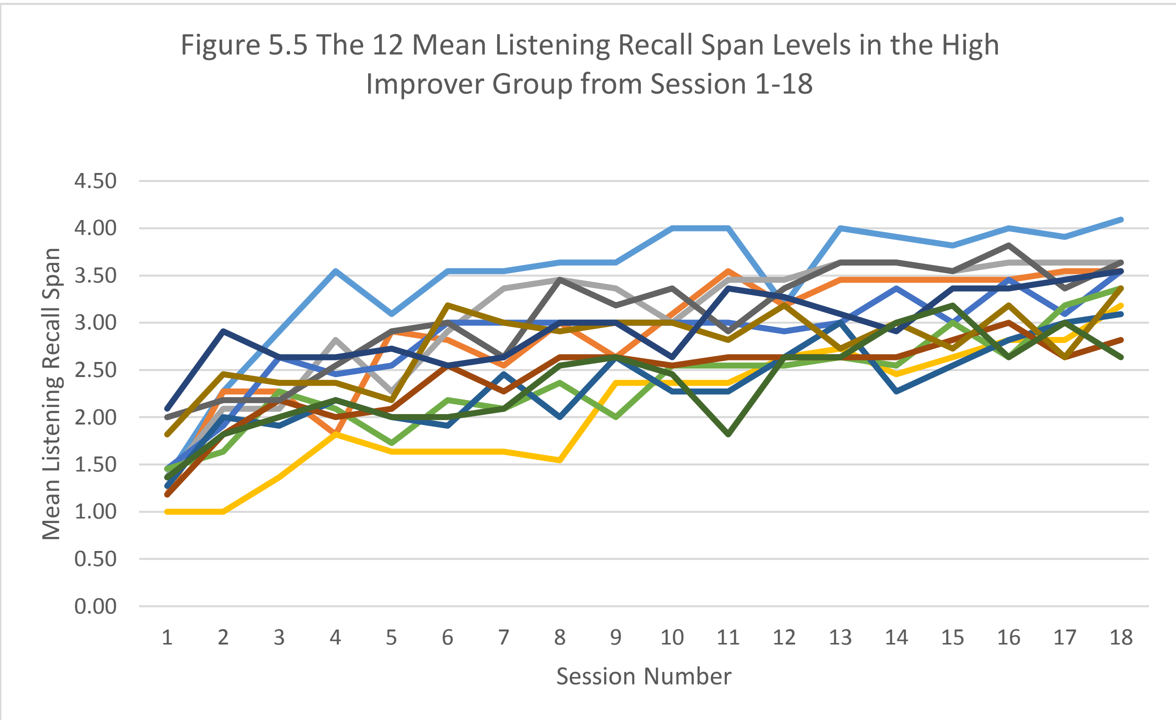


Figure 5.5 Listening Recall span levels from training sessions 1-18 in the high improver group



Figures 5.4 and 5.5 indicate that the 12 trained participants who improved the least from training sessions 1-18 seem to have done so with more fluctuation of span levels between training sessions than the 12 highest improvers. Low improvers seem to display more spikes in improvement level which seem to decrease rapidly. Interestingly, figures 5.4 and 5.5 show that the peak Listening Recall span level achieved within both low and high improver groups is of a similar level, with the peak span level achieved within the low improvers being 3.91 compared to 4.09 in the high improver group. This could suggest that some participants within the low improver group have the potential to reach as high a span level in Listening Recall as their peers in the high improver group, however, they can only maintain this gain for a limited period of time.

5.3.2.3 The Two Lowest and Highest Improvers in Listening Recall Scores

This section will now take a closer look at the two highest and lowest improvers of Listening Recall to see if patterns in their training span levels give any indication as to the variance in their performances throughout training.

Figure 5.6 Two lowest Listening Recall span improvers from Session 1-18

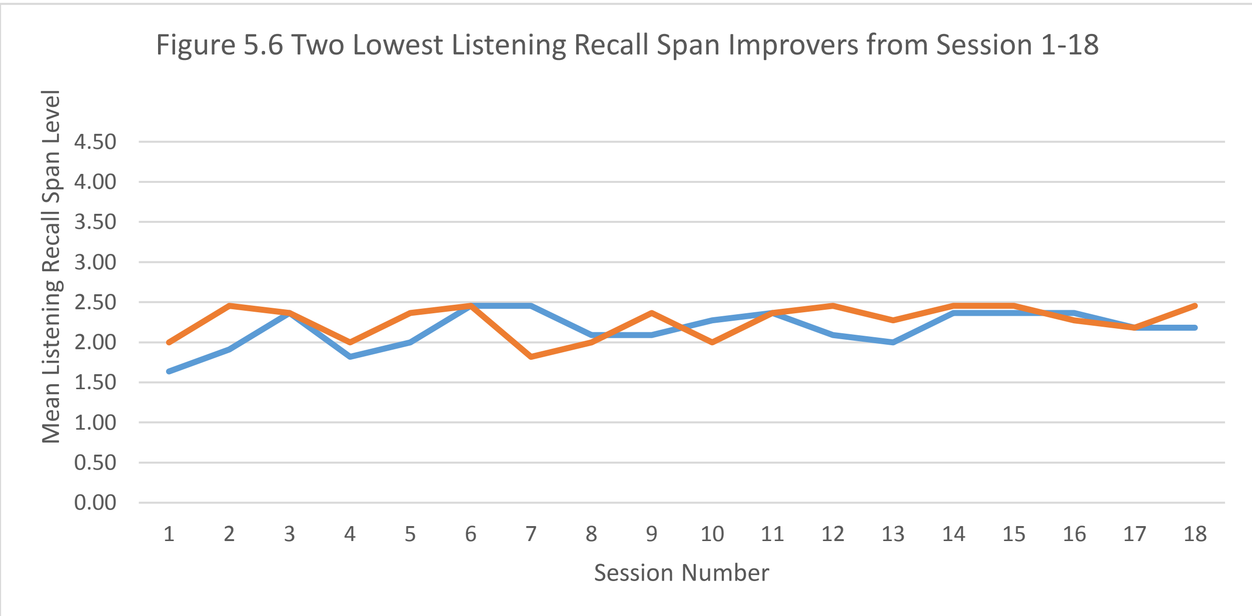


Figure 5.7 Two highest Listening Recall span improvers from Session 1-18

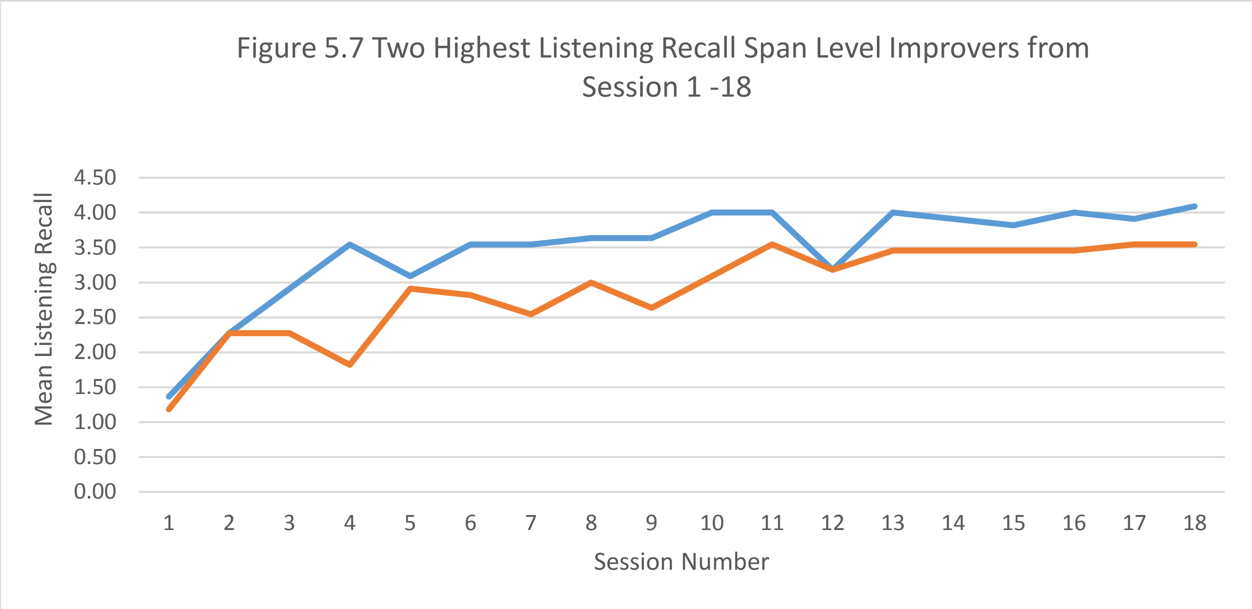


Figure 5.6 shows that the two lowest improvers seem to stay at a fairly stable span level throughout their 18 training sessions, with only an increase in Listening Recall span of 0.45 (see orange line) and 0.55 (see blue line) between training sessions 1-18. In comparison, figure 5.7 shows that the two highest improved participants in Listening Recall seem to improve sharply until approximately training session five, continue to steadily increase until approximately session 12, where any improvement then seems to plateau. Although there are substantial differences in the Listening Recall span level increase from training sessions 1-18 between the two lowest and two highest improvers, interestingly all four participants start at a similar span level, with the highest improvers shown in figure 5.7 being at a lower initial span of 1.18 (see orange line) and 1.36 (see blue line), compared to the lowest improvers in figure 5.6 starting at 1.64 (see blue line) and 2 (see orange line).

5.3.3 Low versus High Span Level Improvement in Odd One Out

This section will now examine more closely the trained participants who had the lowest and highest Odd One Out span level difference from session 1-18. As explained in section 5.3.2, the same method was used in order to divide participants into two subgroups of 12, one representing the lowest Odd One Out improvers and one representing the highest.

5.3.3.1 Initial and Final Span Levels of Low vs High Improvers in Odd One Out

Table 5.2 presents the descriptive statistics (mean, standard deviation and range) at session 1 and session 18 for the low vs high improvers on the Odd One Out task.

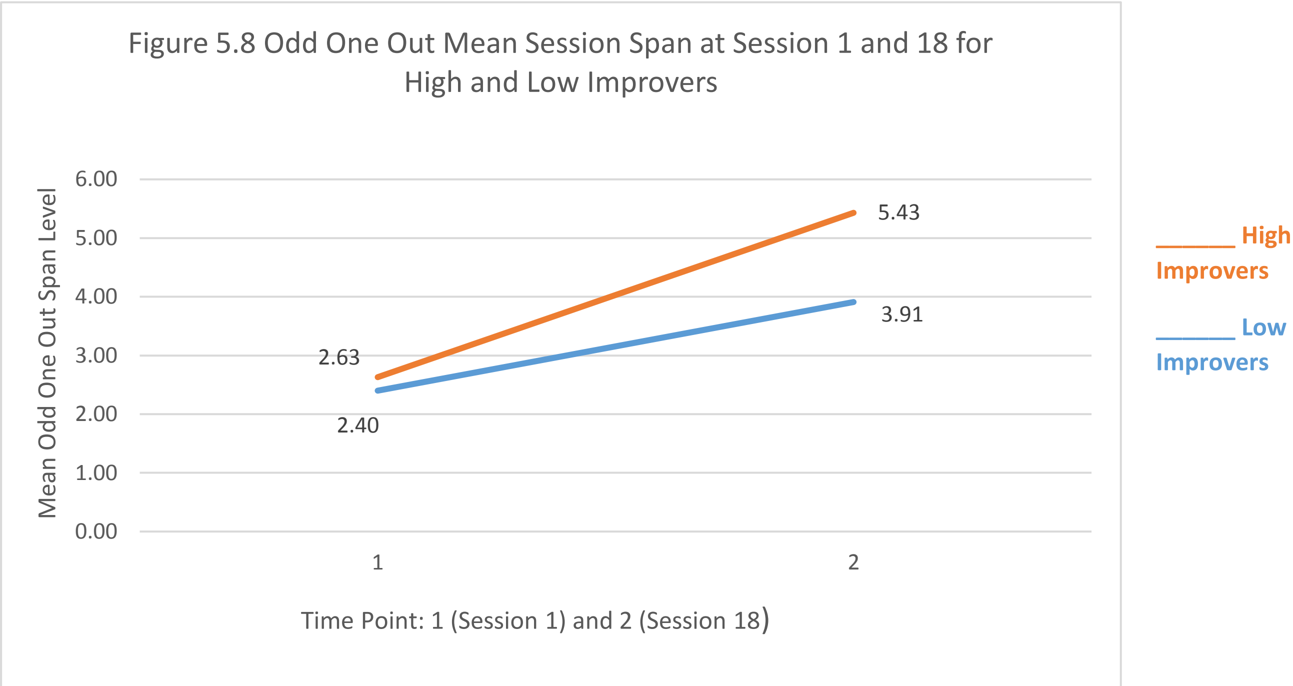
Table 5.2 Mean, standard deviation and range of mean Odd One Out span levels at session 1 and 18 for the low and high improvers

	Mean Listening Recall Span	*SD*	*Range*
Session 1			
Low Improvers (n=12)	2.40	.99	1.00-5.10
High Improvers (n=12)	2.63	.63	1.50-3.60
Session 18			
Low Improvers (n=12)	3.91	1.41	1.80-7.00
High Improvers (n=12)	5.43	.97	4.10-7.40

As is evident in table 5.2, both the trained and active control groups had very similar Odd One Out span levels at session 1, but a notable difference of 1.52 at session 18. Using an independent samples *t*-test, it was confirmed that there was no significant difference in Odd

One Out span between the low and high improvers at session 1 of the training sessions, $t(22) = .69$, $p=.25$ (one tailed). A further independent *t*-test showed that by session 18, there was a significant difference between the mean span scores of the low and high improvers in the Odd One Out task, $t(22)= 3.07$, $p =.003$ (one tailed). Based on these results, illustrated in Figure 5.8, it can be assumed any variability across performance from sessions 1-18 is due to training effects.

Figure 5.8 Odd One Out mean session span at session 1 and 18 for both high and low improvers



5.3.3.2 Trajectory of Low vs High Improvers in Odd One Out

Figures 5.9 and 5.10 show the trajectory of the 12 trained participants in the low and high groups respectively. Each line represents one trained participant's mean Odd One Out span level across all 18 working memory training sessions.

Figure 5.9 Odd One Out span levels from training sessions 1-18 in the low improver group

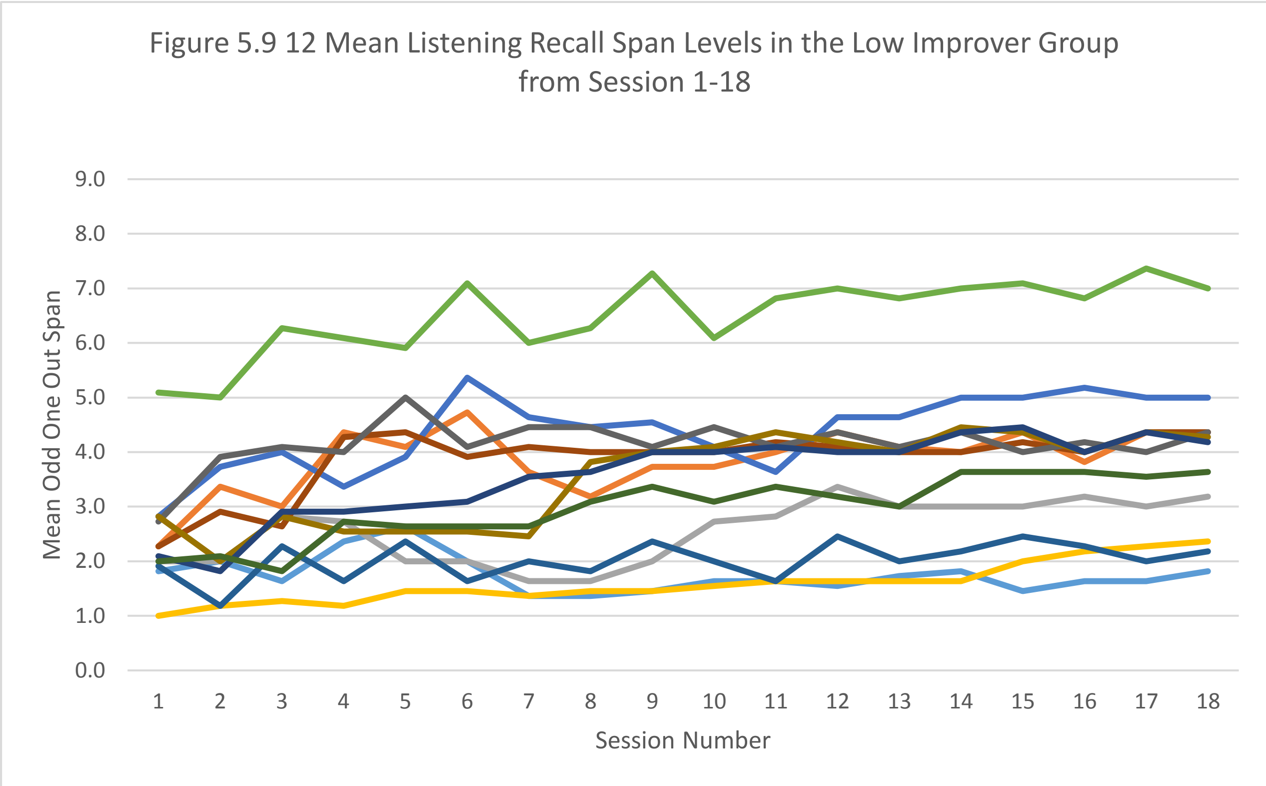
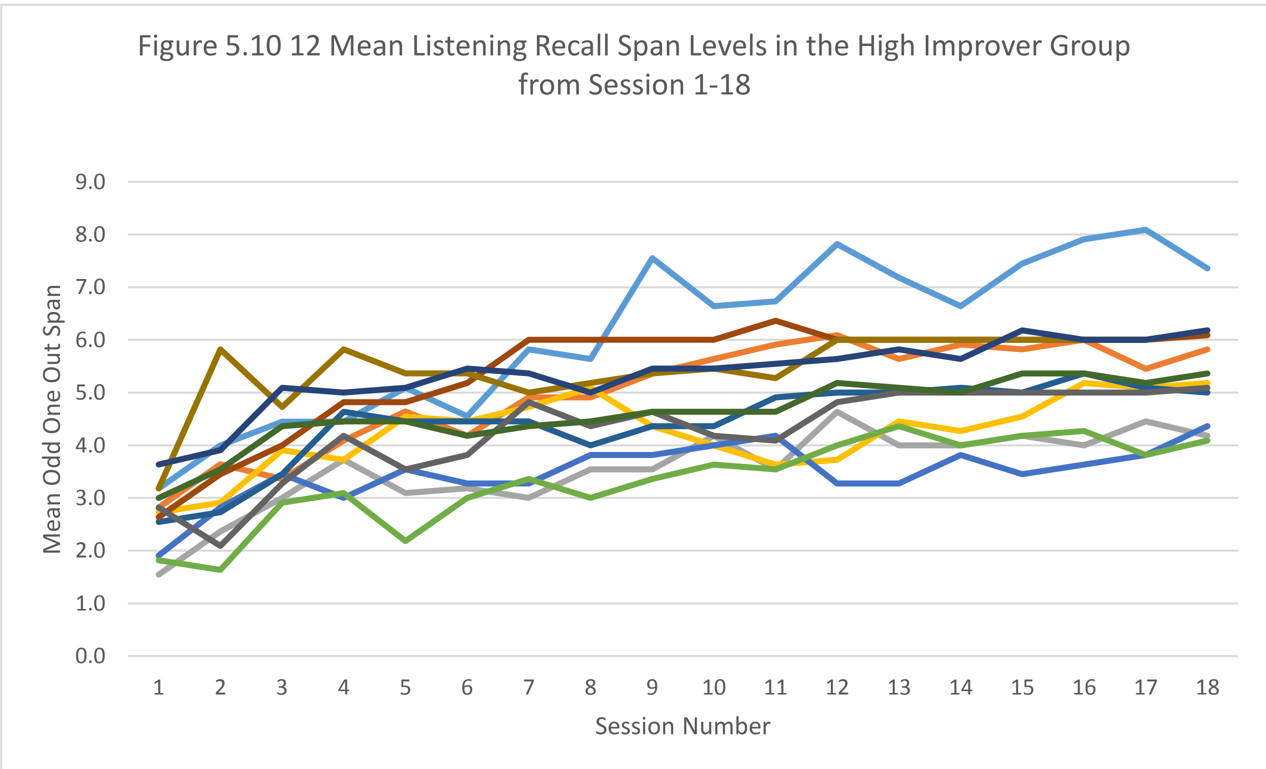


Figure 5.10 Odd One Out span levels from training sessions 1-18 in the high improver group



Similar to figures 5.4 and 5.5, (Listening Recall span from training sessions 1-18 in low and high improver groups, section 5.3.2.2), both the low and high improver groups in the Odd One Out span reached peak span levels that are not too dissimilar between groups, with the low improver group peaking at a span level of 7.4 and the high improver group reaching 8.1 span levels. What is interesting, however, is the large range of initial starting Odd One Out Span levels, compared to Listening Recall in figures 5.4 and 5.5, which can be seen clearly via the participants represented by the bottom yellow line and top green line in figure 5.9 above. Although the mean span level across all high and low improved participants in session one was found to be not significantly different via an independent *t*-test, see section 5.3.3.1, the range of initial mean Odd One Out span levels is between 1 - 5.1 for the low improvers and between 1.8 – 7 for the high improvers (see table 5.2), indicating that children with language disorder can present a varied ability in visuospatial working memory.

5.3.3.3 The Two Lowest and Highest Improvers in Odd One Out Span.

This section will now take a closer look at the two highest and lowest improvers on the Odd One Out task to see if patterns in their training span levels give any indication as to the variance in their performances throughout training.

Figure 5.11 The two lowest Odd One Out span improvers from session 1-18

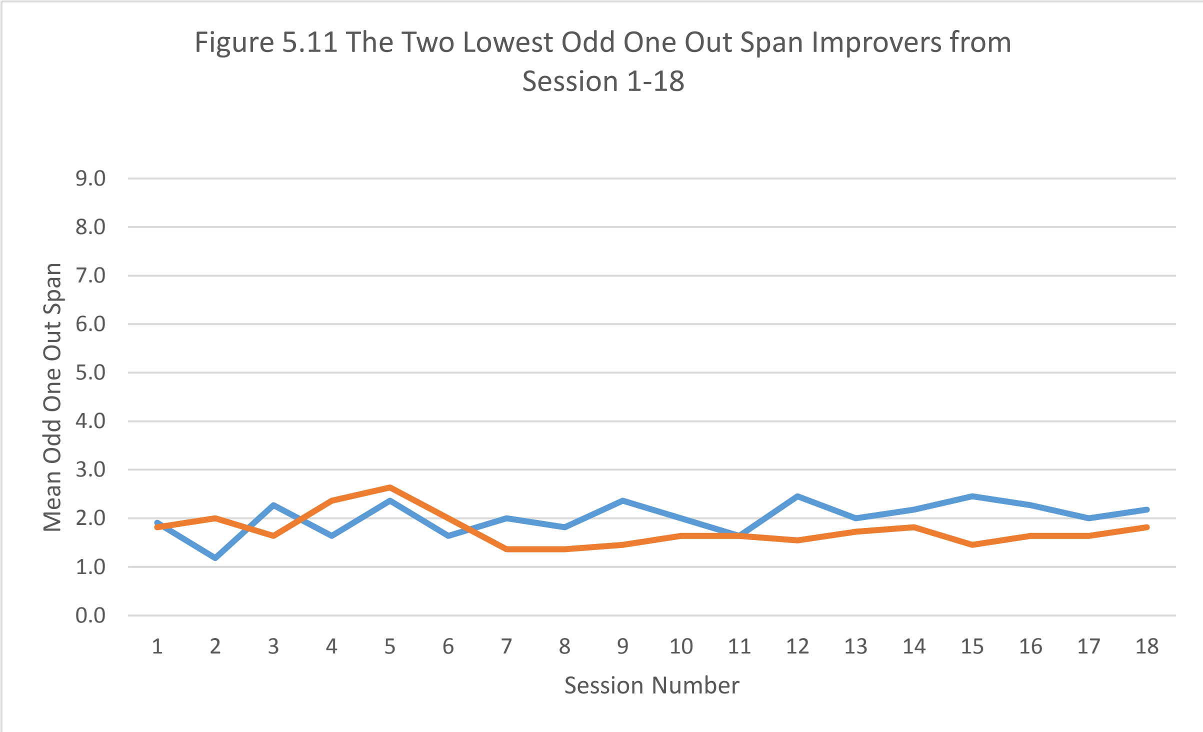
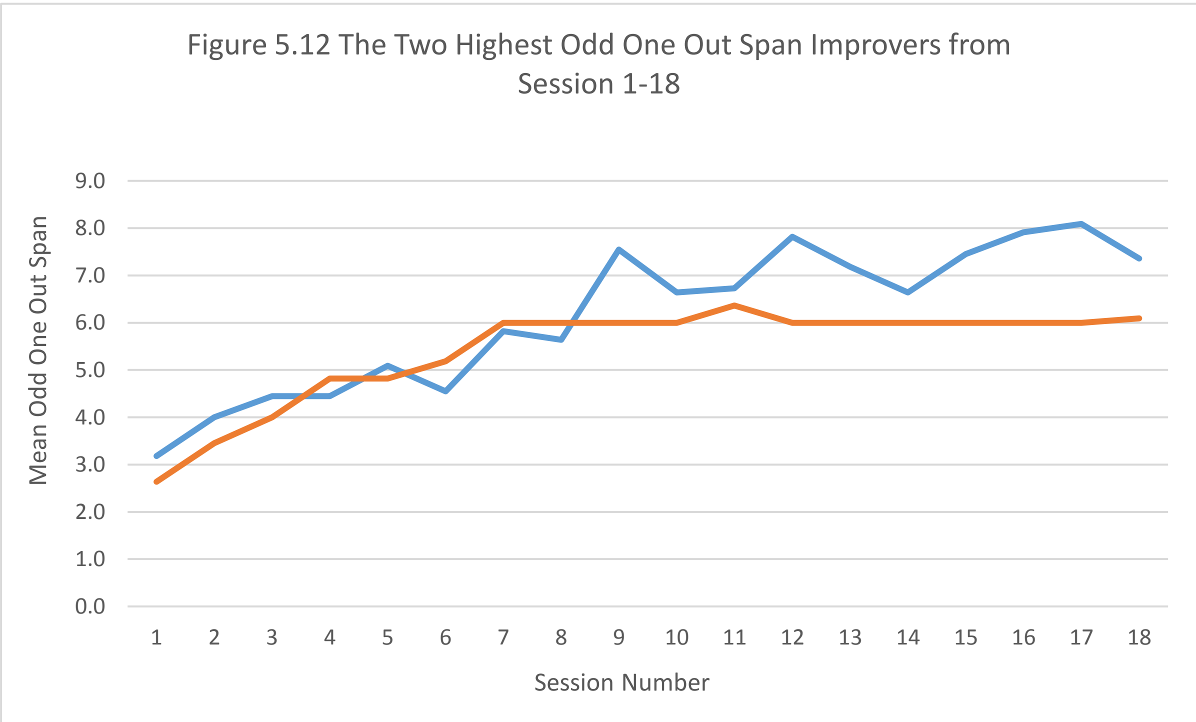


Figure 5.12 The two highest Odd One Out span improvers from session 1-18



Figures 5.11 and 5.12 show clearly the difference in trajectory for the two lowest and highest improvers on the Odd One Out span task. The two lowest improvers, represented in figure 5.11, show minimal increase in their span level differences from session 1-18, and mean span levels remain relatively stable throughout the 18 training sessions. The two highest improvers on Odd One Out span, in figure 5.12, show a steady improvement in the initial working memory training sessions, however, the orange participant shows relatively little change in span level with just a further increase of 0.4 following session 7, and the blue participant only increasing a further 0.5 in span level between sessions 9 and 18. Therefore, it seems that the most progress and improvement in Odd One Out span levels may occur during sessions 1-9.

Figures 5.1 and 5.2, presented in section 5.3.1, showed individual participants' span levels across all 18 training sessions, in both Listening Recall and Odd One Out span tasks. These figures showed a range of trajectories, some comparatively smooth whilst others had clear peaks and troughs. Based on this observation, participants were divided into two further subgroups, with 'spikey' versus 'smooth' profiles. The next section of this chapter will address how these profile types may influence improvement levels in both Listening Recall and Odd One Out trained tasks.

5.4 Was a Spikey Versus Smooth Profile Associated with Improvement Level?

Each trained participant (n=24) has been shown via figures 5.1 and 5.2, section 5.3.1, to respond differently to each trained task across the 18 training sessions. Some participants display a spikey profile, where their mean span score fluctuates between training sessions, whereas some display more of a smooth profile, remaining fairly steady throughout. This section will address whether having a particular profile type, spikey or smooth, was related to overall improvement in both Listening Recall and Odd One Out tasks.

5.4.1 Listening Recall, Spikey Vs Smooth Profile

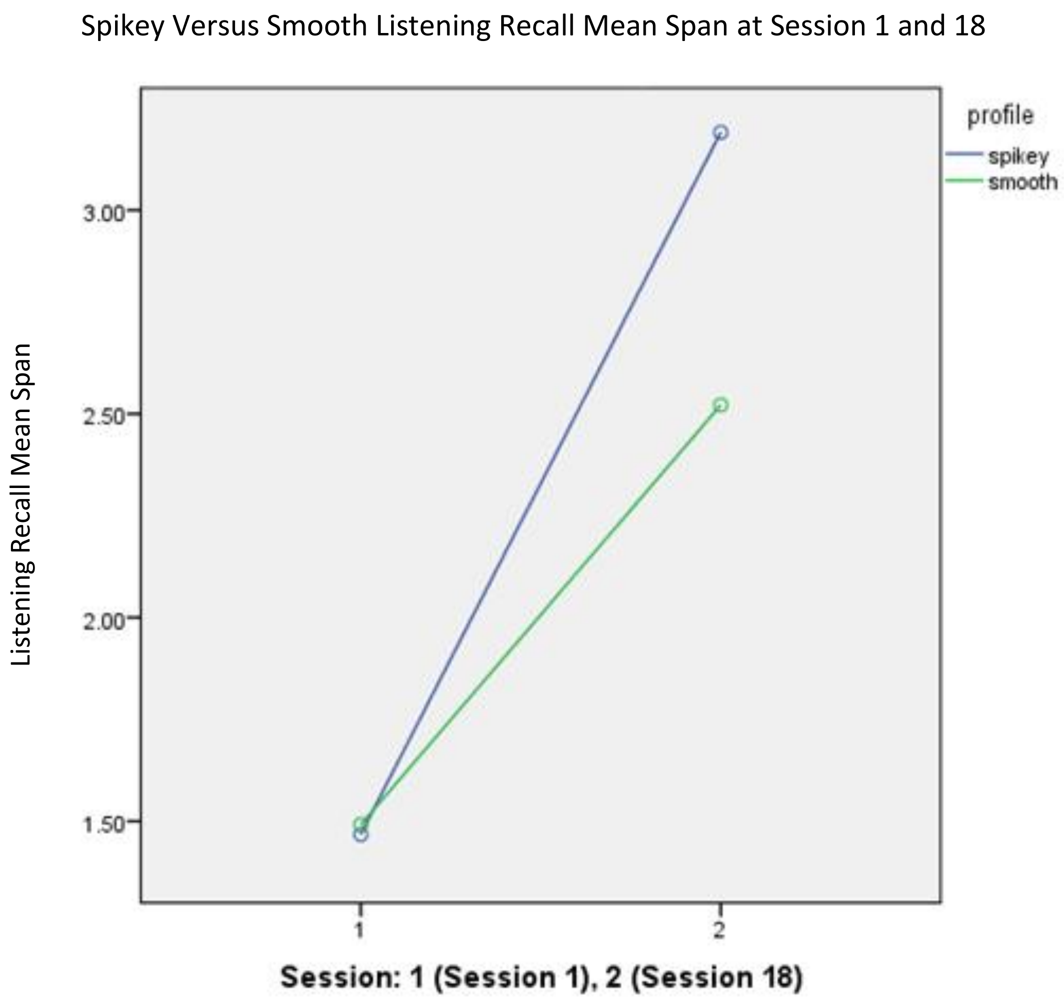
To divide participants into two profile types, 'Spikey' vs 'Smooth', each trained participant was given a mean span level for each training session, giving each trained child 18 mean span levels for Listening Recall and Odd One Out tasks. From this the standard deviation was calculated and those with the highest 12 standard deviations were grouped as 'spikey' and those with the lowest 12 were grouped as 'smooth'.

Table 5.3 Mean and standard deviation of spikey vs smooth profiles for session 1 and 18 of Listening Recall task

	Mean Listening Recall Span	*SD*	*Range*
Session 1			
Spikey Profile (n=12)	1.47	.30	1.00-2.00
Smooth Profile (n=12)	1.49	.34	1.00-2.09
Session 18			
Spikey Profile (n=12)	3.19	.58	4.10-7.40
Smooth Profile (n=12)	2.52	.49	2.00-3.55

As is evident in Table 5.3, both the spikey and smooth profile groups had a very similar Listening Recall span level at session 1, but a notable difference of 0.67 at session 18, illustrated in figure 5.13.

Figure 5.13 Listening Recall mean span levels at sessions 1 and 18 for spikey and smooth profile types



A repeated-measures two way ANOVA showed that there was an effect of session, with significant differences between mean Listening Recall span in sessions 1 and 18, $F(1,23) = 74.81$, $p<.001$, *(partial)* $\eta^2 =.77$.

Profile type (spikey vs smooth) was also found to make a significant difference to Listening Recall span, $F(1,23)= 20.88$, $p <.001$ (*partial)* $\eta^2=.48$. The interaction between session and profile type (spikey versus smooth) was also found to be significant, *F(*1,23) = 62.19, *p<.001, (partial)* η^2 =.73. These results are clearly illustrated in figure 5.13.

Follow up independent samples *t*-tests explored the interaction. There was no significant difference between Listening Recall session 1 span levels for those in the 'spikey' profile versus the 'smooth' profile, *t*(22) = -.18, *p*=.86. However, there was a significant difference between Listening Recall session 18 span levels for those in the 'spikey' profile versus the 'smooth' profile, *t*(22) = 3.01, *p* =.01. Therefore, the group difference at session 18 between smooth and spikey profile children can be assumed to be due to training effects.

It can be inferred, from the repeated-measures two-way ANOVA results together with the follow-up tests and descriptive statistics reported above in section 5.4.1, that at Session 18 those with a spikey profile had a significantly higher Listening Recall span than those with a smooth profile.

5.4.2 Odd one Out, Spikey Vs Smooth Profile

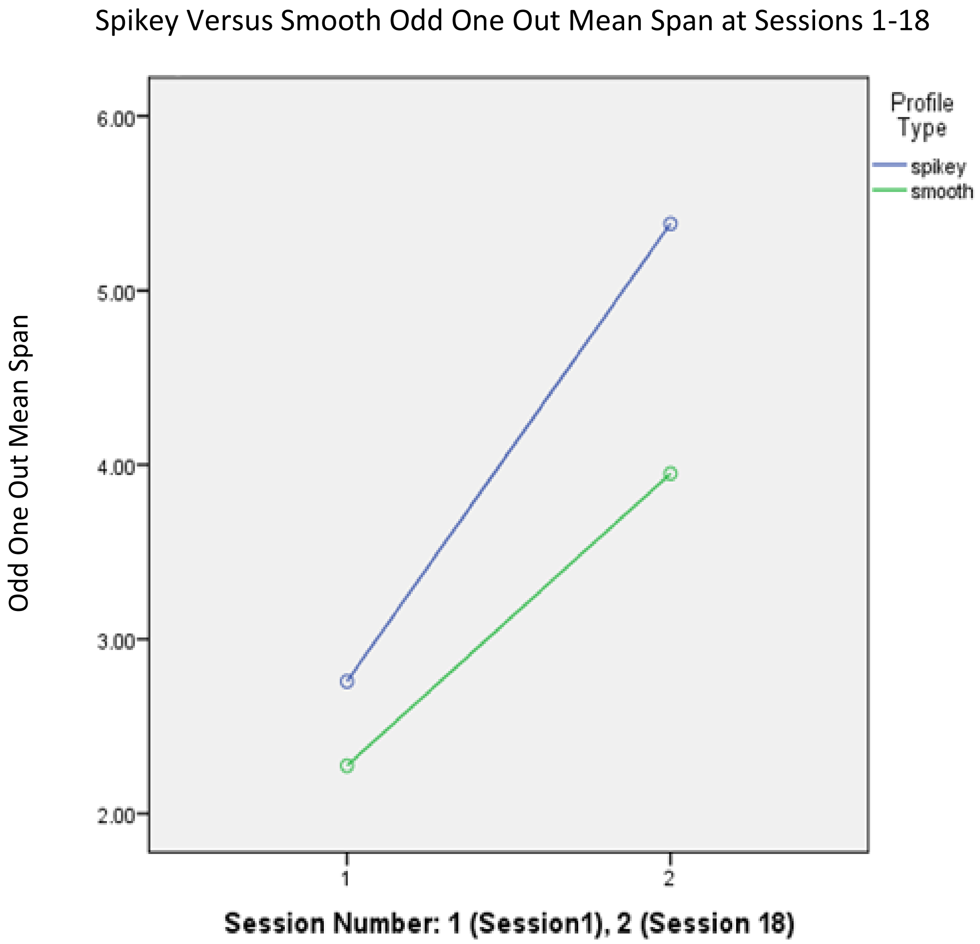
As explained in section 5.4.1, all participants were divided into 'spikey' versus 'smooth' profiles for both Listening Recall and Odd One Out Span tasks. The following section will explore these profiles in relation to the Odd One Out task.

Table 5.4. Mean and standard deviation of spikey vs smooth profiles for session 1 and 18 of Odd One Out task

	Mean Odd One Out Span	*SD*	*Range*
Session 1			
Spikey Profile	2.76	.90	1.50-5.10
Smooth Profile	2.28	.68	1.00-3.60
Session 18			
Spikey Profile	5.38	1.12	4.10-7.40
Smooth Profile	3.95	1.34	1.80-6.20

As is evident in table 5.4, both the spikey and smooth profile groups had a very similar Odd One Out span level at session 1, but a notable difference of 1.43 at session 18, illustrated in figure 5.14.

Figure 5.14 Odd One Out mean span levels at session 1 and 18 for spikey vs smooth profile types



The repeated-measures two way ANOVA results show that there was an effect of session, with a significant difference between mean Odd One Out span in sessions 1 and 18, $F(1,23) = 92.98$, $p<.001$, *(partial)* $\eta^2 = .80$. Profile type (spikey vs smooth) was also found to make a significant difference to Odd One Out span, $F(1,23)=38.80$, $p<.001$ *(partial)* $\eta^2=.63$. The interaction between sessions and profile type was also found to be significant, $F(1,23) = 126.57$, $p<.001$, *(partial)* $\eta^2 = .85$. These results are illustrated in figure 5.14.

Follow up independent samples *t*-tests explored the interaction. There was no significant difference between Odd One Out session 1 span levels for those in the 'spikey' profile versus the 'smooth' profile groups, $t(22) = 1.49$, $p=.15$. However, there was a significant difference between Odd One Out session 18 span levels for those in the 'spikey' profile versus the 'smooth' profile, $t(22) = 2.83$, $p=.01$. It can be inferred from the repeated-measures two-way ANOVA results, the follow-up tests and the descriptive statistics reported in section 5.4.2, that at session 18 those with a spikey profile had a significantly higher Odd One Out score than those with a smooth profile.

5.5 Should The Dosage of The Working Memory Intervention Be Altered?

This section will explore whether the dosage of the working memory intervention implemented in this study could be reduced without affecting outcomes. From looking at different profile types throughout this chapter, it is clear that many children show fluctuations in their span levels for both Listening Recall and Odd One Out trained tasks. The question is whether these fluctuations reach a peak level at a certain point during the 18 training sessions, after which no significant gains are made, suggesting that the intervention may not need to run for the full 18 sessions to achieve potential benefits.

Mean session span scores for all trained participants, for both Listening Recall and Odd One Out trained tasks, were recorded for sessions 1, 6, 12 and 18. A repeated-measures one-way ANOVA was used to explore any significant differences between scores across the four time points. If a significant improvement is found between all scores (at sessions 1, 6, 12 and 18), this indicates the need to continue the working memory intervention until the last session. If, however, no significant difference is found between scores, particularly between session 12 and 18, this could indicate a possible discontinuation point.

5.5.1 Listening Recall Task Dosage Exploration

Table 5.5 shows the mean, standard deviation and range of mean Listening Recall span for sessions 1, 6, 12 and 18.

Table 5.5 Mean and standard deviation of session 1, 6, 12 and 18 Listening Recall span across all trained participants (n=24)

	Mean Listening Recall Span	*SD*	*Range*
Session 1 (n=24)	1.48	.32	1.00-2.09
Session 6 (n=24)	2.36	.60	1.18-3.55
Session 12 (n=24)	2.56	.52	1.64-3.45
Session 18 (n=24)	2.86	.63	2.00-4.09

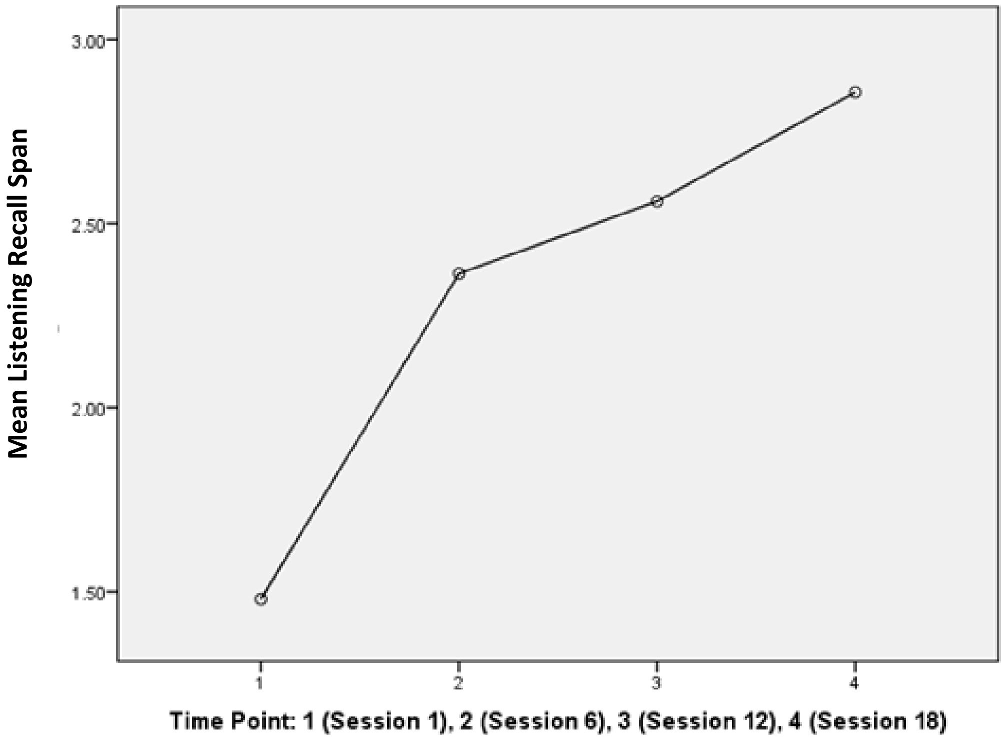
Mauchly's test indicated that the assumption of sphericity had been violated $\chi^2(5) = 21.34$, $p=.001$, therefore Greenhouse-Geisser corrected tests are reported ($\varepsilon=.71$). The repeated-measures one way ANOVA results show that there was a significant overall difference between mean Listening Recall span across training sessions 1, 6, 12 and 18, $F(2.13,49.05) = 66.49$, $p<.001$, *(partial)* $\eta^2 =.74$.

Post hoc Bonferroni analysis indicated that there were significant differences between sessions 1 and 6 ($p<.001$), session 1 and 12 ($p<.001$) and session 1-18 ($p<.001$) showing that participants had a significantly higher span level at each time point compared to their initial starting point at training session 1.

Differences between mean Listening Recall span levels were found to be non-significant between sessions 6 and 12 (p=.22*)* but were again significant between sessions 12 and 18 (p<.001).

Thus, the trained participants had a sharp initial increase in Listening Recall span levels across the first 6 training sessions, which then levelled off between sessions 6 and 12, before making another significant increase between sessions 12 and 18. These results, illustrated in figure 15, indicate that, to achieve the gains observed in the current study, the full 18 training sessions were required for the Listening Recall task.

Figure 5.15 Mean Listening Recall span levels for all trained participants (n=24) in sessions 1,6,12,18



5.5.2 Odd One Out Task Dosage Exploration

Table 5.6 shows the mean, standard deviation and range of mean Odd One Out span for sessions 1, 6, 12 and 18.

Table 5.6 Mean and standard deviation of session 1, 6, 12 and 18 Odd One Out span levels across all trained participants (n=24)

	Mean Odd One Out Span	*SD*	*Range*
Session 1 (n=24)	2.51	.82	1.00-5.10
Session 6 (n=24)	3.83	1.40	1.50-7.10
Session 12 (n=24)	4.46	1.52	1.50-7.80
Session 18 (n=24)	4.67	1.42	1.80-7.40

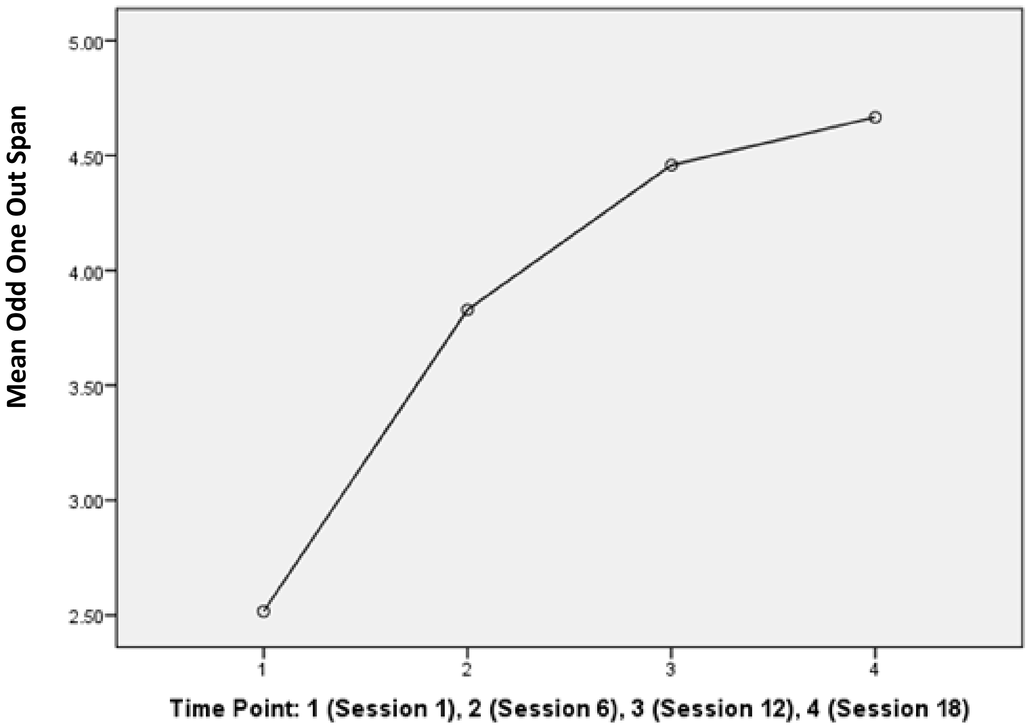
Mauchly's test indicated that the assumption of sphericity had been violated $\chi^2(5) = 21.12$, p=.001*,* therefore Greenhouse-Geisser corrected tests are reported (ε=.70). The repeated-measures one way ANOVA results showed significant differences between mean Odd One Out span levels across training sessions 1, 6, 12 and 18, $F(2.10, 48.31) = 65.00$, p>.001, *(partial)* η^2 =.74.

Post hoc Bonferroni analysis indicated that there were significant differences between sessions 1 and 6 (p<.001), session 1 and 12 (p<.001) and session 1-18 (p<.001) showing that participants scored a significantly higher Odd One Out span level at each time point compared to their initial starting point at training session 1.

Differences between mean Odd One Out span levels were significant between session 6 and 12 (p=.02) but were found to be non-significant between sessions 12 and 18 (p=.20).

Thus, while the trained participants made a significant increase in Odd One Out span levels between sessions 1 to 12 of the working memory training intervention, this tailed off in the remaining 6 sessions. These results, illustrated in figure 5.16, indicate that in the current study, to achieve the gains in Odd One Out span level, the training sessions could potentially have been reduced to 12 sessions.

Figure 5.16 Mean Odd One Out span levels for all trained participants (n=24) in sessions 1,6,12,18



5.6 Overview of Progress by Trained Participants during the Intervention

Individual differences amongst trained participants were firstly examined by separating participants into two subgroups, high and low improvers (see section 5.3.2 for method). An independent *t*-test was conducted to see if there was a significant difference between span levels in both Listening Recall and Odd One Out tasks between low and high improvers at session 1 of the training intervention. Results showed no significant difference between groups for either trained task at session 1, however at session 18, significant differences between groups were evident for both trained tasks. Therefore it can be assumed that any variability across performance from sessions 1-18 is due to training effects.

From looking at the trajectory of low and high improvers for Listening Recall, see figures 5.4 and 5.5, section 5.3.2.2, low improvers tended to show more peaks and troughs throughout their training session span levels compared to high improvers. Although the peak span level achieved by both low and high improvers in Listening Recall was relatively similar, with only a 0.18 difference, this difference in trajectory suggests that the difficulty for low improvers may be related to maintenance of Listening Recall span gains. In comparison, the trajectory for low and high improvers in the Odd One Out task, see figures 5.9 and 5.10, section 5.3.3.2, did not seem to display as many peaks and troughs. The low improver participants in the Odd One Out task did appear to be more spaced out in figure 5.9, representing a varied range of visuospatial ability. Interestingly initial span levels for low improvers in the Odd One Out task ranged from 1.0-5.1, and for the high improvers from 1.5-3.6. Therefore, for both groups, a varied range of visuospatial ability was apparent, in contrast to verbal ability. Previous research has been mixed regarding visuospatial ability in those with language disorder although a meta-analysis has suggested some difficulties (Vugs et al., 2013). As discussed previously, more recent research has suggested that difficulty in visuospatial ability may be associated to the severity and persistence of a child's language disorder (Blom & Boerma., 2020). Therefore a varied range in ability at the start of this intervention study, particularly within the low improver group, supports recent research of mixed abilities in the visuospatial domain for those with language disorder.

The observation of 'peaks and troughs' within training trajectories led to the question of whether improvement can be accounted for by a child who displays a 'spikey profile', with fluctuations between span level throughout training, compared to a 'smooth profile' with relatively steady span level progress throughout training sessions. For both Listening Recall and Odd One Out tasks, participants who displayed a 'spikey' profile made significant gains in both trained task span levels. It could therefore be suggested that in order for children with language disorder to increase span levels and reach a 'peak', they may need to experience a 'trough' in order to have the time to consolidate and secure the previous gains made.

The concept of having enough time to consolidate and secure training gains was also explored further in this chapter by looking at training dosage, questioning whether the number of training sessions could be adjusted in future interventions. Significant differences between mean Listening Recall span across training sessions 1, 6, 12 and 18 were evident, as were significant differences between session 1 and 6, session 1 and 12, and session 1 and 18, indicating that participants had a significantly higher Listening Recall span level at each time point compared to their initial starting point at training session 1. Differences between mean Listening Recall span levels were found to be non-significant between sessions 6 and 12 but were again significant between sessions 12 and 18. This suggests, that in the current study, all 18 training sessions were required for the Listening Recall task, and the non-significance level between sessions 6-12 may again represent the time needed to consolidate any previous gains made, in line with discussion of the 'spikey' profile types above.

Significant differences between mean Odd One Out span levels across training sessions 1, 6, 12 and 18, were also apparent, and there were significant differences between session 1 and 6, session 1 and 12, and session 1 and 18, showing again that participants scored a significantly higher Odd One Out span level at each time point compared to their initial starting point at training session 1. Differences however between mean Odd One Out span levels were found to be significant between session 6 and 12 but non-significant between session 12 and 18. Therefore, indicating that in the current study, to achieve the gains in Odd One Out span level, the training sessions could potentially have been reduced to 12 sessions. Although, it cannot be ruled out that further training in either task could not produce further gains, the current study has given an insight into possible training dosages and patterns of improvement for children with language disorder in both verbal and visuospatial domain trained tasks.

Chapter 6 of this thesis will focus on whether the significant gains made by trained participants for both directly trained tasks, Listening Recall and Odd One Out, transfer to other areas of working memory and produce near transfer effects.

Chapter 6

Near Transfer Effect Results

6.1 Introduction

Significant gains are commonly found in directly trained tasks following a working memory intervention, as shown in Chapter Four of this thesis. Researchers acknowledge the benefits of focussed working memory training to other parts of the working memory system (chapters 1 and 2), i.e. near transfer effects (Sala & Gobet, 2017; chapter 2).

Various meta-analyses have considered the effectiveness of working memory training (Melby-Lervag & Hulme, 2013; Melby-Lervag et al., 2016; Sala & Gobet, 2017; Schwaighofer et al., 2015). For example, a meta-analysis of 23 working memory studies with both clinical and typically developing child and adult samples found evidence for reliable short-term improvements in working memory skills, with verbal working memory showing an initial near transfer effect that was not sustained at follow-up (Melby-Lervag & Hulme, 2013).

Melby-Lervag et al.'s (2016) meta-analysis went on to suggest that working memory training produces short-term improvements on both verbal *and* visuospatial working memory tasks. Verbal working memory tasks, however, were found to yield short-term near transfer effects that were not sustained when reassessed after a delay of a few months post-intervention (Melby-Lervag et al., 2016). Visuospatial working memory was found to have moderate improvements, with some evidence of maintenance at follow-up, however it was often found that the outcome tasks shared features with the working memory training tasks, questioning whether these were true near transfer effects (Melby-Lervag et al., 2016).

However, it is worth noting that a similar intervention to the current study, conducted with typically developing children, targetted two executive-loaded working memory tasks in both the verbal and visuospatial domain This intervention produced near transfer effects in both word recall and counting recall, which were still apparent at a six month follow-up (Henry et al., 2014).

It has been suggested that further working memory intervention studies should be conducted with groups of individuals who are understudied (Rowe et al., 2019a). Children with language disorder fall into this category. They commonly display difficulties with working memory ability, yet research in this field is limited (Chapter 1). A recent study with 32 children with language disorder, matched to chronological age controls, found that following a working memory intervention, which targetted immediate verbal memory, visuospatial working memory and verbal working memory, significant gains were made by the language disorder group in both verbal and visuospatial working memory (Acosta et al., 2019). Thus, children with language disorder have been shown to positively engage with working memory interventions and produce significant gains in areas of related working memory skill.

The current chapter of this thesis aims to answer the following research questions:

i. Do children with language disorder show immediate post-intervention near transfer effects from two executive-loaded working memory trained tasks in the verbal and visuospatial domains?
ii. Can children with language disorder maintain any near transfer effects from two executive-loaded working memory trained tasks in the verbal and visuospatial domains for a period of nine months?

The general consensus from previous research is that working memory training yields improvements in related working memory skills, therefore near transfer effects are often evident (Chapter 2). Various studies, however, have found conflicting findings regarding the maintenance of near transfer effects (Melby-Lervag et al., 2013; Henry et al., 2014). As pointed out above, a similar intervention study with typically developing children obtained immediate post-intervention and six month follow-up near transfer effects in both word recall and counting recall (Henry et al., 2014). Further, a recent study with children with language disorder has also found near transfer effects in both verbal and visuospatial working memory (Acosta et al., 2019). Consequently, the hypotheses of the current study are that following a 6 week face to face adaptive working memory training programme, children with language disorder will produce near transfer effects in both verbal and visuospatial working memory tasks and that these will be maintained at a 9 month follow-up.

6.2 Methods

6.2.1 Participants

The same procedure and participants were adopted as described in chapter 3, please see section 3.2 for a full overview. For a complete overview of inclusion criteria please see section 3.2.1. Those who met the inclusion criteria were randomly allocated to either a working memory intervention group or an active control group. Following random allocation, each group's pre-intervention scores for all near transfer tasks were inspected for normality. To investigate whether there were significant group differences on near transfer working memory tasks, independent t-tests or Mann-Whitney U tests (depending on whether data were parametric or non-parametric) were conducted between trained and active control children for all near transfer tasks at pre-intervention time point (see table 6.1 below).

Table 6.1 Mean raw scores, standard deviation, normal distribution, independent t-test or Mann-Whitney U results for all near transfer tasks at pre-intervention

	Mean	***SD***	***Normal Distribution***	***T-test Result***	***Mann-Whitney U***
Digit Recall					
Trained Group (n=24)	25.83	6.06	*w*(24)=.90, *p*=**.02**		*u*=212.5, *p*=.18
Active Control (n=23)	24.57	4.33	*w*(23)=.97, *p*=.77		
Word List Recall					
Trained Group (n=24)	18.83	3.64	*w*(24)=.93, *p*=.09	*t*(45) = 1.90, *p*=.06	
Active Control (n=23)	17.00	2.91	*w*(23)=.97, *p*=.74		
Block Recall					
Trained Group (n=24)	20.67	6.40	*w*(24)=.99, *p*=.99		*u*=272.5, *p*=.94
Active Control (n=23)	20.00	5.83	*w*(23)=.91, *p*=**.04**		
Counting Recall					
Trained Group (n=24)	16.58	5.49	*w*(24)=.95, *p*=.29	*t*(45)=.80, *p*=.48	
Active Control (n=23)	15.22	6.18	*w*(23)=.95, *p*=.26		

Table 6.1 (continued)

	Mean	*SD*	*Normal Distribution*	*T-test Result*	*Mann-Whitney U*
Backward Digit Recall					
Trained Group (n=24)	11.21	5.66	*w*(24)=.93, *p*=.09		*u*=272.5, *p*=.94
Active Control (n=23)	9.74	5.92	*w*(23)=.89, *p*=.**02**		
Pattern Span					
Trained Group (n=24)	17.29	3.64	*w*(24)=.90, *p*=.**02**		*u*=260.5, *p*=.74
Active Control (n=23)	16.30	2.91	*w*(23)=.95, *p*=.28		

6.2.2 Measures and Procedures

As outlined in section 3.4.2, near transfer measures were administered to each child at pre-intervention, immediately post-intervention and at a 9 month follow-up using the following measures:

i. Digit recall
ii. Word list recall
iii. Block recall
iv. Counting recall
v. Backward digit recall
vi. Pattern Span

For detailed procedural information of each near transfer measure, please refer to section 3.4.2, detailing the training materials and procedures required to administer each near transfer working memory task.

6.2.3 Statistical Analysis

Hierarchical multiple regressions were conducted at two time points, one immediately post-working memory intervention and one at a 9 month follow-up. The regression analyses were used to detect whether group, trained vs active control, was a significant predictor of task performance in each near transfer measure at both post-intervention and at 9 month follow-up. Raw scores were used in the analyses, with age controlled, as a number of children within the sample did not score highly enough to attain a standardised score so there would have been missing data if standardised scores were used. However, as these scores were uncorrected for age, in each regression age was used as an additional 'control' predictor.

There were no significant group differences between the two dependent variables at pre-intervention (see Table 6.1), so if the intervention was effective, group would be a significant predictor of near transfer effect measures at post-intervention and 9 month follow-up.

Hierarchical multiple regression was chosen as a method of statistical analysis so that the influence of pre-intervention performance, and age in months at pre-intervention time point, could be taken account of in the analysis. In the regression analyses the pre-intervention score of the dependent variable was entered at Step 1 (e.g. digit recall pre-intervention score) alongside age in months at the pre-intervention time point, and at step 2 group was entered as a dummy variable.

The regression analyses were carried out separately for each near transfer measure, at post-intervention and 9 month follow-up. This meant that 12 separate regression analyses were conducted. Bonferroni corrections were made to the significance levels from these analyses with a significant *p*-value of 0.004 required (0.05 ÷12). For all regression analyses key statistical checks were carried out (Durbin-Watson, tolerance and VIF statistics, Cook's and Mahalanobis distances, standardised DFBeta, leverage values, plots of standardised residuals, predicted standardised values, standardised residuals and partial plots).

6.3 Results

The following section contains the descriptive statistics (mean, standard deviation and range, and distribution of changes in scores across time points), and hierarchical regression analysis in relation to near transfer effects at immediate post-intervention, and at 9 month follow-up. There is a section relating to each of the six near transfer measures of digit recall, word list recall, block recall, counting recall, backward digit recall and pattern span. In general terms, the descriptive statistics show there were larger changes in near-transfer for the trained group, and that group was a significant predictor of near transfer at both immediate post-intervention and 9 month follow-up.

6.3.1 Digit Recall Results

6.3.1.1 Digit Recall, Descriptive Statistics

Table 6.2 below shows the descriptive statistics for the raw scores of the digit recall task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at a 9 month follow-up.

Table 6.2 Mean, standard deviation and range of digit recall raw scores for both trained and active control groups

		Mean	*SD*	*Range*
Digit Recall	Pre (n=24)	25.8	6.1	9-34
Trained Group	Post (n=24)	31.1	6.0	20-41
	Follow Up (n=24)	32.0	6.2	19-43
Digit Recall	Pre (n=23)	24.6	4.3	15-32
Active Control	Post (n=23)	25.3	4.3	16-35
	Follow Up (n=23)	24.8	4.5	15-34

Table 6.2 above shows that the trained group of children had a mean digit recall score increase from pre-intervention to 9 month follow-up of 6.2 compared to a mean increase of just 0.2 in the active control group.

The percentage of improvements, decreases or no change in digit recall scores for both the trained group (n=24) and active control group (n=23) between pre-intervention and 9 month follow-up, pre-intervention and post-intervention, and post-intervention and 9 month follow-up time points, are shown below in table 6.3.

Table 6.3: Percentage of participants who made an improvement, a decrease or no change in digit recall scores across all time points

		Pre-Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	95.8%	83.3%	58.3%
	Decreased	0%	8.3%	25%
	No Change	4.2%	8.3%	16.6%
Active Control (n=23)	**Improved**	30.4%	56.5%	26.1%
	Decreased	21.7%	17.4%	60.9%
	No Change	47.8%	26.1%	13%

Table 6.3 shows that 95.8% of trained children improved in digit recall scores between pre-intervention and the 9 month follow-up time point compared to 30.4% of active controls. Between pre and post-intervention 83.3% of trained children improved in digit recall compared to 56.5% of active controls. Between post-intervention and 9 month

follow-up, 39.1% of active controls improved or maintained their post-intervention digit recall scores, compared to 75% of the trained group.

6.3.1.2 Digit Recall (Post-Intervention) Hierarchical Regression Analysis

As shown in Table 6.4 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention digit recall scores after controlling for pre-intervention score and age. Step one of the model involved entering pre-intervention scores and age in months as 'control' variables. These two variables accounted for 40.6% of the variance in post-intervention digit recall scores, however, examination of the standardised beta-values indicated that only the pre-intervention score of digit recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 17.9% of the variance in post-intervention digit recall scores. The final model at step two was significant ($F(3,43) = 20.24$, $p<.001$) and predicted 58.5% of the digit recall score variance (Adj. R^2 .556).

Table 6.4 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, digit recall

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	8.60 (-1.35-18.55)	4.94		*p*=.09
Pre-Intervention	.70 (.42-.99)	.14	.62	*p*<.001
Age	.02 (-.08- .12)	.05	.05	*p*=.69
Step 2				
Constant	17.03 (7.74 -26.32)	4.61		*p*=.001
Pre-Intervention	.64 (.39 -.88)	.12	.56	*p*<.001
Age	.03 (-.06-.11)	.04	.07	*p*=.51
Group	-5.06 (-7.42 - -2.70)	1.17	-.43	*p*<.001
Step 1 R^2=.41, p=<.001	**Step 2 ΔR^2= .18, p<.001**			

6.3.1.3 Digit Recall (9 Month Follow-Up) Hierarchical Regression Analysis

As shown in Table 6.5 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up digit recall scores after controlling for pre-intervention score and age. Step one of the model involved entering pre-intervention scores and age in months as ‘control’ variables. These two variables accounted for 45% of the variance in 9 month follow-up digit recall scores, however, examination of the standardised beta-values indicated that only the pre-intervention score of digit recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (*ΔR^2)* showing that this variable was a significant predictor accounting for a further 24% of the variance in 9 month follow-up digit recall scores. The final model at step two was significant ($F(3,43) = 31.29$, $p<.001$) and predicted 69% of the digit recall score variance (Adj. R^2 .664).

Table 6.5 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, digit recall

Pre-Intervention – follow-up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	6.67 (-3.72-17.05)	5.15		*p*=.20
Pre-Intervention	.81 (.52-1.11)	.15	.66	*p*<.001
Age	.01 (-.08- -.12)	.05	.03	*p*=.79
Step 2				
Constant	17.15 (8.37 -25.92)	4.35		*p*<.001
Pre-Intervention	.73 (.50-.96)	.11	.59	*p*<.001
Age	.02 (-.06-.11)	.04	.06	*p*=.55
Group	-6.29 (-8.52 - -4.06)	1.11	-.49	*p*<.001

Step 1 R^2=.45, p=<.001 ***Step 2 ΔR^2= .24, p<.001***

6.3.2 Word List Recall Results

6.3.2.1 Word List Recall, Descriptive Statistics

Table 6.6 below shows the descriptive statistics for the raw scores of the word list recall task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at 9 month follow-up.

Table 6.6: Mean raw scores, standard deviation and range of word list recall scores for both trained and active control groups

		Mean	***SD***	***Range***
Word List Recall	Pre (n=24)	18.8	3.6	13-24
Trained Group	Post (n=24)	22.7	4.1	14-33
	Follow Up (n=24)	23.3	4.2	16-31
Word List Recall	Pre (n=23)	17.0	2.9	11-22
Active Control	Post (n=23)	17.1	3.1	11-23
	Follow Up (n=23)	17.2	2.7	12-21

Table 6.6 above shows that the trained working memory intervention group had a mean word list recall score increase of 4.5 from pre-intervention to the 9 month follow-up, whereas the active control group had an increase of 0.2 in mean word list recall scores.

Table 6.7 below will look at the percentage of trained (n=24) and active control (n=23) participants who made an improvement, decrease or no change in word list recall scores throughout the three time points of the intervention.

Table 6.7 Percentage of participants who made an improvement, a decrease or no change in word list recall scores across all time points

		Pre- Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	95.8%	91.7%	54.2%
	Decreased	0%	8.3%	25%
	No Change	4.2%	0%	20.8%
Active Control (n=23)	**Improved**	30.4%	43.5%	43.5%
	Decreased	17.4%	30.4%	43.5%
	No Change	52.2%	26.1%	13%

Table 6.7 shows that 95.8% of trained children improved in word list recall scores between pre-intervention and the 9 month follow-up compared to 30.4% of active controls. Between pre and post-intervention 91.7% of trained children improved in word list recall compared to 43.5% of active controls. Between post-intervention and 9 month follow-up, 75% of all trained participants further improved or maintained their post-intervention word list recall scores, compared to 56.5% of the active controls. The level of decrease in word list recall scores is higher in active controls than trained participants at each time point and interestingly the percentage of participants who show no change at all in word list recall from pre-intervention to the 9 month follow-up is 52.2% for active controls, compared to 4.2% in the trained participant group.

6.3.2.2 Word List Recall (Post-Intervention) Hierarchical Regression Analysis

As shown in Table 6.8 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention word list recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 51.1% of the variance in post-intervention word list recall scores, however, only the pre-intervention score of word list recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 19.2% of the variance in post-intervention word list recall scores. The final model at step two was significant ($F(3,43) = 33.81$, $p<.001$) and predicted 70.3% of the word list recall score variance (Adj. R^2 .682).

Table 6.8 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, word list recall

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	1.22 (-6.35-8.79)	3.75		*p*=.75
Pre-Intervention	.94 (.65-1.24)	.14	.70	*p*<.001
Age	-.02 (-.05- .09)	.03	.06	*p*=.57
Step 2				
Constant	9.78 (2.96-16.59)	3.38		*p*=.006
Pre-Intervention	.77 (.53-1.01)	.12	.57	*p*<.001
Age	.03 (-.03-.08)	.03	.08	*p*=.33
Group	-4.12 (-5.70 - -2.54)	.78	-.46	*p*<.001

Step 1 R^2=.51, p=<.001 ***Step 2 ΔR^2= .19, p<.001***

6.3.2.3 Word List Recall (9 Month Follow-Up) Hierarchical Regression Analysis

As shown in Table 6.9 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up word list recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 53.7% of the variance in 9 month follow-up word list recall scores, however, only the pre-intervention score of word list recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 23.2% of the variance in 9 month follow-up word list recall scores. The final model at step two was significant ($F(3,43) = 47.76$, $p<.001$) and predicted 76.9% of the word list recall score variance (Adj. R^2 .753).

Table 6.9 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, word list recall

Pre-Intervention – Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	2.27 (-5.19-9.73)	3.70		*p*=.54
Pre-Intervention	.99 (.71-1.29)	.14	.73	*p*<.001
Age	.001 (-.06- .07)	.03	.004	*p*=.97
Step 2				
Constant	11.81 (5.73-17.89)	3.02		*p*<.001
Pre-Intervention	.81 (.59-1.02)	.11	.59	*p*<.001
Age	.01 (-.04-.06)	.02	.03	*p*=.70
Group	-4.60 (-6.01- -3.19)	.70	-.50	*p*<.001

Step 1 R^2=.54, p=<.001 ***Step 2 ΔR^2= .23, p<.001***

6.3.3 Block Recall Results

6.3.3.1 Block Recall, Descriptive Statistics

Table 6.10 below shows the descriptive statistics for the raw scores of the block recall task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post- intervention and 9 month follow-up.

Table 6.10: Mean raw scores, standard deviation and range of block recall scores for both trained and active control groups

		Mean	*SD*	*Range*
Block Recall	Pre (n=24)	20.7	6.4	6-34
Trained Group	Post (n=24)	25.9	5.9	10-36
	Follow Up (n=24)	26.4	5.6	17-36
Block Recall	Pre (n=23)	20.0	5.8	8-30
Active Control	Post (n=23)	21.1	5.6	9-31
	Follow Up (n=23)	21.3	5.0	11-28

Table 6.10 above shows that the mean block recall scores for the trained group of children across all three time points increased by 5.7, whereas for the active control group the increase was smaller with an increase of 1.3.

Table 6.11 below shows the percentage of trained (n=24) and active control (n=23) participants who made an improvement, decrease or no change in block recall scores across the three time points in the intervention.

Table 6.11 Percentage of participants who made an improvement, a decrease or no change in Block Recall Scores across all time points

		Pre to Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	83.3%	91.7%	41.7%
	Decreased	4.2%	8.3%	37.5%
	No Change	12.5%	0%	20.8%
Active Control (n=23)	**Improved**	65.2%	65.2%	34.8%
	Decreased	30.4%	30.4%	52.2%
	No Change	4.3%	4.3%	13%

Table 6.11 shows that 83.3% of trained children improved in block recall scores between pre-intervention and the 9 month follow-up compared to 65.2% of active controls. Between pre and post-intervention 91.7% of trained children improved their block recall scores, compared to 65.2% of active controls. Between post-intervention and 9 month follow-up, 62.5% of all trained participants further improved or maintained their post-intervention block recall scores, compared to 47.7% of the active controls. The level of decrease in block recall scores is considerably higher in active controls than trained participants at each time point, with an overall decrease in block recall scores from pre-intervention to 9 month follow-up being 30.4% of active controls, compared to just 4.2% of trained participants.

6.3.3.2 Block Recall (Post-Intervention) Hierarchical Regression Analysis

As shown in Table 6.12 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention block recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 55.7% of the variance in post-intervention block recall scores, however, only the pre-intervention score of block recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 12.3% of the variance in post-intervention block recall scores. The final model at step two was significant ($F(3,43) = 30.37$, $p<.001$) and predicted 68% of the block recall score variance (Adj. R^2 .657).

Table 6.12 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, block recall

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	8.38 (.11-16.65)	4.11		*p*=.05
Pre-Intervention	.76 (.54-.99)	.11	.75	*p*<.001
Age	-.004 (-.09-.09)	.05	-.01	*p*=.93
Step 2				
Constant	14.89 (7.07-22.71)	3.88		*p*<.001
Pre-Intervention	.74 (.55 -.93)	.10	.73	*p*<.001
Age	-9.52 (-.08-.08)	.04	.00	*p*=.99
Group	-4.30 (-6.43- -2.16)	1.06	-.35	*p*<.001

Step 1 R^2=.56, p=<.001 ***Step 2 ΔR^2= .12, p<.001***

6.3.3.3 Block Recall (9 Month Follow-Up) Hierarchical Regression Analysis

As shown in Table 6.13 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up block recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 50.8% of the variance in 9 month follow-up block recall scores, however, only the pre-intervention score of block recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 16.6% of the variance in 9 month follow-up block recall scores. The final model at step two was significant ($F(3,43) = 29.61$, $p<.001$) and predicted 67.4% of the block recall score variance (Adj. R^2 .651).

Table 6.13 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, block recall

Pre-Intervention – Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	11.75 (3.54-19.96)	4.08		*p*=.006
Pre-Intervention	.70 (.48-.93)	.11	.73	*p*<.001
Age	-.02 (-.11-.07)	.05	-.06	*p*=.62
Step 2				
Constant	18.88 (11.44-26.31)	3.69		*p*<.001
Pre-Intervention	.68 (.49-.86)	.09	.71	*p*<.001
Age	-.02 (-.09-.06)	.04	-.05	*p*=.62
Group	-4.70 (-6.73 - -2.68)	1.01	-.41	*p*<.001

Step 1 R^2=.51, p=<.001 ***Step 2 ΔR^2= .17, p<.001***

6.3.4 Counting Recall Results

6.3.4.1 Counting Recall Descriptive Statistics

Table 6.14 below shows the descriptive statistics for the raw scores of the counting recall task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at a 9 month follow-up.

Table 6.14. Mean raw scores, standard deviation and range of counting recall scores for both trained and active control groups

		Mean	***SD***	***Range***
Counting Recall	Pre (n=24)	16.6	5.5	8-29
Trained Group	Post (n=24)	23.0	7.0	8-41
	Follow Up (n=24)	22.3	6.8	11-41
Counting Recall	Pre (n=23)	15.2	6.2	7-28
Active Control	Post (n=23)	16.3	6.7	6-29
	Follow Up (n=23)	16.8	6.6	7-31

Table 6.14 above shows that the mean counting recall scores for the trained group of children from pre-intervention to 9 month follow-up increased by 5.7, whereas for the active control group the increase was smaller with an increase of 1.6.

Table 6.15 below shows the percentage of trained (n=24) and active control (n=23) participants who made an improvement, decrease or no change in counting recall scores across the three time points in the intervention.

Table 6.15 Percentage of participants who made an improvement, a decrease or no change in counting recall scores across all time points

		Pre-Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	95.8%	87.5%	50%
	Decreased	4.2%	0%	41.7%
	No Change	0%	12.5%	8.3%
Active Control (n=23)	**Improved**	56.5%	43.5%	56.5%
	Decreased	30.4%	30.4%	26.1%
	No Change	13%	26.1%	17.4%

Table 6.15 indicates that 95.8% of trained children improved in counting recall scores between pre-intervention and the 9 month follow-up compared to 56.5% of active controls. Between pre and post-intervention 87.5% of trained children improved their counting recall scores, compared to 43.5% of active controls. During the period between post-intervention and 9 month follow-up, 58.3% of trained participants further improved or maintained their post-intervention counting recall scores, compared to a higher 73.9% of the active controls. Interestingly, from pre-intervention to post-intervention no trained participants decreased their counting recall score, compared to 30.4% of active controls. However, from post-intervention to follow-up, 41.7% of trained participants decreased from their post-intervention score, compared to just 26.1% of active controls

6.3.4.2 Counting Recall (Post-Intervention) Hierarchical Regression Analysis

As shown in Table 6.16 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention counting recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 66.7% of the variance in post-intervention counting recall scores, however, only the pre-intervention score of counting recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 12.9% of the variance in post-intervention counting recall scores. The final model at step two was significant ($F(3,43) = 56.15$, $p<.001$) and predicted 79.6% of the counting recall score variance (Adj. R^2 .782).

Table 6.16: Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, counting recall

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	-2.44 (-11.17-6.29)	4.33		*p*=.58
Pre-Intervention	.96 (.70-1.23)	.13	.74	*p*<.001
Age	.07 (-.03-.17)	.05	.14	*p*=.17
Step 2				
Constant	5.47 (-2.08-13.02)	3.74		*p*=.15
Pre-Intervention	.89 (.68-1.10)	.10	.68	*p*<.001
Age	.09 (.004-.17)	.04	.17	*p*=.04
Group	-5.44 (-7.54- -3.34)	1.04	-.36	*p*<.001

Step 1 R^2=.67, *p*=<.001 **Step 2 ΔR^2= .13, *p<.001***

6.3.4.3 Counting Recall (9 Month Follow-Up) Hierarchical Regression Analysis

As shown in Table 6.17 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up counting recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 62.2% of the variance in 9 month follow-up counting recall scores, however, only the pre-intervention score of counting recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 12.9% of the variance in 9 month follow-up counting recall scores. The final model at step two was significant ($F(3,43) = 43.25$, $p<.001$) and predicted 75.1% of the counting recall score variance (Adj. R^2 .734).

Table 6.17 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, counting recall

Pre-Intervention – Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	1.33 (-7.82- 10.47)	4.54		*p*=.77
Pre-Intervention	.96 (.68-1.23)	.14	.75	*p*<.001
Age	-.04 (-.07-.14)	.05	.07	*p*=.49
Step 2				
Constant	9.09 (.89-17.29)	4.07		*p*=.03
Pre-Intervention	.89 (.66-1.11)	.11	.69	*p*<.001
Age	.05 (-.04-.14)	.04	.10	*p*=.25
Group	-5.34 (-7.62- -3.06)	1.13	-.36	*p*<.001

Step 1 R^2=.62, p=<.001 ***Step 2 ΔR^2= .13, p<.001***

6.3.5 Backward Digit Recall Results

6.3.5.1 Backward Digit Recall, Descriptive Statistics

Table 6.18 below shows the descriptive statistics for the raw scores of the backward digit recall task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at a 9 month follow-up.

Table 6.18 Mean raw scores, standard deviation and range of backward digit recall scores for both trained and active control groups

		Mean	*SD*	*Range*
Backward Digit Recall	Pre (n=24)	11.2	5.7	0-22
Trained Group	Post (n=24)	14.1	5.7	0-25
	Follow Up (n=24)	13.8	5.7	0-25
Backward Digit Recall	Pre (n=23)	9.7	5.9	0-19
Active Control	Post (n=17)	9.5	5.8	0-18
	Follow Up (n=17)	9.6	4.9	0-19

Table 6.18 above shows that the mean score for backward digit recall in trained children from pre-intervention to the 9 month follow-up time point increased by 2.6, whereas for active controls, mean scores decreased by 0.1 over the same time period.

Table 6.19 below shows the percentage of trained (n=24) and active control (n=23) participants who made an improvement, decrease or no change in backward digit scores across the three time points in the intervention.

Table 6.19 Percentage of participants who made an improvement, a decrease or no change in backward digit recall scores across all time points

		Pre-Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	62.5%	58.3%	37.5%
	Decreased	20.8%	20.8%	37.5%
	No Change	16.7%	20.8%	25%
Active Control (n=23)	**Improved**	26.1%	34.8%	34.8%
	Decreased	43.5%	39.1%	43.5%
	No Change	30.4%	26.1%	21.7%

Table 6.19 shows that 62.5% of trained children improved in backward digit recall scores between pre-intervention and the 9 month follow-up compared to 26.1% of active controls. Trained participants improved more between pre-post intervention time points than active controls (trained group improvement was 58.3% compared to 34.8% for active controls) and trained participants also showed a decrease of backward digit recall scores less frequently during this time period (20.8% of the trained group decreased in score, compared to 39.1% of the active controls). However, when looking at percentages of improvement, decrease and no change of backward digit recall scores from post-intervention to follow-up, the percentages between the two groups are relatively similar, indicating that the differences may have levelled out.

6.3.5.2 Backward Digit Recall (Post-Intervention), Hierarchical Regression Analysis

As shown in Table 6.20 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention backward digit recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age in months accounted for 64.5% of the variance in post-intervention backward digit recall scores, however, only the pre-intervention score of backward digit recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing this variable to be a significant predictor accounting for a further 7.5% of the variance in post-intervention backward digit recall scores. The final model ($F(3,43) = 36.87$, $p<.001$) at step two predicted 72% of the backward digit recall score variance (Adj. R^2 .701).

Table 6.20. Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, backward digit recall

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	3.15 (-4.38-10.68)	3.74		*p*=.40
Pre-Intervention	.86 (.64-1.08)	.11	.81	*p*<.001
Age	-.003 (-.09-.08)	.04	-.008	*p*=.94
Step 2				
Constant	7.75 (.46-15.04)	3.61		*p*=.04
Pre-Intervention	.81 (.61-1.01)	.10	.76	*p*<.001
Age	.01 (-.07- .09)	.04	.02	*p*=.86
Group	-3.38 (-5.38- -1.38)	.99	-.28	*p*=.001
Step 1 R^2=.65, *p*=<.001	**Step 2 ΔR^2= .08, *p*<.001**			

6.3.5.3 Backward Digit Recall (9 Month Follow-Up), Hierarchical Regression Analysis

As shown in Table 6.21 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up backward digit recall scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 71% of the variance in 9 month follow-up backward digit recall scores, however, only the pre-intervention score of backward digit recall was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 7.8% of the variance in 9 month follow-up backward digit recall scores. The final model at step two was significant ($F(3,43) = 53.34$, $p<.001$) and predicted 78.8% of the backward digit recall score variance (Adj. R^2 .773).

Table 6.21 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, backward digit recall

Pre-Intervention – Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	.87 (-5.42-7.16)	3.12		*p*=.78
Pre-Intervention	.79 (.61-.98)	.09	.81	*p*<.001
Age	.03 (-.05-.09)	.04	.07	*p*=.47
Step 2				
Constant	5.20 (-.67-11.06)	2.91		*p*=.08
Pre-Intervention	.75 (.58-.91)	.08	.76	*p*<.001
Age	.04 (-.03-.10)	.03	.09	*p*=.26
Group	-3.17 (-4.78- -1.57)	.80	-.28	*p*<.001

Step 1 R^2=.71, p=<.001 ***Step 2 ΔR^2= .08, p<.001***

6.3.6 Pattern Span Results

6.3.6.1 Pattern Span, Descriptive Statistics

Table 6.22 below shows the descriptive statistics for the span levels of the pattern span task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at a 9 month follow-up.

Table 6.22 Mean of raw scores, standard deviation and range of pattern span levels for both trained and active control groups

		Mean	*SD*	*Range*
Pattern Span	Pre (n=24)	3.2	1.0	1-6
Trained Group	Post (n=24)	4.1	1.0	2-6
	Follow Up (n=24)	4.2	1.0	2-6
Pattern Span	Pre (n=23)	3.1	1.0	1-5
Active Control	Post (n=23)	3.2	0.9	1-4
	Follow Up (n=23)	3.3	1.0	1-5

Table 6.22 above shows that the mean span levels for the pattern span task in trained children from pre-intervention to the 9 month follow-up, increased on average by 1.0, whereas the span levels in the active control group increased by 0.2 over the same time period.

Table 6.23 below shows the percentage of trained (n=24) and active control (n=23) participants who made an improvement, decrease or no change in pattern span scores across the three time points in the intervention.

Table 6.23 Percentage of participants who made an improvement, a decrease or no change in Pattern Span Scores across all time points

		Pre to Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	95.8%	87.5%	95.8%
	Decreased	0%	0%	0%
	No Change	4.2%	12.5%	4.2%
Active Control (n=23)	**Improved**	69.6%	52.2%	69.6%
	Decreased	26.1%	17.4%	26.1%
	No Change	4.3%	30.4%	4.3%

Table 6.23 shows that 95.8% of trained children improved in pattern span scores between pre-intervention and the 9 month follow-up compared to 69.6% of active controls. Between the pre and post-intervention time points 87.5% of trained children improved their pattern span, compared to 52.2% of active controls. Between post-intervention and 9 month follow-up, 100% of trained participants either further improved or maintained their post-intervention pattern span level, compared to 73.9% of the active controls. No trained participants were shown to decrease in pattern span at any time point throughout the current study, whereas pattern span decreased in 26.1% of active controls between pre-intervention and 9 month follow-up, 17.4% between pre and post-intervention and 26.1% between the post-intervention and follow-up time point.

6.3.6.2 Pattern Span (Post-Intervention), Hierarchical Regression Analysis

As shown in table 6.24 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention pattern span levels after controlling for pre-intervention span and age. At Step one, pre-intervention span level and age in months accounted for 47.3% of the variance in post-intervention pattern span, however, only the pre-intervention pattern span was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing this variable to be a significant predictor accounting for a further 15.7% of the variance in post-intervention pattern span. The final model ($F(3,43) = 24.41$, $p<.001$) at step two predicted 63% of the pattern span variance (Adj. R^2 .604).

Table 6.24 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, pattern span

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	1.19 (-.28-2.66)	.73		*p*=.11
Pre-Intervention	.65 (.42-.89)	.12	.66	*p*<.001
Age	.004 (-.01-.02)	.008	.06	*p*=.59
Step 2				
Constant	2.39 (1.02-3.75)	.68		*p*=.001
Pre-Intervention	.63 (.43-.82)	.10	.64	*p*<.001
Age	.01 (-.01-.02)	.01	.07	*p*=.46
Group	-7.90 (-1.16- -.42)	.19	-.40	*p*=.001

Step 1 R^2=.47, p=<.001 ***Step 2 ΔR^2= .16, p<.001***

6.3.6.3 Pattern Span (9 Month Follow-Up), Hierarchical Regression Analysis

As shown in table 6.25 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up pattern span levels after controlling for pre-intervention span and age. At Step one, pre-intervention span level and age in months accounted for 51.6% of the variance in 9 month follow-up pattern span, however, only the pre-intervention pattern span was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing this variable to be a significant predictor accounting for a further 12.6% of the variance in 9 month follow-up pattern span. The final model ($F(3,43) = 25.69$, $p<.001$) at step two predicted 64.2% of the pattern span variance (Adj. R^2 .617).

Table 6.25 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, pattern span

Pre-Intervention – Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	.92 (-.69-2.54)	.80		*p*=.26
Pre-Intervention	.80 (.54-1.05)	.13	.70	*p*<.001
Age	.003 (-.02-.02)	.01	.04	*p*=.73
Step 2				
Constant	2.16 (.61-3.70)	.77		*p*=.007
Pre-Intervention	.77 (.54-.99)	.11	.68	*p*<.001
Age	.004 (-.01-.02)	.01	.05	*p*=.63
Group	-.81 (-1.23- -.39)	.21	-.36	*p*<.001

Step 1 R^2=.52, p=<.001 ***Step 2 ΔR^2= .13, p<.001***

6.4. Discussion

This chapter aimed to determine whether group (trained vs active control) was a significant predictor of performance in near transfer measures, digit recall, word list recall, block recall, counting recall, backward digit recall and pattern span, immediately post-intervention and at a 9 month follow-up. Hierarchical multiple regressions show that group membership (trained vs active control) was a significant predictor of all six near transfer measures immediately post-intervention, and at a 9 month follow-up. In addition to the regression analyses, descriptive statistics showed marked differences between the two groups in the post-intervention and 9 month follow-up scores. Between the pre-intervention and the 9 month follow –up, for all the six near transfer assessments there was an appreciable increase in scores for the trained group, and often the untrained group had a very small increase. For a summary of all near transfer results, please refer to table 6.26 below.

Table 6.26: Near transfer measure results summary

Task and Group	Pre-Post intervention		Pre-intervention to 9-month follow-up	
	Mean Increase of Scores	**Variance explained**	**Mean Increase of Scores**	**Variance explained**
Digit Recall				
Trained	5.3	β =-.43 *p*<.001	6.2	β =-.49 *p*<.001
Active control	0.7		0.2	
Word List Recall				
Trained	3.9	β = -.46 *p*<.001	4.5	β = -.50 *p*<.001
Active control	0.1		0.2	
Block Recall				
Trained	5.2	β = -.35 *p*<.001	5.7	β = -.41, *p*<.001
Active control	1.1		1.3	
Counting Recall				
Trained	6.4	β = -.36 *p*<.001	5.7	β = -.36 *p*<.001
Active control	1.1		1.6	
Backward Digit Recall				
Trained	2.9	β = -.28 *p*=.001	2.6	β = -.28 *p*<.001
Active control	-0.2		-0.1	
Pattern Span				
Trained	0.9	β = -.40 *p*=.001	1.0	β = -.36 *p*<.001
Active control	0.1		0.2	

Table 6.26 above shows that as predicted, near transfer effects were evident in both verbal working memory and visuospatial working memory tasks following a face to face adaptive working memory intervention that focused on two executive-loaded working memory tasks within the verbal and visuospatial domains. Group was not only a significant predictor of performance in all near transfer tasks immediately post-intervention but also at a 9 month follow-up time point, indicating training effects were also maintained.

As discussed in section 6.1, researchers have often acknowledged the benefits of working memory training to other parts of the working memory system (Sala & Gobet, 2017; chapter 2). However, in previous working memory research, the maintenance of near transfer effects following intervention has often been questioned (Melby-Lervag & Hulme, 2013 & Melby-Lervag et al., 2016). Previous meta-analysis studies, conducted on computer-based working memory interventions, have found evidence of verbal working memory tasks yielding short-term near transfer effects that were not maintained when reassessed after a delay of a few months post-intervention (Melby-Lervag & Hulme, 2013 & Melby-Lervag et al., 2016). Visuospatial working memory has also been found to have moderate improvements, with some evidence of maintenance at follow-up, however as the outcome tasks often shared features with the working memory trained tasks, this has led to doubt regarding the true nature of these 'near transfer' effects (Melby-Lervag et al., 2016).

Although the current study supports previous literature, which often shows immediate evidence of near transfer effects following working memory interventions, the current study does not support the common finding regarding their short-term maintenance (Melby-Lervag & Hulme, 2013 & Melby-Lervag et al., 2016), as the six near transfer effects in the current study were all still evident at a 9 month follow-up.

As discussed previously similar to the current study, Henry et al.,'s (2014) face to face working memory training intervention trained two executive-loaded working memory tasks within the verbal and visuospatial domain, with typically developing children, and found evidence of near transfer effects in word list recall and counting recall, that were still evident at a 6 month follow-up. Traditionally, computer-based working memory training approaches have often found limited evidence of sustained near transfer effects (Melby-Lervag & Hulme, 2013 & Melby-Lervag et al., 2016). However, both Henry et al.'s (2014) working memory intervention and the current study, used a face to face working memory training approach. When considering why the current study found direct and near far transfer effects that not only were evident immediately post intervention but also at a 9 month follow-up, the use of a non-computerised training method must be considered. As explained above, computerised training traditionally has found limited evidence of sustained near transfer effects, which suggests that the face to face approach, and the pivotal ingredient of human interaction that was involved in the current study, were key to its positive findings.

Supporting this viewpoint, Melby-Lervag & Hulme (2013) discuss the need for future working memory training to move away from the traditional computer-based approach, due to repeated evidence of limited transfer to other areas of working memory and cognitive skill. The current study provides support for this change in working memory research, showing that through adopting a non-computerised working memory training approach, improvements to other related areas of working memory can be made on a large scale and within a wide range of tasks. Working memory ability is fundamental for tasks such as reading comprehension, mathematics/arithmetic, planning of thoughts and actions and maintaining focused behaviour (Holmes, 2012), it is also thought to be closely related to educational attainment (Henry, 2012). Therefore, delivering an executive-loaded working memory intervention that improves performance in a number of working memory measures, which are maintained over a period of 9 months, has the potential to convey an educational benefit.

Overall this chapter has shown that group membership (trained vs active control) is a significant predictor of performance in all six near transfer measures, digit recall, word list recall, block recall, counting recall, backward digit span and pattern span, both immediately post-intervention and for a further 9 month period. These findings support a number of previous studies which have reported immediate near transfer effects, (Melby-Lervag & Hulme, 2013., & Melby-Lervag et al., 2016) and although limited, previous research which has reported near transfer effects that were sustained over several months (Henry et al., 2014).

Although near transfer effects are a positive indicator of the effectiveness of a working memory intervention, gaining far transfer effects, which measure transfer to other cognitive skills and academic ability such as reading comprehension or mathematical tasks, are the main goal of working memory intervention studies. Chapter 7 of this thesis focuses on far transfer effects, and will assess whether group (trained vs active control) is a significant predictor of performance in reading accuracy, reading comprehension, receptive grammar or sentence comprehension.

Chapter 7

Effects of Working Memory Training on Far Transfer Performance

7.1 Introduction

The effectiveness of working memory training is often determined by gains in other areas of working memory (near transfer effects) and other areas of academic and cognitive skills (far transfer). The strongest evidence for the general benefits of working memory training derives from far transfer effects, gains in ability that are not directly related to working memory trained tasks in structure or content (Henry et al., 2014).

Limited evidence of far transfer effects following working memory training has led to pessimism concerning the effectiveness of working memory interventions, with meta-analysis studies commonly finding little or no evidence of far transfer effects (Melby-Lervag & Hulme, 2013; Melby-Lervag et al., 2016; Sala & Gobet, 2017; Schwaighofer et al., 2015; see Chapter 2). Melby-Lervag and Hulme (2013) found no evidence of far transfer effects to nonverbal or verbal ability, inhibitory processes of attention, word decoding or arithmetic. Schwaighofer et al. (2015) found only limited and very small gains in non-verbal and verbal ability, but these were not maintained at follow-up. Weak evidence of far transfer from working memory training to reading comprehension immediately post-intervention and arithmetic at follow-up has been reported (Melby-Lervag et al., 2016), but these findings were questioned due to a decrease between pre-intervention and post-intervention scores for the control group.

Working memory training has traditionally adopted a computer-based approach which has repeatedly found minimal evidence of far transfer effects (Rowe et al., 2019a).

Nevertheless, non-computerised working memory training, with strategy instruction, has been found to produce far transfer effects to visual patterns and arithmetic (Witt, 2011). Furthermore, a face-to-face adaptive working memory intervention that focussed on training two executive-loaded working memory tasks in both the verbal and visuospatial domains found evidence of far transfer in typically developing children to reading comprehension at a 12-month follow-up (Henry et al., 2014).

There are reasons why comprehension might be expected to benefit from working memory training, especially training focussing on the executive system. A verbal executive-loaded working memory intervention could be beneficial to comprehension as Diamond (2013) argues that working memory is critical in making sense of information (such as speech) that unfolds over time. Oral comprehension will therefore require this skill in the verbal domain. Although non-verbal executive-loaded working memory may initially seem irrelevant, at least one model predicts it would benefit reading comprehension. Reading comprehension is seen as dependent upon an individual's ability to construct a coherent meaning-based mental representation of a situation described in text. The 'Situation Model' (De Koning & Van der Schoot, 2013) argues that through using senses, particularly vision, to create non-verbal representations of the concepts within a piece of text, this is thought to increase the likelihood of understanding (De Koning & Van der Schoot, 2013). Training visuospatial ability may therefore have a direct impact on the ability to form 'Situation Models' and improve reading comprehension ability. An executive-loaded working memory intervention that targets both the verbal and visuospatial domain, as in the current study, could therefore be seen as beneficial.

The current chapter of this thesis aims to answer the following research questions:

i. Do children with language disorder show immediate post-intervention far transfer effects in language outcome measures following executive-loaded working memory training in the verbal and visuospatial domains?
ii. Can children with language disorder maintain any such far transfer effects in language outcome measures nine months after the executive-loaded working memory training?

As noted above, computer-based working memory training interventions have shown limited evidence of far transfer effects, but more recent intervention studies which have focused on non-computerised interventions, as in the current study, have produced more positive outcomes (Rowe et al., 2019a). In particular, a similar intervention study with typically developing children obtained far transfer effects in reading comprehension at a 12 month follow-up (Henry et al., 2014). Consequently, the hypotheses in the current study are that following a six-week face-to-face adaptive working memory training programme, children with language disorder will show far transfer effects in reading comprehension and these will be maintained at a 9 month follow-up.

7.2 Methods

7.2.1 Participants

Participants and procedures for the current study were adopted as described fully in chapter 3. For a detailed overview of the participants involved, please refer to section 3.2. Please see section 3.2.1 for inclusion criteria, section 3.2.2 for random allocation procedures and section 3.2.3 for participant screening procedures.

Once randomly allocated, each group's pre-intervention scores for all far transfer tasks were inspected for normality. To investigate whether there were significant group differences on far transfer working memory tasks, independent t-tests or Mann-Whitney U tests (depending on whether data were parametric or non-parametric) were conducted between trained and active control children for all far transfer tasks at pre-intervention time point. As can be seen in Table 7.1, Mann-Whitney U tests revealed significant differences between trained and active control groups in both reading accuracy and reading comprehension scores at pre-intervention time point. Therefore, if group (trained vs active control) is found to be a significant predictor of performance in either reading accuracy or comprehension at either time point, results will need to be interpreted with caution.

Table 7.1 Mean, standard deviation, normal distribution, independent t-test or Mann-Whitney U results for all far transfer tasks at pre-intervention

	Mean	***SD***	***Normal Distribution***	***T-test Result***	***Mann-Whitney U***
Reading Accuracy					
Trained Group (n=24)	24.42	19.02	*w*(24)=.93, *p*=.08		*u*=179.5, **p=.04**
Active Control (n=23)	13.65	13.91	*w*(23)=.87, *p*=.01		
Reading Comprehension					
Trained Group (n=24)	7.42	4.85	*w*(24)=.93, *p*=.09		*u*=165.5, **p=.02**
Active Control (n=23)	4.22	4.18	*w*(23)=.85, *p*=.002		
Receptive Grammar					
Trained Group (n=24)	8.50	3.58	*w*(24)=.98, *p*=.86	*t*(45)=.1.62, *p*=.11	
Active Control (n=23)	6.83	3.52	*w*(23)=.92, *p*=.06		
Sentence Comprehension					
Trained Group (n=24)	21.00	5.85	*w*(24)=.94, *p*=.14	*t*(45)=.1.17, *p*=.25	
Active Control (n=23)	18.91	6.42	*w*(23)=.97, *p*=.58		

7.2.2 Measures and Procedures

As outlined in section 3.4.3, far transfer measures were administered to each child at pre-intervention point, immediately post-intervention and at 9 month follow-up. This enabled differences between the trained and active control groups to be investigated at all time points. For detailed procedural information on each far transfer measure, please refer to section 3.4.3, detailing the training materials and procedures required to administer each far transfer working memory task.

7.2.3 Statistical Analysis

Hierarchical multiple regressions were conducted at two time points, one immediately post working memory intervention and one at 9-month follow-up. Regression analyses were used to detect if group (trained vs active control) was a significant predictor of performance in each far transfer measure immediately post-intervention and at a 9 month follow-up. Raw scores were used in all analyses of far transfer measures as a number of children within the sample did not score highly enough to attain a standardised score. However, as these scores were uncorrected for age, in each regression age was used as an additional 'control' predictor.

As in the previous regression analyses, the relevant pre-intervention score of the dependent variable was entered at Step 1 (e.g. sentence comprehension pre-intervention score) alongside age in months at the pre-intervention time point, and at step 2 group was entered as a dummy variable in order to assess possible group differences.

As eight separate regression analyses were conducted, Bonferroni corrections were made to the significance levels from these analyses with a significant *p* value of 0.006 required

(0.05÷8). For all regression analysis key statistical checks were carried out (Durbin-Watson, tolerance and VIF statistics, Cook's and Mahalanobis distances, standardised DFBeta, leverage values, plots of standardised residuals, predicted standardised values, standardised residuals and partial plots).

7.3 Results

There is a separate section relating to each of the four far transfer measures of sentence comprehension, receptive grammar, reading accuracy and reading comprehension. Each section contains descriptive statistics (mean, standard deviation and range at all three time points, and distribution of changes in scores across time points), and hierarchical regression analysis in relation to a far transfer effect at immediate post-intervention and at 9 month follow-up.

7.3.1. Sentence Comprehension, Assessment of Comprehension and Expression 6-11 (ACE), (Adams et al., 2001)

7.3.1.1 Sentence Comprehension, Descriptive statistics

Table 7.2 below shows the descriptive statistics for the raw scores of sentence comprehension for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at 9-month follow-up.

Table 7.2 Mean, standard deviation and range of sentence comprehension raw scores for both trained and active control groups.

		Mean	*SD*	*Range*
Sentence Comprehension	Pre (n=24)	21.0	5.9	7-29
Trained Group	Post (n=24)	25.8	4.4	15-33
	Follow Up (n=24)	27.5	4.1	18-33
Sentence Comprehension	Pre (n=23)	18.9	6.4	7-29
Active Control	Post (n=23)	20.2	6.4	8-31
	Follow Up (n=23)	22.4	6.1	10-32

The percentage of improvements, decreases or no change in sentence comprehension scores for both the trained group and active control group between pre-intervention and 9 month follow-up, pre-intervention and post-intervention, and post-intervention and 9 month follow-up time points, are shown below in table 7.3.

Table 7.3 Percentage of participants who made an improvement, a decrease or no change in sentence comprehension raw scores across all time points

		Pre-Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	100%	87.5%	79.2%
	Decreased	0%	4.2%	12.5 %
	No Change	0%	8.3%	8.3%
Active Control (n=23)	**Improved**	95.7%	52.2%	87%
	Decreased	0%	26.1%	4.3%
	No Change	4.3%	21.8 %	8.7%

Table 7.3 shows that 100% of trained children improved in sentence comprehension scores between pre-intervention and the 9-month follow-up time point compared to a similar 95.7% of active controls. Between pre and post-intervention 87.5% of trained children improved in sentence comprehension compared to 52.2% of active controls. Between post-intervention and 9 month follow-up, 95.7% of active controls improved or maintained their post-intervention sentence comprehension scores, compared to 87.5% of the trained group.

7.3.1.2 Sentence Comprehension (Post-Intervention), Hierarchical Regression Analysis

As shown in Table 7.4 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention sentence comprehension scores after controlling for pre-intervention score and age. Step one of the model involved entering pre-intervention scores and age in months as 'control' variables. These two variables accounted for 68.5% of the variance in post-intervention sentence comprehension scores, however, examination of the standardised beta-values indicated that only the pre-intervention score of sentence comprehension was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing that this variable was a significant predictor accounting for a further 10.4% of the variance in post-intervention sentence comprehension scores. The final model at step two was significant ($F(3,43) = 53.61$, $p<.001$) and predicted 78.9% of the sentence comprehension score variance (Adj. R^2 .774).

Table 7.4 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, sentence comprehension

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	7.97 (1.13-14.80)	3.39		*p*=.23
Pre-Intervention	.84 (.64-1.04)	.09	.85	*p*<.001
Age	-.02 (-.10-.07)	.04	-.04	*p*=.67
Step 2				
Constant	13.81 (7.60-20.02)	3.08		*p*<.001
Pre-Intervention	.76 (.60-.93)	.08	.77	*p*<.001
Age	-.001 (-.07-.07)	.03	-.002	*p*=.98
Group	-3.98 (-5.72- -2.24)	.86	-.33	*p*<.001

Step 1 R^2=.69, *p*=<.001 **Step 2 ΔR^2= .10, *p*<.001**

7.3.1.3 Sentence Comprehension (9 Month Follow-Up), Hierarchical Regression Analysis

As shown in Table 7.5 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up sentence comprehension scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age in months accounted for 66.2% of the variance in 9 month follow-up sentence comprehension scores, however, only the pre-intervention score of sentence comprehension was a significant individual predictor. In step two, adding group produced a significant change in R^2 (ΔR^2) showing this variable to be a significant predictor accounting for a further 9.7% of the variance in 9 month follow-up sentence comprehension scores. The final model ($F(3,43) = 45.08$, $p<.001$) at step two predicted 75.9% of the sentence comprehension score variance (Adj. R^2 .742).

Table 7.5 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, sentence comprehension

Pre-Intervention – 9 Month Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	10.98 (4.34-17.61)	3.29		*p*=.002
Pre-Intervention	.77 (.58-.97)	.10	.83	*p*<.001
Age	-.02 (-.10-.07)	.04	-.04	*p*=.71
Step 2				
Constant	16.26 (10.03-22.48)	3.09		*p*<.001
Pre-Intervention	.70 (.54-.87)	.08	.76	*p*<.001
Age	.000 (-.07 -.07)	.03	.001	*p*=.99
Group	-3.56 (-5.34- -1.85)	.87	-.32	*p*<.001

Step 1 R^2=.66, *p*=<.001 **Step 2 ΔR^2= .10, *p*<.001**

7.3.2 Test for Receptive Grammar (TROG) (Bishop, 2003)

7.3.2.1 Receptive Grammar, Descriptive Statistics

Table 7.6 below shows the descriptive statistics for the raw scores of receptive grammar for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at a 9 month follow-up.

Table 7.6 Mean, standard deviation and range of receptive grammar raw scores for both trained and active control groups

		Mean	*SD*	*Range*
Receptive Grammar	Pre (n=24)	8.5	3.6	2-16
Trained Group	Post (n=24)	11.3	4.4	4-18
	Follow Up (n=24)	12.1	3.8	6-18
Receptive Grammar	Pre (n=23)	6.8	3.5	2-14
Active Control	Post (n=23)	8.4	3.7	3-14
	Follow Up (n=23)	9.4	3.5	4-16

The percentage of improvements, decreases or no change in receptive grammar scores for both the trained group (n=24) and active control group (n=23) between pre-intervention and 9-month follow-up, pre-intervention and post-intervention, and post-intervention and 9 month follow-up time points, are shown below in table 7.7.

Table 7.7 Percentage of participants who made an improvement, a decrease or no change in receptive grammar scores across all time points

		Pre-Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	95.8%	70.8%	70.8%
	Decreased	0%	2.5%	2.5 %
	No Change	4.2%	4.2%	4.2%
Active Control (n=23)	**Improved**	91.3%	65.2%	69.6%
	Decreased	0%	8.7%	21.7%
	No Change	8.7%	26.1%	8.7%

Table 7.7 shows that 95.8% of trained children improved in receptive grammar scores between pre-intervention and the 9 month follow-up time point compared to a slightly smaller 91.3% of active controls. Between pre and post-intervention 70.8% of trained children improved in receptive grammar scores compared to 65.2% of active controls. Between post-intervention and 9 month follow-up, 78.3% of active controls improved or maintained their post-intervention receptive grammar scores, compared to a slightly lower 75% of the trained group.

7.3.2.2 Receptive Grammar (Post-Intervention), Hierarchical Regression Analysis

As shown in Table 7.8 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention receptive grammar scores after controlling for pre-intervention score and age. At Step one, pre-intervention scores and age accounted for 54% of the variance in post-intervention receptive grammar scores, however, only the pre-intervention score of receptive grammar was a significant individual predictor. In step two, adding group did not produce a significant change in R^2 (*ΔR^2)* showing this variable to not be a significant predictor, accounting for only a further 2.5% of the variance in post-intervention receptive grammar scores. The final model at step two ($F(3,43) = 18.59$, $p < .001$) predicted 56.5% of the receptive grammar score variance (Adj. R^2 .534).

Table 7.8 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, receptive grammar

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	6.46 (.20-12.72)	3.11		*p*=.04
Pre-Intervention	.98 (.66-1.30)	.16	.82	*p*<.001
Age	-.04 (-.12-.04)	.04	-.15	*p*=.27
Step 2				
Constant	7.96 (1.50-14.41)	3.20		*p*=.02
Pre-Intervention	.90 (.57-1.23)	.16	.76	*p*<.001
Age	-.03 (-.11-.05)	.04	-.11	*p*=.44
Group	-1.4 (-3.21-.41)	.90	-.17	*p*=.13

Step 1 R^2=.54*, p*=<.001 ***Step 2 ΔR²= .03, p*=.13**

7.3.2.3 Receptive Grammar (9 Month Follow-Up), Hierarchical Regression Analysis

As shown in Table 7.9 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up receptive grammar scores after controlling for pre-intervention score and age. At Step one pre-intervention scores and age in months accounted for 58.1% of the variance in 9 month follow-up receptive grammar scores, however, only the pre-intervention score of receptive grammar was a significant individual predictor. In step two, adding group, did not produce a significant change in R^2 (ΔR^2) showing this variable to not be a significant predictor, accounting for only a further 3.5% of the variance in 9 month follow-up receptive grammar scores. The final model at step two ($F(3,43) = 22.93$, $p<.001$) predicted 61.6% of the receptive grammar score variance (Adj. R^2 .589).

Table 7.9 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, receptive grammar

Pre-Intervention – 9 Month Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	5.38 (-.02-10.78)	2.68		*p*=.05
Pre-Intervention	.85 (.57-1.13)	.14	.79	*p*<.001
Age	-.01 (-.08-.06)	.03	-.05	*p*=.73
Step 2				
Constant	6.99 (1.50-12.47)	2.72		*p*=.01
Pre-Intervention	.76 (.48-1.05)	.14	.71	*p*<.001
Age	.002 (-.07-.07)	.03	.01	*p*=.96
Group	-1.50 (-3.04-.04)	.76	-.20	*p*=.06

Step 1 R^2=.58, p=<.001 ***Step 2 ΔR^2= .04, p=.06***

7.3.3 Neale Analysis of Reading Ability – Second Revised British Edition (NARA II) (Neale, 1997), Reading Accuracy

7.3.3.1 Reading Accuracy Descriptive Statistics

Table 7.10 below shows the descriptive statistics for the raw scores of the NARA reading accuracy task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at 9 month follow-up.

Table 7.10 Mean, standard deviation and range of NARA reading accuracy raw scores for both trained and active control groups

		Mean	*SD*	*Range*
Reading Accuracy	Pre (n=24)	25.5	18.7	2-67
Trained Group	Post (n=24)	31.3	16.5	2-64
	Follow Up (n=24)	35.7	16.8	2-75
Reading Accuracy	Pre (n=23)	17.4	13.4	1-48
Active Control	Post (n=23)	20.4	13.3	2-48
	Follow Up (n=23)	24.9	13.9	4-53

The percentage of improvements, decreases or no change in reading accuracy scores for both the trained group (n=24) and active control group (n=23) between pre-intervention and 9 month follow-up, pre-intervention and post-intervention, and post-intervention and 9 month follow-up time points are shown below in table 7.11.

Table 7.11 Percentage of participants who made an improvement, a decrease or no change in reading accuracy scores across all time points

		Pre-Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	91.7%	75%	83.3%
	Decreased	0%	12.5%	4.2 %
	No Change	8.3%	12.5%	12.5
Active Control (n=23)	**Improved**	87%	60.9%	82.6%
	Decreased	0%	8.7%	0%
	No Change	13%	30.4%	17.4%

Table 7.11 shows that 91.7% of trained children improved in reading accuracy scores between pre-intervention and the 9 month follow-up time point compared to 87% of active controls. Between pre- and post-intervention 75% of trained children improved in reading accuracy scores compared to 60.9% of active controls. Between post-intervention and 9 month follow-up, 100% of active controls improved or maintained their post-intervention reading accuracy scores, compared to a slightly lower 95.8% of the trained group.

7.3.3.2 Reading Accuracy (Post-Intervention), Hierarchical Regression Analysis

As shown in Table 7.12 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention reading accuracy scores after controlling for pre-intervention score and age. In step one of the model, pre-intervention scores and age in months accounted for 87.2% of the variance in post-intervention reading accuracy scores, however, only the pre-intervention score of reading accuracy was a significant individual predictor. In step two, adding group did not produce a significant change in R^2 (ΔR^2) showing this variable is not a significant predictor, accounting for only a further 1.2% of the variance in post-intervention reading accuracy scores. The final model at step two ($F(3,43) = 109.96$, $p<.001$) predicted 88.4% of the reading accuracy score variance (Adj. R^2 .877).

Table 7.12 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, reading accuracy

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	12.32 (-.32-24.97)	6.28		*p*=.56
Pre-Intervention	.95 (.83-1.07)	.06	.96	*p*<.001
Age	-.08 (-.21-.06)	.07	-.07	*p*=.27
Step 2				
Constant	17.31 (4.28-30.33)	6.46		*p*=.01
Pre-Intervention	.91 (.78-1.03)	.06	.91	*p*<.001
Age	-.06 (-.19-.08)	.07	-.05	*p*=.40
Group	-4.06 (-7.86- -.26)	1.88	-.12	*p*=.03

Step 1 R^2=.87, p=<.001 ***Step 2 ΔR^2= .01, p=.03***

7.3.3.3. Reading Accuracy (9 Month Follow-Up), Hierarchical Regression Analysis

As shown in Table 7.13 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up reading accuracy scores after controlling for pre-intervention score and age. In step one of the model, pre-intervention scores and age in months accounted for 87.9% of the variance in 9 month reading-accuracy scores, however, only the pre-intervention score of reading accuracy was a significant individual predictor. In step two, adding group did not produce a significant change in R^2 (ΔR^2) showing this variable is not a significant predictor, accounting for only a further 0.8% of the variance in 9 month follow-up reading accuracy scores. The final model at step two ($F(3,43) = 112.15$, $p<.001$) predicted 88.7% of the reading accuracy score variance (Adj. R^2 .879).

Table 7.13 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, reading accuracy

Pre-Intervention – 9 month Follow-Up	**B**	**SE B**	**β**	***p***
Step 1				
Constant	11.97 (-1.64-25.58)	6.75		*p*=.08
Pre-Intervention	1.05 (.92-1.17)	.06	.95	*p*<.001
Age	-.05 (-.19-.09)	.07	-.04	*p*=.50
Step 2				
Constant	16.31 (2.04-30.59)	7.08		*p*=.03
Pre-Intervention	1.01 (.88-1.14)	.07	.92	*p*<.001
Age	-.03 (-.18-.11)	.72	-.03	*p*=.65
Group	-3.54 (-7.70-.62)	2.06	-.09	*p*=.09

Step 1 R^2=.88, p=<.001 **Step 2 ΔR^2= .01, p=.09**

7.3.4 Neale Analysis of Reading Ability – Second Revised British Edition (NARA II) (Neale, 1997) Reading Comprehension

7.3.4.1 Reading Comprehension, Descriptive Statistics

Table 7.14 below shows the descriptive statistics for the raw scores of the reading comprehension task for both the trained group (n= 24) and the active control group (n=23) at pre-intervention, post-intervention and at nine-month follow-up.

Table 7.14 Mean, standard deviation and range of reading comprehension scores for both trained and active control groups

		Mean	*SD*	*Range*
Reading Comprehension	Pre (n=24)	7.4	4.9	0-15
Trained Group	Post (n=24)	8.7	5.2	0-18
	Follow Up (n=24)	10.9	4.2	16-31
Reading Comprehension	Pre (n=23)	4.2	4.2	0-12
Active Control	Post (n=23)	4.7	4.7	0-15
	Follow Up (n=23)	6.1	5.2	0-18

The percentage of improvements, decreases or no change in reading comprehension scores for both the trained group (n=24) and active control group (n=23) between pre-intervention and 9 month follow-up, pre-intervention and post-intervention, and post-intervention and 9 month follow-up time points, are shown below in table 7.15.

Table 7.15 Percentage of participants who made an improvement, a decrease or no change in reading comprehension scores across all time points

		Pre-Follow up	*Pre-Post*	*Post-Follow up*
Trained Group (n=24)	**Improved**	87.5%	54.2%	79.2%
	Decreased	4.2%	16.7%	0 %
	No Change	8.3%	29.2%	20.8%
Active Control (n=23)	**Improved**	78.3%	34.8%	78.3%
	Decreased	0%	13%	0%
	No Change	21.7%	52.2%	21.7%

Table 7.15 shows that 87.5% of trained children improved in reading comprehension scores between pre-intervention and the 9 month follow-up time point compared to 78.3% of active controls. Between pre and post-intervention 54.2% of trained children improved in reading comprehension scores compared to 34.8% of active controls. Between post-intervention and 9 month follow-up, 100% of both trained and active controls improved or maintained their post-intervention reading comprehension scores.

7.3.4.2 Reading Comprehension (Post-Intervention), Hierarchical Regression Analysis

As shown in table 7.15 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of post-intervention reading comprehension scores after controlling for pre-intervention score and age. In step one of the model, pre-intervention scores and age in months accounted for 87.3% of the variance in post-intervention reading comprehension scores, however, only the pre-intervention score of reading comprehension was a significant individual predictor. In step two, adding group did not produce a significant change in R^2 (ΔR^2) showing this variable is not a significant predictor, accounting for only a further 0.4% of the variance in post-intervention reading comprehension scores. The final model at step two ($F(3,43) = 101.72$, $p<.001$) predicted 87.7% of the reading comprehension score variance (Adj. R^2 .868).

Table 7.16 Hierarchical multiple regression analysis concerning the predictors of immediate post-intervention, reading comprehension

Pre-Intervention – Post-Intervention	**B**	**SE B**	**β**	***p***
Step 1				
Constant	2.59 (-1.38-6.57)	1.97		*p*=.20
Pre-Intervention	1.07 (.93-1.21)	.07	.96	*p*<.001
Age	-.02 (-.07-.02)	.02	-.06	*p*=.32
Step 2				
Constant	3.33 (-.85-7.51)	2.07		*p*=.12
Pre-Intervention	.1.04 (.89-1.19)	.07	.93	*p*<.001
Age	-.02 (-.06-.03)	.02	-.05	*p*=.45
Group	-.69 (-5.70 - -2.54)	.61	-.07	*p*=.27

Step 1 R^2=.87, p=<.001 ***Step 2 ΔR^2= .004, p=.27***

7.3.4.3 Reading Comprehension (9 Month Follow-Up), Hierarchical Regression Analysis

As shown in table 7.16 below, a hierarchical multiple regression was undertaken to examine whether group (trained vs active control) was a significant predictor of 9 month follow-up reading comprehension scores after controlling for pre-intervention score and age. In step one of the model, pre-intervention scores and age in months accounted for 90% of the variance in 9 month follow-up reading comprehension scores, however, only the pre-intervention score of reading comprehension was a significant individual predictor. In step two, adding group did not produce a significant change in R^2 (ΔR^2) showing this variable is not a significant predictor, accounting for only a further 0.7% of the variance in 9 month follow-up reading comprehension scores. The final model at step two ($F(3,43) = 140.64$, $p<.001$) predicted 90.7% of the reading comprehension score variance (Adj. R^2 .901).

Table 7.17 Hierarchical multiple regression analysis concerning the predictors of 9 month follow-up, reading comprehension

Pre-Intervention – 9 month Follow-Up

	B	**SE B**	**β**	***p***
Step 1				
Constant	3.31 (-65-7.26)	1.96		*p*=.10
Pre-Intervention	1.22 (1.08-1.36)	.07	.97	*p*<.001
Age	-.02 (-.06-.02)	.02	-.05	*p*=.37
Step 2				
Constant	4.42 (.35-8.5)	2.02		*p*=.03
Pre-Intervention	1.17 (1.03-1.32)	.07	.93	*p*<.001
Age	-.01 (-.06-.03)	.02	-.03	*p*=.57
Group	-1.04 (-2.24-.16)	.60	-.09	*p*=.09

Step 1 R^2=.90, *p*=<.001   Step 2 ΔR^2= .007, *p*=.09

7.4 Discussion

This chapter investigated whether group membership (trained vs active control) was a significant predictor of performance on far transfer measures (sentence comprehension, receptive grammar, reading accuracy and reading comprehension) at the post-intervention time point and at a 9 month follow-up. Hierarchical multiple regressions showed that group was a significant predictor of sentence comprehension, even after controlling for age and pre-intervention score, at both the immediate post-intervention and 9 month follow-up time points. This suggested that being in the working memory training group was related to greater improvements in sentence comprehension than being in the active control group. For a summary of all far transfer results, please refer to the summary of results, table 7.18, below.

Table 7.18 Far transfer measure results summary

Task and Group	Pre-Post intervention		Pre-intervention to 9-month follow-up	
	Mean Increase of Scores	Variance explained	Mean Increase of Scores	Variance explained
Sentence Comprehension				
Trained	4.8	**β** =-.33 *p*<.001	6.5	**β** =-.32 *p*<.001
Active control	1.3		3.5	
Receptive Grammar				
Trained	2.8	**β** = -.17 *p*=.13	3.6	**β** = -.20 *p*=.06
Active control	1.6		2.6	
Reading Accuracy				
Trained	5.8	**β** = -.12 *p*=.03	10.2	**β** = -.09, *p*=.09
Active control	3		7.5	
Reading Comprehension				
Trained	1.3	**β** = -.07 *p*=.27	3.5	**β** = -.09 *p*=.09
Active control	0.5		1.9	

Previous meta-analytic studies have found limited evidence for far transfer effects following working memory intervention studies (Melby-Lervag & Hulme, 2013; Melby-Lervag et al., 2016; Sala & Gobet, 2017; Schwaighofer et al., 2015). The lack of evidence of far transfer effects to receptive grammar, reading accuracy and reading comprehension in the current study, once Bonferroni corrections had been made, is therefore in line with previous findings. However, as group was a significant predictor of sentence comprehension performance, at both post-intervention and 9 month follow-up in the current study, the current face-to-face adaptive working memory intervention that targeted both verbal and visuospatial domains has been shown to produce far transfer effects for children with language disorder.

Although group was a significant predictor of sentence comprehension at both time points, the current study had hypothesised that group would also be a significant predictor of reading comprehension, immediately post-intervention and at a 9 month follow-up. This was based on previous literature already discussed, that reported a similar non-computerised working memory training intervention study, with typically developing children, finding reading comprehension gains for trained children at a 12-month follow-up (Henry et al., 2014). Contrary to the current study's prediction, there were no significant differences between trained and active control groups in reading comprehension at either time point. It should be noted that Mann-Whitney U tests indicated significant differences between the trained and active control groups in both reading comprehension and reading accuracy scores at pre-intervention, despite random allocation to groups (see section 7.2.1). Therefore, if any significant group differences had occurred in reading accuracy or comprehension in the current study, they would have needed to be interpreted with caution.

The current study's sample of children had language difficulties and would, therefore, be expected to have problems processing and storing larger chunks of information. The finding that group was a significant predictor of sentence comprehension suggests that the intervention did improve the trained children's ability to divide attention between processing and storage, which is essential for comprehension ability. However, it is possible they were only able to utilise this skill for smaller chunks of text such as sentences, and/or for information that was heard rather than read (as reading skills in the current study's sample of children were generally very weak). For reading measures, from a total of 47 participants, 18 children did not score highly enough to attain a standardised score in reading accuracy and 20 children did not score highly enough to attain a standardised score in reading comprehension. Table 7.19 below shows the descriptive statistics of reading measures for those participants who did score highly enough to attain a standardised score.

Table 7.19: Mean, standard deviation and range of standardised reading accuracy scores and standardised reading comprehension scores

Group and Task	Mean	SD	Range
Trained			
Reading Accuracy (n=17)	97.47	5.35	88-113
Reading Comprehension (n=16)	89.47	5.64	82-102
Active Control			
Reading Accuracy (n=12)	88.4	13.19	72-112
Reading Comprehension (n=11)	84.9	8.31	75-104

The weak reading levels of the children in the current study could therefore explain why children with language disorder significantly improved in sentence comprehension ability, but not reading comprehension, in contrast to the finding of significant gains in reading comprehension for typically developing children (Henry et al., 2014).

As discussed in section 7.1 of this chapter, the development of 'Situation Models', a coherent meaning-based visuospatial mental representation of a situation described in a piece of text, may be more difficult for children with language disorder. In the current study, listening comprehension was confined to sentence level, whereas reading involved comprehension beyond sentence level, therefore involving the development of more complex 'Situation Models'. This may explain the differences found between the current study and Henry et al (2014). Furthermore, it is also important to consider that as the working memory intervention itself was delivered orally, it may be that improvements in dividing attention between processing and storage were confined to the oral domain. Alternatively, it may just be that children's comprehension was poor in the current study as the children had difficulties comprehending what they had read. The 'Simple View of Reading' (Gough & Tunmer, 1986) considers reading to have two basic components, decoding (the ability to identify words in print) and linguistic comprehension (ability to understand spoken language) (Nation, 2019). Having strong decoding and linguistic comprehension skills are therefore crucial in being able to proficiently comprehend text. As those with language disorder also had generally weak reading skills in the current study's sample of children, 'The Simple View of Reading' (Gough & Tunmer, 1986), may help to explain why the trained group did not make significant gains in reading comprehension as found with typically developing children.

As findings suggest that the working memory intervention improved the trained child's ability to divide attention between processing and storage, it could be expected that a more generalised transfer could also be evident to other comprehension based tasks such as the Test for Receptive Grammar (Bishop, 2003). Although both the sentence comprehension and receptive grammar tasks require the use of comprehension ability, the sentence comprehension task only requires a trained child to remember the general concept of an oral sentence whereas the receptive grammar task also requires the child to fully comprehend specific grammatical knowledge which was not targetted or trained in the current study.

In Summary, the findings of this study contrast with the general view that working memory training does not produce far transfer effects (Melby-Lervag & Hulme, 2013; Melby-Lervag et al., 2016; Sala & Gobet, 2017 and Schwaighofer et al., 2015). While findings for receptive grammar, reading accuracy and reading comprehension were consistent with this view, significant gains observed in sentence comprehension provide evidence of a far-transfer effect. As training methods are often computer-based, using alternative approaches, as the current study does, could be important in achieving evidence for gains in other areas of academic ability and cognitive skills (Melby-Lervag & Hulme, 2013).

Chapter 8 Discussion

8.1 Introduction

The research reported in this thesis addressed the question of whether a working memory intervention that focused on executive-loaded working memory tasks in both the verbal and visuospatial domains of the central executive would benefit children with language difficulties. More specifically the questions concerned whether the working memory intervention would have a direct benefit on the trained tasks administered, whether there would be any near transfer effects to other areas of working memory that were not directly trained, and finally whether there would be any far transfer effects to other academic skills such as reading or sentence comprehension. The research also concerned issues related to individual differences and dosage.

This chapter will discuss the findings of the current study's working memory intervention, how this relates to previous literature, the limitations of previous work, and the novel contributions the current study adds to our knowledge. Individual differences between trained participants will then be explored, followed by a discussion about the dosage required for both Odd One Out and Listening Recall tasks. Future research will be considered, along with clinical and educational implications of the findings. Finally, some of the current study's limitations will be discussed.

8.1.1. Effects of the Working Memory Intervention

Hierarchical multiple regression analyses showed that group (trained vs active control) was a significant predictor of performance in the two trained working memory tasks, Listening Recall and Odd One Out, at both post-intervention and at a 9 month follow-up (after controlling for pre-intervention score and age). Thus, those with language difficulties responded to a targeted working memory intervention in both the verbal and visuospatial domains and, importantly, could maintain improvements over a nine month period.

Hierarchical multiple regressions also showed group membership (trained vs active control) to be a significant predictor of performance, at both time points, in all near transfer effects that were measured (again, controlling for pre-intervention scores and age): digit recall, word list recall, block recall, counting recall, backward digit recall and pattern span. Group membership was therefore not only a significant predictor of performance on the trained tasks, but also on all other areas of working memory (near transfer) which were measured at both time points.

This study also aimed to investigate whether group was a significant predictor of performance, immediately post-intervention and at a 9 month follow-up, in the far transfer effects measured (controlling for pre-intervention scores and age): reading accuracy and reading comprehension, receptive grammar, and sentence comprehension. Hierarchical multiple regression showed that group (trained vs active control) was a significant predictor of performance in sentence comprehension, which was not only evident immediately post-intervention but also remained evident at a 9 month follow-up. No other far transfer effect measures showed any significant difference between groups at any time point when Bonferroni adjustments were carried out.

8.2 Previous Literature Comparisons

Working memory interventions have become one of the most debated issues in cognitive psychology with conflicting arguments regarding outcomes of interventions and near and far transfer effects (Szewcyzk, 2016). Working memory training interventions are often deemed successful by the presence of near or far transfer effects, i.e., where gains in other areas of working memory (near transfer) or other academic skills (far transfer) are achieved. The current study indicates that children with language difficulties respond to a targeted working memory intervention in both the verbal and visuospatial domains and, importantly, group (trained vs active control) remained a significant predictor of performance over a nine month period. It is generally accepted that working memory interventions produce gains in directly trained working memory tasks. For example, a meta-analysis of 87 papers in this field found large effects for trained task improvements when looking at measures which were identical or closely related to the working memory trained tasks; and these gains were also reported to be maintained at a five month follow-up (Melby-Lervag et al., 2016). It is important to stress that this meta-analysis considered only computer-based working memory interventions (see below for more on this issue). Nevertheless, the findings of the current study using a face-to-face intervention provide robust support for the conclusions of Melby-Lervag et al. (2016).

Working memory interventions have also been suggested to yield a short-term improvement in verbal and visuospatial tasks, although these often are reported not to be sustained after a delay of a few months (Melby-Lervag et al., 2016). The current study does not support this conclusion, as group was found to be a significant predictor of both directly trained tasks and all six near transfer effect measures at both post-intervention and also the 9 month follow-up time points. These findings indicate that improvements achieved through an adaptive face to face working memory intervention can be maintained for longer periods.

Although near transfer effects give some indication of the effectiveness of working memory interventions, the strongest evidence of effectiveness can be seen in far transfer effects. These demonstrate a gain in ability that is not directly related to a trained working memory task in either structure or content (Henry et al., 2014). Far transfer effects are rarely present in any working memory intervention study, regardless of the participant group. Melby-Lervag et al. (2016) discuss in their meta-analysis of 87 papers that far transfer effects on reading comprehension and arithmetic were present in only two studies, however, even these findings were later described as weak and not sustainable on closer inspection.

The meta-analysis undertaken by Melby-Lervag et al. (2016) was conducted with studies that solely used computer based working memory interventions. Rowe et al. (2019a) was the first to review the effectiveness of non-computerised working memory interventions in children aged 4-11 years. 18 publications were included, with only three papers using direct working memory interventions without strategy instruction, as used in the current study. From these three papers, near and far transfer effects were mixed. One study, an executive-loaded working memory training intervention with typically developing children using a face to face approach, showed positive gains in direct transfer, near transfer (counting recall and word recall) and a far transfer effect in reading comprehension at a 12 month follow-up (Henry et al., 2014). As the current study also found a far transfer effect in a comprehension outcome, the current study supports the idea that improving the ability to divide attention between processing and storage could have a specific benefit for comprehension ability (Henry et al., 2014). As the current study also has found a far transfer gain, this highlights the importance of further exploration into face to face working memory intervention studies in comparison to the more traditional computer based programmes, which previous research has indicated give very little gains in far transfer effects.

Although limited, the few working memory intervention studies that have emerged over recent years for children with language difficulties have also shown promising results. It is important to note that these studies were not included in Melby-Lervag et al. (2016) meta- analysis paper discussed above. For example, Acosta et al. (2019) found that children with language disorder responded positively to a working memory intervention which targets immediate verbal memory, i.e. what an individual can repeat immediately after perceiving it, visuospatial working memory and verbal working memory. Those with language disorder significantly improved at post-intervention time point on all function measures of short-term verbal memory, verbal working memory, short-term visuospatial memory, visuospatial working memory, attention and processing speed, and lexical-semantic processing ability. Significant gains in all measures were also evident in the language disorder group, compared to chronological age matched controls.

Near transfer gains have also been reported in visuospatial short-term memory, verbal short-term memory and central executive following a 'Cogmed Working Memory Training' (computerised) program in a group of 12 children with low language ability (Holmes et al., 2015). Henry and Botting (2017), in a review paper point out that once corrections had been made for multiple comparisons in the Holmes et al. (2015) study, only visuospatial gains remained as a significant improvement for children with low level language ability. In the current research, even after Bonferroni corrections for multiple comparisons had been made, group was still a significant predictor of performance in all six near transfer measures (digit recall, word list recall, block recall, counting recall, backward digit recall and pattern span) at both post-intervention and also the 9 month follow-up time points.

Research into working memory interventions for children with language difficulties still remains relatively limited, however. It is also interesting to note that an important difference between Holmes et al. (2015) and this study together with the one conducted by Henry et al. (2014) as well as Acosta et al. (2019) was the use of a computer based program, rather than a face to face approach. As positive gains in both near and far transfer effects have been reported in the studies that have used a face to face approach, this is certainly an area that needs further exploration especially with children with language difficulties who have shown gains in direct, near and far transfer effects following the current working memory training intervention.

Having looked at previous working memory research, the following section will discuss its limitations and how the current study helps to extend our understanding about the response to working memory intervention for children with language difficulties.

8.2.2. Limitations in previous research: what does the current study add to our knowledge?

Evidence of a reading comprehension far transfer effect has been reported by Henry et al. (2014), where typically developing trained participants scored significantly higher than active controls at a 12 month follow-up. However, there was no pre-intervention measurement of reading comprehension for trained or active control participants taken in this study, which is a methodological weakness. Therefore, it cannot be said definitively that the significant group difference in reading comprehension at the 12 month follow-up was due to the intervention itself. Promisingly though, the participant groups did not differ on any other measure at pre-intervention time point, i.e. matrices, verbal similarities, number skills, word recall, digit recall, block recall, word reading, Listening Recall, counting recall or

Odd One Out span, so it is plausible that this effect could apply to reading comprehension measures as well.

In terms of the current study, despite the prediction based on previous literature, group (trained vs active control) was not a significant predictor of performance in reading comprehension either immediately post-intervention or at a 9 month follow-up. In fact, in the current study, there were 'unwanted' significant differences between the trained and active control groups in both reading comprehension and reading accuracy scores at pre-intervention time point, despite random allocation to groups. Therefore, if group had been revealed to be a significant predictor of performance in either reading accuracy or comprehension at any time point, results would have needed to be interpreted with caution. Many children in the current study did not score highly enough to be given a standardised score for either reading comprehension or reading accuracy, and 11% of participants scored zero throughout the study, with no improvement at all in either reading measure. In hindsight, the reading measure was simply too difficult for the current sample. This could explain why alternatively group predicted performance in sentence comprehension for this group of children, rather than reading comprehension as found by Henry et al. (2014). Children in the current study had language difficulties, therefore, they tended to display problems processing and storing larger chunks of information. Improving these skills enabled children with language difficulties to improve in the ability to divide attention between processing and storage. However, they could only utilise this skill for smaller chunks of text such as sentences, and for information that was heard rather than read (as reading skills generally were very weak).

A useful suggestion for understanding the differences in listening and reading performance outcomes in the present intervention is the possibility that the children with language disorder found it difficult to create ‘Situation Models’, discussed previously in section 2.2.2. Through creating a coherent meaning based mental representation of a situation described in a piece of text, readers develop a visuospatial representation of the situation which allows for a better understanding (De Koning & Van der Schoot, 2013). In order to achieve ‘Situation Model’ level of comprehension and move beyond word, phrase and sentence comprehension, individuals need to be proficient readers (De Koning & Van der Scoot, 2013). Due to the reading difficulties experienced by the children in the current study, children with language disorder may have found it challenging during reading comprehension tasks, where it was required to read short narratives that went beyond sentence level, to create a ‘Situation Model’. In comparison, sentence comprehension was measured through listening comprehension, with no reading requirement, therefore children in the current study may have been able to form ‘Situation Models’ much more easily for sentence comprehension tasks. Comprehension is a domain general skill and the modality of initial information input (listening comprehension or reading comprehension), should not influence the creation of a ‘Situation Model’ (Gernsbacher et al., 1990). This could therefore explain why group was a significant predictor for sentence comprehension ability for children with language disorder in the current study, compared to the significant gains found in reading comprehension for typically developing children by Henry et al. (2014).

Previous working memory intervention studies with those with language disorder tend to have had very small sample sizes, and this together with the small number of investigations highlights a need for further investigation into working memory interventions for children with language difficulties. The current study, with a sample size of 47 children, has begun this process, but much more research is needed to be undertaken with even bigger sample sizes in order to give a definitive answer as to how children with language difficulties respond to working memory interventions.

In addition, further exploration of face to face working memory interventions is also required. Previous research in this field has often been pessimistic about positive gains, and positive outcomes relating to near and far transfer effects have been limited. The current study however, found that group was a significant predictor of performance in direct transfer effects for both trained tasks, all six near transfer effects, and a far transfer effect for sentence comprehension at both post-intervention and a 9 month follow-up. This indicates that children with language difficulties not only respond to an adaptive face to face working memory training intervention, but can also maintain or make further improvements in other areas of working memory and sentence comprehension. In relation to previous research, the only other working memory intervention study that has yielded similar outcomes with a language related far transfer effect at a follow-up time point is the Henry et al. (2014) study, which also adopted a face to face approach. Future research, therefore, needs to explore this area and carry out working memory intervention studies with children with language difficulties using much larger sample sizes.

In summary, the current study helps to extend our understanding about the response to working memory intervention for children with language difficulties. Although promising results were emerging from the limited studies available for children with language difficulties, an executive-loaded working memory intervention study that used a larger sample size and a face to face approach was essential to expand understanding in this area and give a more definitive answer as to whether children with language difficulties are able to benefit from a working memory intervention. The current study is the start in trying to change the pessimistic and conflicting viewpoint of previous research in the working memory intervention field. With future research concentrating on face to face adaptive training interventions rather than the more traditional and non-productive computer based programmes, there may be a shift in the far transfer gains produced via working memory training interventions with individuals both typically developing and with language difficulties.

Due to the vast heterogeneity of the language disorder population (chapter 1, section 1.1), the following section will explore individual differences of the trained participants within the current study. Initial span levels, profile types and dosage effects will be explored in relation to their effects on individual improvements in trained tasks.

8.3 Individual differences in Trained Participants

Differences in findings between participant groups within working memory training studies have raised questions about the types of children who may benefit from working memory intervention (Rowe et al., 2019a), about individual differences within groups of children and about the gains made during the intervention. The current study explored individual differences in the trained group of participants to gain a clearer understanding of these children and their heterogeneity, and explore factors that might account for different outcomes. Other studies have also adopted this approach and as discussed previously, a recent study conducted by Blom and Boerma (2020), explored whether severity and persistence of language disorder impacted working memory performance.

First, the current study looked at whether those who made the *most* improvement during the 18 training sessions generally started with a lower memory span, on the grounds that a child with low span at the outset would have greater room for improvement during the training sessions. For both trained tasks, independent *t*-tests showed that at session 1, there were no significant differences between high or lower improvers, however, by session 18, a further independent *t* test revealed a significant difference between the two groups. This therefore indicates that the differences in improvement between the trained participants by the end of the working memory training intervention for both trained tasks cannot be linked to the participant's initial memory span levels. Therefore, it is reassuring to know that all children have the same opportunity to improve their working memory span levels during the training sessions regardless of their initial abilities.

Continuing to explore individual differences, it became apparent following the working memory training intervention sessions that some children displayed a 'spikey profile' throughout the 18 sessions. In other words, some participants displayed a fluctuation in span levels between sessions, whereas other participants increased span levels rather smoothly. This led to the investigation of whether profile type, spikey versus smooth, made a difference to improvement level. A repeated measures two-way ANOVA found that those with a spikey profile gained significantly higher scores by session 18 in both trained tasks compared to those who displayed a smooth profile. This is an interesting finding and indicates the need for further exploration into the individual's optimum span level reached during training sessions, and what this means in terms of levels of working memory improvement made. It may be that in order to reach an optimum span level, children with language disorder need extra time to process and consolidate the gains achieved in previous training sessions. This could therefore equate to the fluctuation in span levels between sessions, for those who displayed a spikey profile. Spikey profiles could also be related to the ‘Overlapping Waves Model’ (Siegler, 1996), discussed in section 2.1.2. According to this model, children have a number of different strategies in their repertoires at any given time, however, they transfer reliance to different strategies with age and experience (Kwong & Varnhagen, 2005). As children learn a task, they show increasing reliance to more effective strategies and decrease their reliance on less effective strategies (Kwong & Varnhagen, 2005). Therefore, a ‘spikey profile’ may relate to children trying out a greater range of different strategies, some successful and others less so.

Rowe et al. (2019a) discuss that few working memory intervention studies report dosage effects, so this was seen as an important area to explore to extend understanding further. Having observed training effects, it was important to consider training dosage and whether the full 18 training sessions were required to achieve the observed improvements. Repeated measures one-way ANOVAs showed that in Listening Recall training, participants made an initial sharp increase with a significant difference between sessions one and six, which then levelled off between sessions six and 12, before making another significant increase again between sessions 12 and 18. This indicates that participants required the full 18 training sessions in Listening Recall to achieve their optimal outcome. It is possible that the non-significant span increase between sessions 6 and 12 reflected the time period required to secure and consolidate the initial Listening Recall span increase achieved from session 1-6, before allowing a further span increase to occur in the final third of the training sessions. As explained above, this is a similar concept to the possible time period needed to secure progression during training sessions for those who displayed a ‘spikey profile’. This insight indicates that the Listening Recall span task requires the full 18 training sessions, however it also cannot be ruled out that by having further training sessions, further gains may also be achieved. This finding helps to extend our understanding of Listening Recall training in those with language difficulties and will inform the implementation of future studies.

In contrast to the findings for Listening Recall, a further repeated measures one-way ANOVA of Odd One Out scores showed that following a sharp increase from session one to 12, there was no significant difference between sessions 12 and 18. This suggests that for future studies and interventions it may be possible to reduce the number of working memory training sessions for the Odd One Out span task. Although visuospatial span has been shown to plateau earlier than verbal span in the current study, a precise number of sessions required cannot be definitively stated. Again, it cannot be ruled out that further training in the Odd One Out span could accrue further benefits and gains at a later point.

The following section will discuss how the current study can be used in future research, with a main focus on future research increasing sample sizes, exploring strategy mechanisms further and considering participant inclusion criteria and language disorder severity.

8.4 Future Research

When considering how the working memory intervention delivered in the current study could be used in future research, as discussed previously, it would be beneficial to repeat it with a larger sample size. Previous research in this field has often been carried out with very small sample sizes of 12 participants and under (Holmes et al., 2015; Wener et al., 2011). Acosta et al.'s (2019) study was carried out with a sample size of 32 participants and the current study has increased this to 47 participants, 24 trained and 23 in the active control group, so extending this sample size even further would be the next step in future research. Realistically however, this would be difficult to achieve with just one researcher. Delivering an intervention of this kind to a number of children requires hours of time, therefore, increasing participant numbers may be problematic without significant resources.

A further interesting question is, who is able to deliver this intervention? Could teaching assistants in primary schools across the country be a part of a much bigger sample size? From experience of visiting schools and delivering this intervention, school staff were very interested in this work, enthusiastic about its potential outcomes, and extremely eager to try it out for themselves. With appropriate training, and if given the correct resources, future research could look at whether teaching assistants or language support staff in schools could implement this working memory training intervention themselves, record and report the progress made. This could then potentially give a much larger sample size and extend our knowledge and understanding in this field much further. Through school staff implementing this working memory training intervention themselves, this would also enable training to be conducted in children's everyday context, an area of working memory training that is receiving increased interest (Rowe et al., 2019a). Interventions conducted in children's everyday contexts have been suggested to have the potential to increase working memory skills and produce transfer to 'real world' skills such as attention, language and academic ability (Rowe et al., 2019a).

Whilst delivering the working memory intervention study to children in the trained group, it became apparent that some of them used various techniques or strategies to try to improve their working memory capacity during the Odd One Out task. This was not suggested or encouraged by the researcher, but the children had a range of unique methods to increase their span levels during training sessions. The use of self-generated strategies, observed in the current study, supports the finding of Feng and Fuchs (2015) that 28% of participants used strategies during working memory training without any instruction to do so. The use of strategy during training is also related to Siegler's 'Overlapping Waves Model' discussed previously, which postulates that children try out a range of different strategies when performing a task, with an increased reliance formed for more effective strategies and a decreased reliance for less effective strategies over time (Kwong & Varnhagen, 2005). Children in the current study often saw the training sessions as a game, and would remember their previous span levels and try to improve upon their last scores, often becoming quite competitive with themselves throughout training. One strategy method that stood out from observations made was the use of gestures in aiming to improve span levels of the Odd One Out task. One participant would move a part of the body to correspond to the Odd One Out shape's location. For example, if the left or right shape was the odd one out, the child would shrug their corresponding shoulder; or if the middle shape was the odd one out, the child would nod their head. This then led to a pattern of gestures being performed by the child whilst they answered each Odd One Out trial. This child on average, had a difference of 1.9 span levels between their highest and lowest Odd One Out span scores achieved across the 18 training sessions, indicating that the use of gestures and physical movements may have helped to increase span level.

The most successful use of gestures observed in a participant during the Odd One Out training sessions, however, was a child that used finger tapping on a table to represent the position of the Odd One Out shapes. If the Odd One Out shape was located on the left, the child would tap their index finger on the table; the middle shape was represented by the middle finger; and the right hand shape by their ring finger. The child, therefore, presented a sequence of taps on the table whilst recalling the order of the Odd One Out shapes during each training session. This child showed the most improvement across all trained participants and had a difference of 4.2 span levels between their lowest average span and highest average span level across the 18 training sessions. What was particularly interesting about this highest Odd One Out improver and their finger tapping gestures was that this child played the piano. Complementing Ericcson and Kintsch's (1995) Long Term Working Memory theory, discussed in section 1.2, experts and skilled performers are said to have an expanded working memory capacity in activities that they have practiced and acquired knowledge and specialist memory skills for. It would be very interesting to look at gestures and improvements in working memory capacity in future research. Could strategies be taught that involve physical gestures when performing the working memory training in order to improve span capacity even further? The two participants discussed here show that they were able to increase their working memory span in the visuospatial domain through the incorporation of gestures and physical movements, therefore, this would be worth further exploration.

The current study used rigorous criteria to determine which participants could be included or, in turn, excluded from the current research. This was to ensure that the children included in the current study had language difficulties as a primary need. As this was the focus of the current research, these strict criteria were necessary. A question, however, for future research arises as to whether it is necessary to have such rigorously defined criteria when it would be hoped that this intervention could work for any child with a language difficulty.

In setting these stringent inclusion criteria for the current study, there were children within a school's cohort who staff felt would benefit from the working memory training intervention but did not fulfil the study's criteria. One particular group of children that this applied to were those with Autism Spectrum Disorder. Staff were reassured that these children would still benefit from the intervention, just at a later date when school staff had been trained in how to deliver the working memory training tasks themselves. Future research could, therefore, adjust the inclusion criteria and base inclusion on a formal observation of who would benefit, looking at any child who presents with a language difficulty regardless of any other co-occurring conditions.

A further observation made during the working memory training was the wide spectrum of differences in language abilities across the participants. In the current study all participants attended a mainstream school, however some were part of a specialist speech and language provision within a mainstream school, and others were solely part of a mainstream classroom with ongoing speech and language support. This seemed to vary dependent on area, and as the study was conducted across two counties, Dorset and Hampshire, schools differed between counties in their approaches. Children were included within the current study if they scored a scaled score of seven or below in two of the Clinical Evaluation of Language Fundamentals (CELF 4) subtests (Semel et al.,2006). This is rather lenient when considering language difficulty, however, it was felt that if children scored below average in two core areas of language, they were having sufficient difficulty to require additional support and benefit from the working memory intervention. In doing so, this study included children with a range of language needs and it was noted that those in specialist language provisions within mainstream schools in Hampshire seemed to have more complex needs. It would be beneficial to explore these differences in future research in order to gain a deeper understanding of working memory training and language difficulty. Does a child with a more complex language difficulty respond differently to the working memory intervention than a child who falls just below average in two core areas? Complementing Blom and Boerma's

(2020) recent findings that severity and persistence of language disorder impacts visuospatial working memory ability, this would be an interesting area and one that could be further explored in future work to give us further insight into working memory training and those with language difficulties.

As well as considering how the current study can be used in future research, the limitations of the current study also need to be discussed. The following section of this chapter will move on to focus on the limitations of the current study with a main focus being the extent of reading accuracy and reading comprehension difficulties within the language disorder group of children. If this study was therefore repeated or extended further in the future with children with language disorder, the inclusion of these outcome measures would need to be reviewed.

8.5 Study Limitations

One of the main limitations of this study was the fact that it was not a double blind experiment and the researcher was aware of the participant's group membership. Unfortunately for a PhD with one researcher, it would have been impossible to conduct a double blind experiment. Some mitigating steps were taken, however. It was ensured that the participants were unaware if they were a part of the trained or active control group, and that all children assumed that they did the same activities as their peers. The school teachers and support staff were also unaware of those who were trained or active control participants and were only told on completion of the study when the intervention and all follow-up measurements were completed. The researcher did ensure that, although not blind to the conditions, both groups of children found their sessions engaging, fun and interactive regardless of the random group allocation.

A second limitation of this study, as discussed previously, concerned the reading accuracy and reading comprehension measurements. Although group was not found to be a significant predictor of reading accuracy or reading comprehension at either post-intervention or 9 month follow-up time points, the trained and active control groups did differ significantly on these measures at the pre-intervention time point. Therefore, any reading accuracy and reading comprehension results from this research need to be looked at with caution. As discussed previously, reading scores in both accuracy and comprehension differed greatly within this group of children. Many did not achieve a standardised score in either accuracy or comprehension as their reading ability was too low. It was envisaged that children within this group would have had a basic level of reading and the extent and scope of their reading difficulties were not predicted. In fact, it was predicted that after the intervention, group would be a significant predictor of reading comprehension performance, based on previous findings (Henry et al., 2014) with typically developing children at a 12 month follow-up. This limitation would definitely need consideration again, if conducting future research with a similar cohort; and the inclusion of any reading measurements would need to be carefully considered and adapted to children's current levels of ability.

Moving forward, the following section will discuss clinical and educational implications. The importance of a non-computerised training approach will be discussed alongside the positive feedback that was received from both children and teachers regarding an interactive face to face training approach that caused fairly minimal disruption to the school day.

8.6 Clinical and Educational Implications

The working memory training intervention implemented as part of this research was a short, face to face training program that the majority of children seemed to engage with and enjoy. As previous studies have mainly used computer based working memory training programs with little human interaction, through conducting a face to face intervention the children's positive reactions and responses were able to be gauged. Schools were very open to participating in the research and were supportive of an intervention to improve working memory in those with language difficulties. Many teachers discussed that working memory was a crucial area that they felt their pupils needed to improve in so were extremely supportive and accommodating in allowing the research to take place.

In terms of the working memory intervention itself, children enjoyed taking part, and all children who started the training program happily completed all 18 training sessions. Having a personal interaction during the working memory training sessions and building a relationship between the researcher and child was a fundamental aspect of keeping motivation and engagement levels high. The children in the trained group loved to try and improve their previous score in both tasks, and this made the training sessions more of a game to them, in fact some of the children referred to the sessions as a 'memory game'. This was very positive feedback from them, as it was clear that they did not feel they were training their working memory but having fun. The active control group still came along for 18 sessions, but all memory element of their tasks were removed. Nevertheless, both groups of children seemed to enjoy the time spent at each training session and always came along willingly when called out of their lessons to participate.

In terms of disruption to a school day, each training session lasted approximately ten minutes, three times a week. Some sessions towards the end of the 6 week schedule lasted a little longer as the trained child's working memory capacity had often improved so the trials took longer to administer. Even so, the majority of the children in this group are familiar with working outside of the classroom with language support staff or their speech and language therapist, so leaving their classroom setting for a short period of time did not seem strange to them, their teachers or their peers.

The intervention itself is a simple design that is easy to follow and does not require extensive training to be able to be administered. With minimal training, teaching assistants or language support staff would be capable of running a 6 week intervention program independently for pupils in their own school setting. This links in with future research, as school staff could be trained to deliver this working memory training intervention in order to roll this work out with a much larger sample across a wider area.

8.7 Conclusion

Working memory intervention research has produced conflicting results and received a mixture of support and condemnation over recent years, with differing outcomes found in near and far transfer effects (Melby-Lervag et al., 2016; Rowe et al., 2019a). Working memory intervention studies concerning those with language difficulties have had limited focus, with very few published findings, and of those available, often with extremely small sample sizes used (Holmes et al., 2015; Wener et al., 2011). The current study aimed to extend our understanding of working memory intervention in those with language difficulties by administering a six week face to face working memory training intervention to a reasonably large sample of children, which focused on two executive-loaded working memory tasks within the verbal and visuospatial domain. Hierarchical multiple regression found group (trained vs active control) was a significant predictor of performance at both post-intervention and 9 month follow-up in both of the directly trained tasks (Listening Recall and Odd One Out), all 6 near transfer effects (digit recall, word list recall, block recall, counting recall, backward digit recall and pattern span), and in one far transfer effect (sentence comprehension). In terms of previous research, directly trained working memory tasks often show a significant gain, in line with this research, however near effects are less frequently apparent, and far transfer effects are rarely reported. It is suggested that the current study's findings were possibly due to its implementation of a short, face to face intervention, compared to the commonly used computer based working memory training programs previously administered. Within the limited previous research that has used a face to face working memory training intervention, one other study has found a far transfer effect in reading comprehension at a 12 month follow-up in typically developing children using the same training programme (Henry et al., 2014). Although predicted, group membership was not a significant predictor of reading comprehension in the current study, but it did find a robust difference between scores on one measure of sentence comprehension immediately post-intervention and at a 9 month follow-up. This, therefore, suggests that improving the ability to divide your attention between processing and storage can have a specific benefit to listening comprehension. As children with language difficulties often have problems processing larger amounts of text, it is thought plausible that children in the current study responded better to sentence comprehension tasks (which were heard rather than read), rather than reading comprehension tasks, which required participants to read short narratives beyond sentence level.

Reading proficiency has also been linked to the ability to form 'Situation Models' of comprehension, a coherent meaning based mental representation of a situation described in a piece of text (De Koning & Van der Schoot, 2013). As explained previously the sample of children in the current study scored very low on reading accuracy measures, with a large majority not scoring enough to attain a standardised score. Due to these reading difficulties, the children in the current study may have found it challenging during reading comprehension tasks to develop a visuospatial representation of the situation, which would have allowed for a better understanding to occur (De Koning & Van der Schoot, 2013). In comparison, sentence comprehension was measured via listening comprehension, with no reading requirement, therefore children in the current study may have been able to form 'Situation Models' easier for sentence comprehension tasks. Also, in a single sentence form, creating the 'Situation Model' could be deemed simpler than when processing a larger narrative. Gernsbacher et al. (1990) discuss that comprehension is a domain general skill and as such the modality of initial information input (listening comprehension or reading comprehension), should not influence the creation of a 'Situation Model'. This could therefore explain why children with language disorder in the current study significantly improved in sentence comprehension ability, compared to significant gains found in reading comprehension for typically developing children by Henry et al. (2014).

Overall, the current study has provided promising results in a very limited field. Working memory intervention studies in children with language difficulties require further exploration. Future research needs to increase participant numbers and if findings are replicated in a further independent study, potentially the intervention could be rolled out over a widespread training program across schools nationally. This would enable teaching assistants and language support staff to implement this training program independently in their own school settings. This would also create further opportunities to deliver this training program within children's everyday context and the potential to transfer improved working memory to 'real world' skills. The schools visited in the current study seemed very excited by this research and it is envisaged that this would be the case in other primary school settings. Teachers in general wholeheartedly recognised the benefits of training working memory in children with language difficulties. Providing a training intervention that was interactive, engaging, enjoyed by the majority of children involved, relatively inexpensive to administer, and minimally disruptive to the school's working day was widely appreciated by all school staff involved in this project. The current research has extended and furthered our understanding of working memory interventions in children with language difficulties and provided the foundation for future research to explore these promising findings further.

References

Acosta, V., Hernandez, S., & Ramirez, G. (2019). Effectiveness of a working memory intervention program in children with language disorders. *Applied neuropsychology. Child*, *8*(1), 15–23.

Adams, C., Hesketh, A., Reeves, D., Cooke, RE., & Crutchley, A. (2001). *Assessment of Comprehension and Expression (6-11)*. No publisher name.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.

Archibald, L. M., & Gathercole, S. E. (2006a). Visuospatial immediate memory in specific language impairment. *Journal of Speech, Language, and Hearing Research, 49*(2), 265–277.

Archibald, L., & Gathercole, S. (2006b). Short-term and working memory in SLI. *International Journal of Language and Communication Disorders*, *41*, 675–693.

Archibald, L., & Joanisse, M. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of speech, language, and hearing research*: JSLHR. 52. 899-914.

Atkinson, R. C., & Shiffrin, R. M. (1968). Chapter: Human memory: A proposed system and its control processes. In Spence, K. W., & Spence, J. T. *The psychology of learning and motivation* (Volume 2). New York: Academic Press. pp. 89–195.

Baddeley, A. (2000). The episodic buffer: New component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423

Baddeley, A. (2003a). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189–208

Baddeley, A. (2003b). Working memory: looking back and looking forward, *Nature Reviews Neuroscience,* 4 (10): 829-839.

Baddeley, A., & Hitch, G. (1974). Working memory. In Bower, G. A. (Ed.), *Recent advances in learning and motivation*, *8*(pp. 47–99). New York: Academic Press.

Baddeley, A. D., & Larsen, J. D. (2007). The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative. *Quarterly Journal of Experimental Psychology*, 60, 497– 504.

Baddeley, A.D., Thomson, N., & Buchanan, M. (1975).Word length and the structure of short-term memory, *Journal of Verbal Learning and Verbal Behavior,* 14 (6) 575-589.

Bellair, J., Clark, S., & Lynham, S. (2014). Can any label work for both intervention and research purposes? Commentary on Bishop, D. V. M., Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders,* 49, 381– 415.

Bishop, D.V. (2003). *Test of Receptive Grammar Version 2* (TROG 2)

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders* 49 (4):381 – 415.

Bishop, D. V. M., Adams, C. V., & Norbury, C. F. (2004). Using nonword repetition to distinguish genetic and environmental influences on early literacy development: A study of 6-year-old twins. *American Journal of Medical Genetics: Neuropsychiatric Genetics*, *129*, 94–96.

Bishop, D.V.M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry*, 37: 391-403.

Bishop D.V.M., Snowling M.J., Thompson, P.A., & Greenhalgh, T. CATALISE consortium. (2016) CATALISE: a multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PeerJ Preprints* 4:e1986v1

Blom, E., & Boerma, T. (2020). Do children with developmental language disorder (DLD) have difficulties with interference control, visuospatial working memory, and selective attention? Developmental patterns and the role of severity and persistence of DLD. *Journal of Speech, Language, and Hearing Research*. 63. 1-15.

Catts, H. W., & Kamhi, A. G. (Ed). (2005). *The connections between language and reading disabilities*, (pp. 25-40). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers

Chiat, S. 2001, "Mapping theories of developmental language impairment: Premises, predictions and evidence", *Language and Cognitive Processes,* vol. 16, no. 2-3, pp. 113-142.

Claessen, M., Leitão, S., Kane, R., & Williams, C. (2013). "Phonological processing skills in specific language impairment", *International Journal of Speech-Language Pathology*, vol. 15, no. 5, pp. 471-483.

Coady, J., & Evans, J. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairment (SLI). *International journal of language & communication disorders / Royal College of Speech & Language Therapists.* 43. 1-40

Cocchini, G., Logie, R.H., Della Sala, S., MacPherson, S.E., & Baddeley, A.D. (2002). Concurrent performance of two memory tasks: evidence for domain-specific working memory systems. *Memory and Cognition,* 30(7):1086-95.

Conrad, R. (1964). Acoustic confusions in the immediate memory. *British Journal of Psychology,* 55: 75–84.

Cornoldi, C., Carretti, B., Drusi, S., & Tencati, C. (2015). Improving problem solving in primary school students: The effect of a training programme focusing on metacognition and working memory. *The British journal of educational psychology.* 85 (3).

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*(2), 163-191.

Dahlin, K. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing*. 24. 479-491.

De Koning, B., & Schoot, M. (2013). Becoming part of the story! Refuelling the interest in visualization strategies for reading comprehension. *Educational Psychology* Review. 25. 261-287.

Della Sala, Gray, Baddeley & Wilson, (1997). Pattern Span

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*, 135–168.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136–1146.

Ebbels, S. (2014), Introducing the SLI debate. *International Journal Language and Communication Disorders*, 49: 377–380.

Ecker, U.K.H., Lewandowsky, S., Oberauer, K., & Chee, A.E.H. (2010). "The Components of Working Memory Updating: An Experimental Decomposition and Individual Differences", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 36, no. 1, pp. 170-189.

Ellis Weismer, S., Evans, J., & Hesketh, L. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 1249–1260

Engle, R. W., & Kane, M. J. (2004). *Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control.* In B. H. Ross (Ed.),*The psychology of learning and motivation: Advances in research and theory, Vol. 44* (p. 145–199). Elsevier Science.

Ericsson, K. A., & Kintsch, W. (1995). *Long-term working memory.* American Psychological Association.

Farquharson, K., & Franzluebbers, C. E. (2014). Comparing the effects of working memory–based interventions for children with language impairment. *EBP Briefs*, *9*(1), 1-11.

Fellman, D., Jylkkä, J., Waris, O., Soveri, A., & Ritakallio, L., Haga, S., Salmi, J., Nyman, T., & Laine, M. (2020). The role of strategy use in working memory training outcomes. *Journal of Memory and Language*. 110.

Fisher, S., Lai, C.S.L., & Monaco, A. (2003). Deciphering the genetic basis of speech and language disorders. *Annual review of neuroscience*. 26. 57-80.

French, J.R.P. (1953). "Experiments in field settings" ch.3 pp.98-135 in Festinger,L., & Katz,D. *Research methods in the behavioral sciences* (New York: Holt, Rinehart & Winston).

Friel-Patti, S., 1999. Specific Language Impairment. *Topics in Language Disorders*, 20(1), pp.1-13.

Gathercole S. E. (2006). Nonword repetition and word learning: The nature of the relationship [Keynote]. *Applied Psycholinguistics*, 27, 513– 543.

Gathercole, S.E., & Baddeley, A.D. (1990). "Phonological memory deficits in language disordered children: Is there a causal connection?, Journal of Memory and Language, vol. 29, no. 3, pp. 336-360.

Gathercole, S.E., & Pickering, S (2001). Working Memory Test Battery (WMTB - C)

Gaulin, C., & Campbell, T. (1994). Procedure for assessing verbal working memory in normal school-age children: Some preliminary data. *Perceptual and motor skills*. 79. 55-64.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 430–445.

Goldstein, S., Naglieri, J.A., & SpringerLink eBook Collection 2014, *Handbook of Executive Functioning,* Springer New York, New York, NY.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6–10.

Graf Estes, K., Evans, J., & Else-Quest, N. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 50. 177-95.

Haines, G. (2015, June 12). The SLI debate: Diagnostic criteria and terminology. Retrieved February 1, 2017, from https://slt.buckshealth.link/wp-content/uploads/2015/07/SLI-debate-12.6.15.pdf

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory.* Wiley.

Henry, L. A. (2001). How does the severity of a learning disability affect working memory performance? Memory, 9, 233-247.

Henry, L. (2012). *The development of working memory in children*. Los Angeles: SAGE.

Henry, L.A., & Botting, N. (2017). Working memory and developmental language impairments. *Child Language Teaching and Therapy*, *33*(1), pp. 19–32.

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry 53,* 37–45.

Henry, L. A., Messer, D. J., & Nash, G. (2014). Testing for near and far transfer effects with a short, face-to-face adaptive working memory training intervention in typical children. *Infant and child development*, *23*, 84-103.

Hick, R., Botting, N., & Conti-Ramsden, G. (2005). Cognitive abilities in children with specific language impairment: Consideration of visuo-spatial skills. *International journal of language & communication disorders / Royal College of Speech & Language Therapists.* 40. 137-49.

Holmes, J. (2012). Working memory and learning difficulties. *Dyslexia Review Summer 2012*, 7-10.

Holmes, J., Butterfield, S., Cormack, F., Van Loenhoud, A., Ruggero, L., Kashikar, L., & Gathercole, S. (2015). Improving working memory in children with low language abilities. *Frontiers in Psychology*. 6:519.

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development*, *77*, 1822–1841.

Jackson, E., Leitao, S., & Claessen, M. (2016). The relationship between phonological short-term memory, receptive vocabulary, and fast mapping in children with specific language impairment. *International journal of language & communication disorders / Royal College of Speech & Language Therapists*. 51.(10)

Jackson, E., Leitao, S., Claessen, M., & Boyes, M. (2019). Fast mapping short and long words: Examining the influence of phonological short-term memory and receptive vocabulary in children with developmental language disorder. *Journal of Communication Disorders.* 79:11-23.

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A. 2008 May 13;105(19)

Kate Nation (2019) Children's reading difficulties, language, and reflections on the simple view of reading, Australian Journal of Learning Difficulties, 24:1, 47-73,

Kim, Y. G. (2016). Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology, 141*, 101-120.

Kwong, T. E., & Varnhagen, C. K. (2005). Strategy Development and Learning to Spell New Words: Generalization of a Process. *Developmental Psychology, 41*(1), 148–159.

Laws, G., & Bishop, D. (2003). A comparison of language abilities in adolescents with Down syndrome and children with specific language impairment. *Journal of Speech Language and Hearing* Research, 46: 1324-1339

Li, D., Christ, S. E., & Cowan, N. (2014). Domain-general and domain-specific functional networks in working memory. *NeuroImage*, *102 Pt 2*(0 2), 646–656.

Leonard, L. (2014). *Children with specific language impairment*. 1st ed. Cambridge, Mass: MIT Press.

Logie, R. (1995). Visual spatial working memory. London: Psychology Press.

Lukács, A., Ladanyi, E., Fazekas, K., & Kemeny, F. (2015). Executive functions and the contribution of short-term memory span in children with specific language impairment. *Neuropsychology.* 30. 10.

Marton, K. (2008). "Visuo-spatial processing and executive functions in children with specific language impairment", International Journal of Language & Communication Disorders, vol. 43, no. 2, pp. 181-200.

Mason, K., Rowley, K., Marshall, C. R., Atkinson, J. R., Herman, R., Woll, B., & Morgan, G. (2010). Identifying SLI in Deaf children acquiring British Sign Language: Implications for theory and practice. *British Journal of Developmental Psychology, 28*, 33-49.

McArthur, G.M., Hogben, J.H., Edwards, V.T., Heath, S.M., & Mengler, E.D. (2000). On the "specifics" of specific reading disability and specific language impairment, *The Journal of Child Psychology and Psychiatry and Allied Disciplines,* vol. 41, no. 7, pp. 869-874

Mcleod, S. (2012, 3rd August). *Working Memory Model | Simply Psychology*. Retrieved from https://www.simplypsychology.org/working%20memory.html

Mcleod, S. (2017,3rd August). Multi store model of memory. Retrieved from, https://www.simplypsychology.org/multi-store.html

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General, 141*(2), 302–320.

Melby-Lervåg M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*. 49(2), 270-91.

Melby-Lervåg, M., Redick, T.S. & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review, *Perspectives on Psychological Science*, vol. 11, no. 4, pp. 512-534.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100.

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology*, *19*, 78-94.

Morrison, A.B., & Chein, J.M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory, *Psychonomic Bulletin & Review*, vol. 18, no. 1, pp. 46-60.

Neale. (1997) Neale Analysis of Reading Ability – Second Revised British Edition (NARA II)

Passolunghi, M.C., & Costa, H.M. (2016). Working memory and early numeracy training in preschool children. *Child neuropsychology : a journal on normal and abnormal development in childhood and adolescence*. 22. 1-18.

Pauls, L. J., & Archibald, L. M. D. (2016). Executive functions in children with specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 59*(5), 1074-1086.

Peng, P., & Fuchs, D. (2015). A randomized control trial of working memory training with and without strategy instruction: Effects on young children's working memory and comprehension. *Journal of learning disabilities.* 50 (1).

Pickering, S. J., & Gathercole, S. E. (2004). Distinctive working memory profiles in children with special educational needs. *Educational Psychology, 24*(3), 393–408.

Pimperton H., & Nation, K. (2014) Poor comprehenders in the classroom: teacher ratings of behavior in children with poor reading comprehension and its relationship with individual differences in working memory, *Journal of Learning Disabilities,* 47(3):199-207.

Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: what does the task measure? *International Journal of Language and Communication Disorders*, *50*(1), pp. 106–118.

Reilly, S., Bishop, D., & Tomblin, B. (2014a). Terminological debate over language impairment in children: forward movement and sticking points. *International Journal of Language & Communication Disorders*, 49(4), pp.452-462.

Reilly, S., Tomblin, B., Law, J., McKean, C., Menash, F. K., Morgan, A., Goldfield, S., Nicholson, J. M., & Wake, M. (2014b). Specific language impairment: A convenient label for whom? *International Journal of Language and Communication Disorders*, *49*, 416-451.

Roche, Jeffrey D. (2014). Cogmed Working Memory Training Product Review. *Journal of attention disorders* (1087-0547), 18 (4).

Roulstone, S., Law, J., Rush, R., Clegg, J., & Peters, T. (2011). Investigating the Role of Language in Children's Early Educational Outcomes. *PsycEXTRA Dataset.*

Rowe, A., Titterington, J., Holmes, J., Henry, L., & Taggart, L. (2019a). Interventions targeting working memory in 4-11 year olds within their everyday contexts: A systematic review. *Developmental Review.* 52. 1-23.

Rowe, A., Titterington, J., & Taggart, L. (2019b). A classroom-based intervention targeting working memory, attention and language skills in 4–5 year olds (RECALL): study protocol for a cluster randomised feasibility trial. *Pilot Feasibility Stud* 5, 82.

Rudner, M., & Rönnberg, J., (2007). The role of the episodic buffer in working memory for language processing, *Cognitive Processing,* vol. 9, no. 1, pp. 19-28.

Sala, G., & Gobet, F. (2017). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology, 53*(4), 671–685.

Schwaighofer, M., Fischer, F., & Bühner, M. (2015) Does Working Memory Training Transfer? A Meta-Analysis Including Training Conditions as Moderators, Educational Psychologist, 50:2, 138-166.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). Clinical evaluation of language fundamentals, fourth edition (CELF-4). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138*(4), 628–654.

Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking*. New York, NY: Oxford University Press

Spaulding, T.J., Plante, E., & Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: evidence for separate attentional capacities. *Journal of Speech Language and Hearing Research,*51(1):16-34.

St Clair-Thompson, H., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology, 30*(2), 203–219.

Stavrakaki, S. (2015). *Specific language impairment current trends in research.* Amsterdam: John Benjamins Publishing Company.

Swanson, H.L., Howard, C.B., & Suez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities,* 39(3): 252–269.

Swanson, H.L., & Jerman, O. (2006) Math Disabilities: A Selective Meta-Analysis of the Literature. *Review of Educational Research*, 76, 249-274.

Swanson, H.L., Xinhua, Z., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: a selective meta-analysis of the literature. *Journal of Learning Disabilities*. 42(3):260-87.

Szewczyk, R. (2016). The effectiveness of working memory training – points to consider for future research, *Polish Psychological Bulletin*, 47 (1): 43-50.

Thompson, T., Waskom, M., Garel, K., Cardenas-Iniguez, C., Reynolds, G., Winter, R., Chang, P., Pollard, K., Lala, N., Alvarez, G. and Gabrieli, J., 2013. Failure of Working Memory Training to Enhance Cognition or Intelligence. *PLoS ONE,* 8(5), p.e63614.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39*(6), 1284-1294.

Uittenhove, K., Chaabi, L., Camos, V., & Barrouillet, P. (2019). Is working memory storage intrinsically domain-specific? *Journal of Experimental Psychology* Gen. 148(11):2027-2057.

Van, Dijk. (1985). Cognitive situation models in discourse production. The expression of ethnic situation models in prejudiced stories. *In: J.P. Forgas, (Ed.) Language and social situations.(pp. 61-79). New York: Springer, 1985.*

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Verbal and Visuo-spatial Working Memory: A Case for Domain-general Time-Based Resource Sharing. *Psychological Science*. 21. 384-390.

Volckaert, A., & Noel, M.P. (2015). Training executive function in preschoolers reduce externalizing behaviors. *Trends in Neuroscience and Education.* 4 (1-2)

Vugs, B.A.M., Cuperus, J.M., Heniks, M.P.H., & Verhoeven, L.T.W. (2013). Visuospatial working memory in specific language impairment: A meta-analysis, *Research in Developmental Disabilities,* 34 (9): 2586-2597.

Wener, S.E., & Archibald, L.M. (2011). Domain-specific treatment effects in children with language and/or working memory impairments: A pilot study. *Child Language Teaching and Therapy*. 2011; 27 (3):313-330.

Wetherell, D., Botting, N. & Conti-Ramsden, G. (2007). Narrative skills in adolescents with a history of SLI in relation to non-verbal IQ scores. Child Language Teaching and Therapy: 23(1), pp. 95-113.

Williams, D., & Lind, S.E. (2013). Comorbidity and diagnosis of developmental disorders. In C. Marshall (Ed.), *Current Issues in Developmental Psychology* (pp. 19-45). Oxfordshire, UK: Taylor & Francis.

Witt, M.U. (2011). School based working memory training: Preliminary finding of improvement in children's mathematical performance. *Advances in Cognitive Psychology, 7*, 7 - 15.

Appendix A – Ethical Approval Letter

CITY UNIVERSITY LONDON

Research Office
Northampton Square
London EC1V 0HB

Tel: +44 (0) 20 7040 5704

www.city.ac.uk

21 January 2016

Dear Emma, Lucy, Shula and David

Reference number: PR/LCS/PhD/15-16/02

Name: Emma Christopher, Lucy Henry, Shula Chiat, David Messer

Title of project: Improving Working Memory in Children with Language Difficulties

Your application has been reviewed and I am happy to approve the application from today's date.

Best of luck with your project.

Please note the above reference number which identifies this application and **must be quoted in all correspondence.**

Kind regards

[redacted]

Language and Communication Science Division

City University London

Northampton Square

London EC1V 0HB

[redacted]

Appendix B – Parent Information Sheet

CITY UNIVERSITY LONDON

Information Sheet

Improving Working Memory in Children with Language Difficulties

Dear Parent,

I am a PhD student from City University London researching whether we can improve working memory in children with language difficulties. I would like to invite your child to take part in my study, but before you decide whether this is something you would be interested in, let me explain what I am doing and why and how this would involve your child.

Please take your time to read the following information carefully and discuss it with the staff at your school. Furthermore, you are also welcome to contact me if you have any questions or would like any further information.

What is the purpose of the study?

The purpose of this study is to assess the benefits of conducting a short face to face working memory intervention with primary aged children who have difficulties with language. Working memory allows us to store information within our minds for short periods of time and then use it for our current thinking. It is often described as our mental workspace, and it is important for learning and educational achievement.

Children with language difficulties commonly find working memory tasks difficult. This research will look at whether working memory can be enhanced through intensive training in activities that tax memory. I would like to visit your child at their school and spend 10 minutes, 3 times a week with them, for 6 weeks in total within their normal school day. I will undertake activities with your child that require the use of their working memory and will then assess the benefits gained through conducting these activities by looking at your child's working memory ability and other aspects of their thinking and learning skills over a 9 month period.

Why have you been invited?

This study is working with schools across Dorset, Hampshire and Wiltshire. Your child has been invited to take part, along with others at his/her school, because your child is within the primary age range and has been identified to us as having language difficulties as a primary need.

Does your child have to take part?

No. Participation in this study is voluntary, so it is up to you and your child to both decide whether or not to take part. If you and your child wish to participate, you will be asked to sign a parental consent form and your child will also be asked to give his/her consent in a child friendly manner. You can both choose to drop out of the study at any time. Withdrawing from the study will not affect your child in any way.

What will happen if your child takes part?

If you and your child decide to take part, I will visit your child at his/her school and carry out some initial language, working memory and educational attainment activities which take approximately 2 hours to complete and can be done over several sessions.

Following this I will then begin the working memory games with your child for 10 minutes 3 times a week; totalling 18 sessions. All sessions will take place during your child's normal school day and no one will visit your home or require you to travel anywhere specific. During these sessions your child will undertake two different working memory tasks which will be fun and set at your child's own level.

Once your child has completed all 18 intervention sessions, follow up sessions will take place with your child at school immediately after the intervention sessions. I will repeat these after 1 term and also after 9 months. These will give us further information regarding your child's language, working memory and educational attainment after the intervention and will take approximately 2 hours to complete.

It is important for us to have an *active control group* in this study as a comparison to the working memory training intervention. Your child will be randomly allocated to 1 of 2 groups during this study. One group will undertake working memory activities during the 18 sessions and one group will be part of the active control group. These children complete very similar activities, but only use processing skills which have no memory requirements. If your child is randomly allocated to this active control group they will still complete 18 sessions with me, have the same pre-intervention, post-intervention, 9 month follow up sessions. They will be unaware of any differences between the task they are undertaking and that of their peers who may be allocated to the working memory intervention group.

Please be assured that if the intervention is effective and your child has been in the active control group that his or her school will be given my full support to use the training with any child they feel would benefit. They will have all resources that have been used throughout the study, and teaching staff will be given a training session on the intervention that I have applied.

What are the benefits of taking part?

- This study will help us learn more about the benefits of working memory training for children with language difficulties.
- It will help us understand which aspects of your child's learning would benefit from working memory training, and the duration that these benefits would last from a relatively short 6 week intervention.
- Your child will hopefully enjoy working with me and also enjoy the games that I will bring to each session.

What are the disadvantages and risks of taking part?

- Your child will be taken out of normal lesson-time. However, I will always discuss this with your child's class teacher and consider the best times to visit. As the working memory intervention process itself has been designed to only be 10 minutes, 3 times a week, this should help to reduce some of the classroom disruption.
- If your child becomes distressed, tired etc. throughout the process or decides that they no longer wish to continue at any point, I will stop the activities immediately and return them to the care of their teacher.

Will your child's data be kept confidential?

Your child's data will be kept entirely confidential. The data will only be accessible by me and my 3 PhD supervisors; Professor Lucy Henry, Professor Shula Chiat and Professor David Messer. Information gathered regarding language, working memory and educational skills will be shared with your child's class teacher as this will be of help when planning classroom activities and in gauging your child's progress.

Your child will be allocated a participation number from the onset of the study and all data, in line with the Data Protection Act, will be stored under this number and not your child's name. Any report written will not mention your child by name and they will not be identifiable under any circumstance during the whole process. All data collected will be kept securely in locked cabinets.

What will happen when the research study stops?

To maintain your confidentiality, when the project is finished all paper records will be shredded. All computer data will be kept anonymously in password protected files, and destroyed after 10 years in line with City University London policy.

What will happen to results of the research study?

The findings of this study will be used to write my PhD thesis. Findings may also be included in articles that are published within journals relating to this study. Please be assured as explained previously; your child will not be named or identifiable in any written work produced.

What will happen if I don't want to carry on with the research study?

If you no longer wish your child to be involved in the study, you are free to withdraw at any time, even if you have already signed a consent form. You do not have to tell us why you do not want to carry on, and it will not affect your child in any way.

What if there is a problem?

If you have any problems, concerns or questions relating to this study, please ask to speak with any member of the research team. If however you would like to speak to someone independent about any aspect of the study, you can contact the University's Senate Research Ethics Committee. You would need to phone 020 70403040 and ask to speak with the Secretary to the Senate Research Ethics Committee and inform them that the name of the project is**: Improving Working Memory in Children with Language Difficulties.**

Alternatively you could also write to the Ethics Secretary:

[redacted], Secretary to Senate Research Ethics Committee, Research Office, E214, City University London, Northampton Square, London, EC1V 0HB

Email: [redacted]

City University London holds insurance policies which apply to this study. If you feel you have been harmed or injured by taking part in this study you may be eligible to claim compensation. This does not affect your legal rights to seek compensation. If you are harmed due to someone's negligence, then you may have grounds for legal action.

Who has reviewed the study?

This study has been approved by the Language and Communication Science Proportionate Review panel, School of Health Sciences Research Ethics Committee, City University London.

Further information and contact details

If you have any questions about any aspects of this study, you can contact me at:

[redacted]

Alternatively you may also contact my PhD supervisors:

Lucy Henry: [redacted]

Shula Chiat: [redacted]

David Messer: [redacted]

Thank you for taking the time to read this information sheet and I look forward to hopefully working with your child on this exciting project in the near future!

Many thanks

[redacted]

Emma Christopher

PhD Student, Division of Language and Communication Science,

City University London

Appendix C – Parental Consent Form

CITY UNIVERSITY
LONDON

Consent Form

Title: Improvement of Working Memory in Children with Language Difficulties

Please tick boxes to confirm that you have read:

1.	I agree to take part in the above City University London research project. I have had the project explained to me, and I have read the participant information sheet, which I may keep for my records. I understand this will involve: • My child taking part in a 6 week working memory training or processing task training intervention programme at their school on a 1:1 basis with the researcher. • My child undertaking assessments in language, working memory and academic attainment prior to the intervention, post-intervention and at 9 month follow up. • Completion of a parental report requiring me to rate statements from 1-4 relating to my child. • Allowing teachers to see my child's scores to help implement future classroom planning and gain a better picture of my child's abilities.	☐
2.	This information will be held and processed for a research study only. I understand that any information I provide is confidential, and that no information that could lead to the identification of any individual will be disclosed in any reports on the project, or to any other party. No identifiable personal data will be published. The identifiable data will not be shared with any other organisation.	☐
3.	I understand that my participation is voluntary, that I can choose not to participate in part or all of the project, and that I can drop out at any stage of the project without being penalized or disadvantaged in any way.	☐
4.	I agree to City University London processing this information about me. I understand that this information will be used only for the purpose(s) set out in this statement and my consent is conditional on the University sticking to its duties and obligations under the Data Protection Act 1998.	☐
5.	English is my main language spoken at home.	☐
6	I agree to take part in the above study.	☐

Name of Participant (Child) ____

Name of Parent/Guardian ____

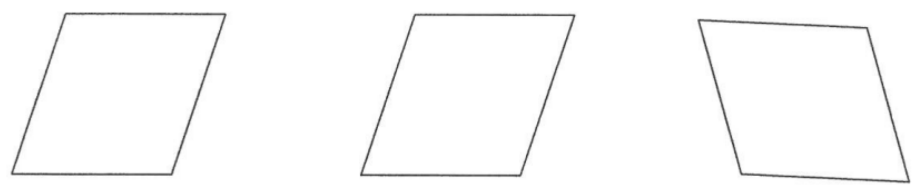
Parent/Guardian Signature ____

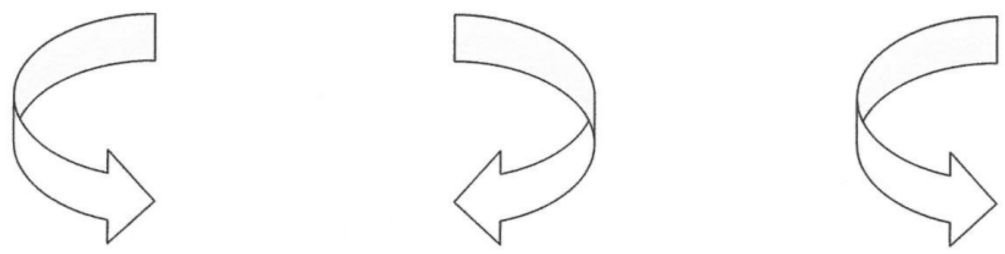
Date____________________________________

Name of Researcher _____________________ Researcher's Signature_____________________
Date ____________________________________

When completed, 1 copy for participant; 1 copy for researcher file.

Appendix D: Example of Odd One Out Card Strips





Appendix E: Example of Odd One Out Response Sheets

Span Level 1

Span Level 2

Appendix F: Training Session Record Form

Working memory training sessions – Record Sheet and Instructions

This training involves administering 11 trials of practice with the selected working memory span task. All trials are presented and scored exactly like the original span test you administered for the pre-test. These brief instructions explain START RULES and MOVING ON RULES. **The key issue is to establish the length of lists (i.e. LEVEL) that will be administered at the start of the training session and how the LEVEL will change across the 11 training trials.**

START LEVEL:

- Start the very first set of training trials at the child's SPAN LEVEL obtained from the pre-test, unless span level was ZERO, in which case start training trials at a level ONE.
- Start each subsequent training session at the span level that was MOST COMMONLY ADMINISTERED in the previous training session (obtain by looking at all span levels administered over the previous 11 training trials)
- Always administer the FIRST TWO training trials at the START LEVEL.

MOVING ON RULES:

To decide what level to administer on each of the remaining 9 training trials, look at the two immediately preceding trials you have just administered and apply the following rules:

- If both trials were fully correct – INCREASE the level by ONE
- If both trials were incorrect – DECREASE the level by ONE
- If one trial was correct and one trial was incorrect – STAY at the same level

CONSTRUCTING TRAINING TRIALS

Refer to the training sentences/odd one out cards provided to construct appropriate lists for your training trials.

- Sentences: We suggest that you draw a line between 'groups' of sentences as you administer them to remind you of the correct items to administer and to provide a visual reminder of the correct answers to help in judging whether the child's responses were correct.
- Odd one out cards: We suggest that you hold the list of cards in your hand (but concealed from the child) in the correct order as a visual reminder of whether the child is pointing to the correct spatial locations and in the correct order.
- **We recommend 18 training sessions, 2 or 3 times per week**

WORKING MEMORY TRAINING TASK:
LISTENING SPAN OR **ODD ONE OUT SPAN**

☐ ☐

CHILD'S NAME:

Training Trial	SPAN LEVEL for administration	CORRECT	INCORRECT
1	Use start level		
2	Use start level		
3	Refer to rules		
4	Refer to rules		
5	Refer to rules		
6	Refer to rules		
7	Refer to rules		
8	Refer to rules		
9	Refer to rules		
10	Refer to rules		
11	Refer to rules		

TODAY'S DATE:

MOST COMMONLY ADMINISTERED SPAN LEVEL (calculate at end of session):

TRAINING SESSION NUMBER (1-18):

Appendix G – Listening Recall Intervention Sentences Examples:

Pigs lay eggs	**F**	**eggs**
Ducks have fur	**F**	**fur**
Walls have ears	**F**	**ears**
Elephants are huge	**T**	**huge**
Elephants live in sand	**F**	**sand**

Tables sing songs	F	songs
Houses can walk	F	walk
The sky is blue	T	blue
Vans grow on trees	F	trees
Lions sail boats	F	boats
Milkmen deliver milk	T	milk
Tables have wings	F	wings
Triangles are round	F	round
Glasses help people to hear	F	hear
Ducks have eight legs	F	legs
Slugs have hands	F	hands
Mice eat cheese	T	cheese

Cows live on farms	**T**	**farms**
Telephones can dance	**F**	**dance**
Girls have wings	**F**	**wings**
People can talk	**T**	**talk**
Trees have leaves	**T**	**leaves**
You eat yogurt with a spoon	**T**	**spoon**
Postmen deliver post	**T**	**Post**
Stars twinkle at night	**T**	**night**
Pianos can talk	**F**	**Talk**
Windows have legs	**F**	**Legs**
You can see with your ears	**F**	**Ears**
Flowers grow on heads	**F**	**heads**
Trees go to bed	**F**	**Bed**
People have two arms	**T**	**Arms**
The moon comes out at night	**T**	**Night**

Appendix H: Mean standardised listening recall scores, standard deviation & range at pre-intervention.

	Mean	*SD*	*Range*
Listening Recall			
Trained Group (n=24)	94.65	16.42	66-`128
Active Control (n=23)	93.36	24.10	57-139