

Article

A short and engaging adaptive working memory intervention for children with Developmental Language Disorder: Effects on language and working memory

Lucy A. Henry ^{1*}, Emma Christopher ², Shula Chiat ³ and David J. Messer ⁴

¹ Division of Language and Communication Science, City, University of London, UK; lucy.henry.1@city.ac.uk

² Division of Language and Communication Science, City, University of London, UK; emma.christopher@city.ac.uk

³ Division of Language and Communication Science, City, University of London, UK; shula.chiat.1@city.ac.uk

⁴ Faculty of Wellbeing, Education & Language Studies, Open University, UK; David.Messer@open.ac.uk

* Correspondence: lucy.henry.1@city.ac.uk

Abstract: Recent research has suggested that working memory training interventions may benefit children with Developmental Language Disorder (DLD). The current study investigated a short and engaging adaptive working memory intervention that targeted executive skills and aimed to improve both language comprehension and working memory abilities in children with DLD. Forty-seven 6- to 10-year-old children with DLD were randomly allocated to an executive working memory training intervention (n=24) or an active control group (n=23). A pre-test/intervention/post-test/9-month-follow-up design was used. Outcome measures included assessments of language (to evaluate far transfer of the training) and working memory (to evaluate near transfer of the training). Hierarchical multiple regression analyses controlling for pre-intervention performance and age found group to be a significant predictor of sentence comprehension and of performance on six untrained working memory measures at post-intervention and 9-month follow-up. Children in the intervention group showed significantly higher language comprehension and working memory scores at both time points than children in the active control group. The intervention programme showed potential to improve working memory and language comprehension in children with DLD and demonstrated several advantages: it involved short sessions over a short period; caused little disruption in the school day; and was enjoyed by children.

Keywords: Working memory training, intervention, developmental language disorder, children

1. Introduction

The current investigation concerns a short and engaging adaptive working memory intervention that aimed to improve both language comprehension and working memory in children with developmental language disorder. As well as looking at ‘far transfer’ to two measures of language comprehension (sentence comprehension and receptive grammar), the study included six measures of ‘near transfer’ to assess whether the intervention led to gains in working memory tasks that were not directly trained.

The working memory system is widely accepted as describing the way information is processed and retained [1] and can be seen as a mental workspace which encompasses several skills used for ‘online’ activities during daily life [1,2]. There are several approaches to understanding working memory [3-6] and one of the most influential is Baddeley and Hitch’s [7] ‘Multi Component Model of Working Memory’, and its subsequent revision by Baddeley [8]. This model consists of several interacting components, including a limited attentional capacity control system, termed the central executive, which is assisted by two further ‘passive’ components, the phonological loop, for holding speech-based information, and the visuospatial sketchpad, for holding visual and spatial information [9,10]. The central executive relies heavily, but not exclusively on the frontal lobes,

[11], and is actively responsible for attentional control, directing and allocating resources in activities when there is a need for the retention of information together with the processing of other information [9,10]. We use the term executive working memory (EWM) to refer to working memory assessments that involve input from the central executive (e.g., concurrent storage and processing) and to interventions that target this component. There also is a fourth component, the episodic buffer, that provides links to long-term memory and some additional multi-modal storage, but this component will not be considered further.

Researchers have emphasised that understanding working memory development is crucial for helping children maximize their intellectual progress [12]. There are significant associations between the components of the working memory system and other cognitive processes, especially those related to language [13], literacy [14] and intelligence [15,16] (but see [17]). It is notable that many groups of children with developmental conditions have significantly lower abilities when assessed on the components of working memory compared to children with typical development (TD) who have a similar chronological or sometimes even mental age [1]. Examples of this are: language disabilities [18]; dyslexia [19]; intellectual disabilities [20]; and autism [21].

These associations between working memory and other cognitive processes, as well as the lower working memory performance of children with developmental conditions, underpin the interest in using interventions to improve working memory and related neuropsychological abilities. An investigation of a computerised working memory intervention targeting the central executive found that, following training, performance on components of the working memory system improved, and importantly, non-verbal intelligence also improved [22]. These effects were reported for children with attention deficit hyperactivity disorder and for a small sample of adults with TD [22]. This initial study was followed by further investigations of this claim, which have been summarised in subsequent systematic reviews and meta-analyses. In general, these reviews have concluded that training can result in near transfer, involving an increase in other working memory abilities that were not part of the intervention procedure, although the evidence for the persistence of these effects is variable. The reviews have also concluded that interventions generally do not result in far transfer to other cognitive abilities such as those associated with mathematics, intelligence, literacy, or language.

However, systematic reviews to date have either been concerned with computer-based interventions that attempt to improve at least one component of the working memory system [23-26] or non-computerised interventions involving mainly children with typical development [27]. Consequently, although these reviews suggest that working memory training does not result in far transfer, their focus has usually been on computer-based interventions or interventions with TD groups, rather than non-computer interventions targeting children with developmental conditions.

It remains important to understand how working memory interventions might benefit children with developmental conditions. Rowe et al. [27] emphasised the potential importance of exploring working memory training interventions that target executive skills with children who have language or other disabilities. This is particularly important as a meta-analysis by Peijnenborgh et al. [28] of working memory interventions with children who have developmental conditions showed that previous research had focussed mainly on children with attention deficit hyperactivity disorder (10/13 studies), or the type of developmental condition was not specified. In addition, several investigations since 2018 challenge the pessimistic conclusions reached in the reviews of this topic. For example, a single case experimental design investigation involving seven children with working memory difficulties [29] used the Cogmed computer-based intervention [30]. This study found near transfer effects for all children, and far transfer in language, reading, or mathematics for three children with 'convincing but modest training effects across multiple measures' [22] (p.1).

Several recent investigations of working memory interventions have concerned children with Developmental Language Disorder (DLD) or a related diagnosis. DLD is a persistent neurodevelopmental condition, involving a primary difficulty with language that is unexplained by any other syndrome or circumstances [31,32]. DLD is estimated to have a prevalence of 7% [33].

Newly emerging evidence suggests a more promising role for working memory interventions in children with DLD. For example, in a study using a pre-test/intervention/post-test design, Holmes et al. [34] used the Cogmed working memory intervention with a sample of children who had Low Language Ability. They identified significant training gains for several working memory components and a far transfer effect to performance IQ, although this effect was non-significant with Bonferroni corrections. Acosta, Hernandez and Ramirez [35] also used a pre-test/intervention/post-test design, developing a working memory training which targeted short-term verbal memory and EWM in both verbal and visuospatial domains, delivered using a combination of individual and group sessions. They found significant training gains in all components of the working memory system for children with language difficulties, as well as far transfer to a lexical-semantic task; they also noted greater gains in their group with language difficulties than in a comparison group with TD. However, neither of these studies included an untreated or active control group, which would provide stronger evidence in favour of the intervention [24].

Acosta-Rodríguez, Ramírez-Santana, and Hernández-Expósito [36] used a comprehensive multi-technique non-computer classroom-based intervention that included executive working memory training as one element. They included children with DLD and TD, allocating half of the children in each group to receive the intervention and half to receive no treatment. Acosta-Rodríguez and colleagues [36] reported immediate post-intervention improvement on several assessments of language comprehension, verbal EWM and semantic fluency in trained children with DLD compared to the untreated DLD control group. The children with DLD receiving the intervention made greater gains than the equivalent typical group. In a single case experimental design, Maleki Shahmahmood et al. [37] used a mixed non-computer and computer-based intervention with 10 Iranian children who were identified with a primary language impairment. The intervention focused on phonological short-term memory and EWM, as well as suggestions about strategies. Following training over five weeks, increases in working memory and morpho-syntactic abilities were reported. A subsequent five-week language intervention also increased morpho-syntactic abilities, but gains were not detected for working memory, suggesting an effect of working memory training on language, but not *vice versa* [37].

In two other investigations, far transfer effects of working memory training have been reported on syntax with French children who have DLD [38,39]. Stanford et al. [38] used iPad or computer-based training (Magic Memory) in groups with DLD and typical development (TD). They also included active control groups who followed an alternative general scholastic training without working memory content, and randomly allocated their participants to each condition. The training process involved practising five verbal working memory tasks assessing two components of the working memory system (verbal short-term memory and EWM). These working memory training tasks were adaptive, such that as the child's performance improved or declined, the task difficulty was adjusted to match the child's current performance. Improvements in both aspects of working memory were found immediately post-intervention for DLD and TD groups. Furthermore, for the DLD group there were improvements in syntax involving 3rd-person accusative clitics, which are often seen as a marker for DLD in French-speaking children [38]. Delage et al. [39] used the same Magic Memory intervention, reporting improved working memory in DLD and TD groups compared to active control groups immediately post-intervention, with additional far transfer gains on sentence repetition and complex syntax in the DLD intervention group. Therefore, the findings reported since 2018 are not only promising support for working memory interventions for children with DLD, but also

suggest that children with DLD may gain more from working memory interventions compared to individuals with TD [24,34-36,38-41].

The present investigation of the effectiveness of an executive working memory intervention for children with DLD builds on these findings. We focussed on children with DLD because of the suggestion that children with developmental conditions may be more likely to benefit from working memory interventions, as there is more room for improvement. Children with DLD are known to have difficulties with working memory, which encompass both short-term and executive working memory [35,42-45]. A face-to-face rather than a computer-based intervention was chosen because, when the study was planned, meta-analyses suggested computer-based forms of intervention were not effective [24,26]. The intervention targeted EWM skills, based on previous evidence that this is the common feature of effective working memory interventions [27]; and evaluated far transfer effects to receptive grammar and language comprehension skills. It was predicted that the intervention would improve receptive grammar abilities, given evidence reported above that working memory interventions can improve morpho-syntax in children with DLD, possibly by increasing information processing abilities. It was also predicted that the intervention would improve language comprehension, based on previous findings and work by Diamond [46] and others [47] who emphasise the strong links between language comprehension and EWM, both of which require the concurrent processing and retention of continuous input that unfolds over time. In this context, Melby-Lervåg et al.'s [24] report of a small significant effect of working memory training on reading comprehension (and arithmetic) is interesting, although when they excluded studies that had a significant decline pre- to post-intervention in relevant control groups the effect was not regarded as noteworthy.

The current intervention followed a programme used by Henry et al. [48], who reported promising longer-term findings for reading comprehension in children with TD. This programme comprised a relatively short, face-to-face, enjoyable, adaptive (an important feature of effective interventions [49]) intervention targeting EWM skills. The intervention employed both verbal and non-verbal EWM tasks as recommended by Danielsson et al. [50], who only found significant training effects to untrained working memory measures in samples with a developmental condition (intellectual disabilities) when a mixed working memory approach was adopted that included both verbal and visuo-spatial components. In the current study, participants were assessed before intervention, and re-assessed immediately post-intervention and nine months later, evaluating maintenance over a longer period than any previous study of children with DLD.

The overall research question was: For children with DLD, are there immediate and longer-term effects of EWM training? We assessed immediate and longer-term effects for directly trained EWM tasks, working memory tasks that were not trained (near transfer) and language comprehension (far transfer).

2. Materials and Method

2.1 Participants

Fifty-one participants, age six to eleven years with language difficulty as a primary need, were recruited from six mainstream primary schools (four with a specialist language provision) across Dorset and Hampshire (UK). Study inclusion criteria were: age 6-11 years; language difficulty identified as a primary need (and two or more language scores at or below one standard deviation of the mean on screening); nonverbal IQ of 70 or above; no formal diagnosis of autism; and English as a first spoken language at home.

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the relevant university Ethics Committee. Written informed consent was sought and obtained from all parents/guardians of participating children, and verbal assent was obtained from the children themselves. Schools recommended suitable children, and following the consent process, screening tasks were undertaken. These included: the four core subtests from Clinical Evaluation of Language Fundamentals 4 (CELF-4-UK)

[51]; the Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI-II) [52] Block Design and Matrix Reasoning subtests to obtain a measure of non-verbal IQ (Perceptual Reasoning); establishing that the child had no formal diagnosis of autism through discussion with the school (to ensure any language difficulty identified could not be related to the autism diagnosis); and ensuring the child’s first language at home was English (to ensure any language difficulty identified could not be due to the child speaking English as an additional language). Two children were excluded due to perceptual reasoning index scores falling below 70; one was undergoing an autism assessment at the time of the intervention; and one did not speak English as a first language. Forty-seven children (mean age = 8 years 0 months, SD = 15 months; range = 6;2 to 10;6) fulfilled the inclusion criteria.

2.2 Random Allocation of Children to groups

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 was aware of the participant’s group membership as practical limitations made it impossible to conduct a double-blind experiment. Screening scores for both groups are given in Table 1, together with independent samples *t*-tests for group differences (effect sizes, Cohen’s *d*, with 95% confidence intervals). No significant group differences on the screening variables were found, although there was a tendency for the active controls to show numerically lower language scores.

Table 1: Mean screening scores (SD) for children in the working memory training and active control groups. Independent samples *t*-tests for group differences are also presented.

Screening scores	Working memory training group (n=24)	Active control group (n=23)	Group differences
Age in months	96.17 (13.96)	96.17 (16.00)	$t(45) = -.002, p=.999, d=.00 (-.57 .57)$
Non-verbal IQ (perceptual reasoning index, WASI-II) ¹	92.04 (14.69)	91.00 (12.44)	$t(45) = .26, p=.795, d=.08 (-.50 .65)$
Concepts and following directions, CELF-4-UK ²	5.12 (3.01)	3.78 (2.88)	$t(45) = 1.56, p=.125, d=.46 (-.13 1.03)$
Word structure, CELF-4-UK ²	4.75 (1.89)	3.78 (2.19)	$t(45) = 1.62, p=.112, d=.47 (-.11 1.05)$
Recalling sentences, CELF-4-UK ²	4.96 (2.56)	4.09 (2.35)	$t(45) = 1.21, p=.231, d=.35 (-.23 .93)$
Formulated sentences, CELF-4-UK ²	3.67 (2.70)	2.52 (1.93)	$t(45) = 1.67, p=.102, d=.49 (-.10 1.07)$

¹Standardised scores (Mean=100; SD=15); ²Scaled scores (Mean=10; SD=3)

2.3 Screening tasks

The CELF-4-UK [51] was administered to determine whether language difficulty was a primary need, as this test is commonly used amongst speech and language professionals. Four language subtests ‘Concepts and Following Instructions’, ‘Word Structure’, ‘Recalling Sentences’ and ‘Formulated Sentences’ were administered to all participants.

Concepts and Following Directions assesses the child’s ability to interpret, recall, and execute oral commands. The child listens to a set of instructions and points to the relevant

pictures, with options presented visually. For example, a child may be asked to 'point to a fish next to a house' or 'point to all the shoes on the top row'. The commands increase in length and complexity and the task stops after seven consecutive scores of zero. Word Structure assesses knowledge of grammatical rules using a sentence completion task. Various grammatical rules are presented (irregular plurals, possessive pronouns, future tenses and irregular past tenses), and the child completes an orally presented sentence related to an illustration, e.g., 'Here is one book. Here are two ____'. There was no discontinuation point and all children were administered 32 sentences. Recalling Sentences evaluates a child's ability to recall and produce sentences varying in length and syntactic complexity. Orally presented sentences are repeated back by the child and scored for errors. Full recall with no errors gains a score of three, one error a score of two, two to three errors a score of one and more than four errors a score of zero. After five consecutive scores of zero, the subtest was discontinued. Formulated Sentences required the child to formulate semantically and grammatically correct sentences using a 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 WASI-II [52] assessed non-verbal IQ using the Block Design and Matrix Reasoning subtests to obtain a perceptual reasoning index (PRI) score. Block Design required a child to analyse and re-produce abstract visual stimuli. There were 13 items, the first four required the design to be reproduced 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. Matrix Reasoning 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.

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

All children were administered outcome assessments before the working memory intervention commenced, immediately after 18 intervention sessions were completed and at a 9-month follow-up. Outcome measures related to direct training effects (the two trained executive working memory tasks), far transfer effects (measures of sentence comprehension and receptive grammar), and near transfer effects (short-term and executive working memory tasks that were not trained). Raw scores are reported for all outcome measures as standardised tests were not available for some tasks (Odd One Out, Pattern Span), and age was controlled in all relevant analyses. See Table 2 for mean scores in both groups at all time points.

2.4.1 Direct effects

Listening Recall from the Working Memory Test Battery (WMTB-C) [53] was administered to assess verbal EWM. The task required the child to listen to an orally presented sentence, say whether the sentence was true or false (e.g., 'Dogs have tails'), and then recall the final word of the sentence ('tails'). It began with two practice trials. Following this, a block of six trials with a list length of one (i.e., one sentence) was presented. If the child obtained at least four out of six trials correct, the task increased in difficulty by presenting two sentences (judgements were made for each sentence, in turn, and then the two final words recalled). Two practice items were administered, followed by a block of six trials at list length two. Again, if four trials out of six were correct, list length increased to three items and so on until fewer than four correct items in a block were achieved, at which point the test was discontinued. If the first four trials in any block were correct, credit was given for trials 5 and 6, and the task moved up a level. Total trials correct scores were used as outcome measures, and span scores (highest list length achieved

successfully) were used for determining the starting point for the working memory intervention. Test-retest reliability ranges from .38 to .84.

The Odd One Out Span task [54] assessed a child's EWM in the visuospatial domain. This task was slightly adapted to be administered on a tablet using a PowerPoint presentation, and the numbers of trials increased to ensure sensitivity at low performance levels. The researcher read through the instructions with the child, and the task began with two practice trials. Each trial showed three black and white images of hard-to-name, but simple forms, presented in the middle of the screen in a horizontal row. The child was asked to point to the 'odd one out' (one item was slightly different to the other two). Following this, the child was shown an empty response grid containing three boxes, horizontally arranged to match the presentation screen, and asked to recall the spatial location of the odd one out by pointing to the correct location on the response grid (left, middle or right). The formal task began with six trials at a list length of one. If the child obtained four or more trials correct, the child was moved up to the next list length of two, i.e., two odd one out judgements, in turn, and a response screen with two empty grids, one on top of the other, for recalling the two spatial locations. The task proceeded with longer list lengths, provided four out of six trials were correct. If fewer than four trials out of six were correct, the task was discontinued. Total trials correct were used as the outcome measure, and span scores (highest list length achieved successfully) were used for determining the starting point for the working memory intervention. Reliability for the original version of this task is .80 [54].

2.4.2 Far Transfer

The Sentence Comprehension subtest from the Assessment of Comprehension and Expression (ACE) [55] was administered to evaluate the ability to understand sentences of increasing length and complexity, requiring the child to decode verbal concepts of quantity, space, time, and emotion. The task has a 'broad floor' so allows even children with very low language abilities to succeed on some items [55]. For many items, 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, e.g., 'The helicopter flew above the clouds'. After hearing the sentence, the child was asked to choose which picture fit the sentence best. For other items, children were presented with one or two sentences supported by a picture, sometimes with a set of printed words round the picture, and asked a question related to these. Responses required either a spoken word or a finger-point to one of the items in the picture or printed words, for example, from a multiple choice of words for feelings 'sad, hungry, cheerful, sleepy' that are spoken by the tester and printed on the page. The subtest had 35 items, which were all administered as there was no discontinuation point. Raw scores were used. Test-retest reliability is .61 and internal consistency is .66.

The Test for Receptive Grammar Version 2 (TROG-2) [56] measures the understanding of grammatical constructs. The test covers 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. A discontinuation rule applied if the child answered five consecutive items incorrectly. Raw scores in terms of number of blocks passed were used. The parallel form reliability for standardised scores on this test is .71.

2.4.3 Near Transfer

Four tasks assessed short-term working memory abilities (two measures of phonological short-term memory and a measure each of spatial and visual short-term memory). For five of the tasks (all from the WMTB-C [53]), the tasks proceeded by presenting blocks of six trials, beginning with list lengths of one item. If four trials out of six were correct within a block, list length increased by one, and so on, until fewer than four correct items

in a block were achieved, at which point the test was discontinued. If the first four trials in any block were correct, credit was given for trials 5 and 6, and the task moved up a level. Total trials correct scores were used as outcome measures for all five of the following tasks to ensure comparability with the non-standardised measures (which took the same format).

Digit Recall (WMTB-C) assessed verbal short-term memory and involved the experimenter reading out a list of numbers which the child repeated back in the correct order. Test-retest reliability for Digit Recall ranges from .81 to .82.

Word List Recall (WMTB-C) assessed verbal short-term memory and involved the experimenter reading out a list of one syllable words which the child repeated back in the correct order. Test-retest reliability for Word List Recall ranges from .64 to .80.

Block Recall (WMTB-C) assessed spatial short-term memory and involved showing the child a board with nine randomly spread out identical raised cubes (3cm x 3cm, coloured grey). The researcher tapped a series of one or more spatial positions on the blocks and asked the child to recall the positions by tapping the same sequence immediately afterwards. Test-retest reliability for Block Recall ranges from .43 to .63.

Counting Recall (WMTB-C) assessed executive working memory and 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 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, as per the other tests. Test-retest reliability for Counting Recall ranges from .48 to .74.

Backward Digit Recall (WMTB-C) assessed executive working memory and involved the experimenter reading out a list of numbers which the child had to repeat backwards. Test-retest reliability for Backward Digit Recall ranges from .53 to .71.

A measure of Pattern Span was developed, based on that of Della Sala et al. [57], to assess visual short-term memory using a similar procedure to the WMTB-C tasks. Pattern span required the child to look at the position of a coloured-in red square on a 2x2 black and white grid, and then recall that position on an identical but empty black and white grid. This task had three trials per block, and each block comprised one list length (list length corresponded to the number of red squares) for one grid size, with list lengths and grid sizes systematically increasing in difficulty as the task progressed. The first three trials (block 1) required the child to recall one red square on a 2x2 grid. The second three trials (block 2) required them to recall two red squares on a 2x2 grid. Next, the grid size was increased to 2x3, with three trials recalling two red squares (block 3). Next the number of squares was increased with three further trials recalling three red squares (block 4). The grid size then increased again to 3x3 and there were a further three trials recalling three red squares (block 5), and three trials recalling four red squares (block 6). This repeating pattern of grid size increases and list length increments continued in blocks of three trials, until the last set of trials required the child to recall seven red squares on a 4x5 grid. The researcher discontinued the task if the child did not score four or more correct in any set of six trials with the same list length (i.e. number of red squares to recall, regardless of grid size). Scores were total trials correct over all blocks. Internal consistency was calculated by determining separate trials correct scores for all trial 1s, all trial 2s and all trial 3s within each block (for a similar approach see [5]). Correlations were $r=.71$ between trial 1 and trial 2 scores, $r=.79$ between trial 1 and trial 3 scores, and $r=.82$ between trial 2 and trial 3 scores).

2.5 Study Procedure

For the screening phase, the researcher visited the child at school and administered the CELF-4-UK subtests and the WASI-II perceptual reasoning scale. Following random allocation to working memory training and active control groups, children completed assessments in working memory (eight tasks), Sentence Comprehension and Receptive Grammar (measures of reading accuracy and comprehension were also administered, but

due to floor effects are not considered further). Next, children received the intervention (see below) and, following the intervention, all outcome measures were administered again. After a delay of around 9-months, outcome measures were administered once more to assess maintenance.

Those assigned into the working memory intervention group undertook the full intervention protocol over a six-week period. During the intervention, all children were visited three times a week within their school setting at the same time each week, in a quiet area either in a separate classroom, library area or intervention room. Each child spent around 10 minutes with the researcher administering 11 trials similar to those in the Listening Recall task [53] and 11 trials similar to those in the Odd One Out Span task [54]. The task was adaptive, reflecting the child's performance from trial-to-trial, therefore, the researcher continuously monitored performance, first to determine previous span levels on each task, second to ensure that the task administered first alternated from session to session, and third to record every span performance score within each training session. The researcher always checked that the child was aware of the task instructions before administering the 11 trials of each training task. The same procedure occurred for the active control group; however, the working memory element was removed from the training tasks. In this way, the inclusion of an active control group should rule out training benefits accruing from global stimulation, individual time with the researcher, or maturation [23,38]. All children in both groups enjoyed the sessions and were happy to attend each time. In the event of school absence, the researcher saw the child on another day such that all children completed all 18 training sessions. Further details about the executive working memory training tasks are provided below.

2.5.1 Listening Recall Training Task (adapted from Listening Recall, WMTB-C [53])

This training task focused on the child's verbal EWM. A different set of listening recall stimuli to the outcome assessment materials was used to eliminate practice effects and overlap.

The listening recall training materials comprised six sets of sentence stimuli for every list length between one and seven, which could be drawn from during the training session to produce trials at the appropriate level. The first training trial began at the child's span level, established from pre-intervention testing. If the span level was previously zero, the researcher started at span level one. The researcher orally presented the requisite number of sentences depending on the start span level, each was judged as true or false by the child, and the child attempted to recall the final word from each sentence. Two trials at the initial span level 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. The training 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.

2.5.2 Odd One Out Span Training Task (adapted from Odd One Out Span [54])

This training task focused on the child's visuospatial EWM and was based on the Odd One Out task. A different set of odd one out stimuli to the outcome assessment materials was used to eliminate practice effects and overlap. The task was presented with paper-based materials, rather than on a tablet computer, but otherwise had the same

format as the Odd One Out task. Odd one out stimuli were each presented as black and white line drawings on 10cm x 30cm white cards. The response card depicted one or more blank horizontal grids, each containing three empty boxes. The child was asked to remember the spatial location of the ‘odd one out’ items by pointing to the correct position/s on the grid (see Figure 1 for examples).

As for the Listening Recall training task, the first training trial began at the child’s span level, established from pre-testing, unless the span level was previously zero, in which case the researcher started at span level one. Again, the working memory training was designed to be adaptive, so increased in difficulty if the child’s memory span performance increased, stayed at the same if the level was exactly at the child’s level, or decreased in difficulty if span had been exceeded. At the end of the 11 training trials, the most frequently presented list length was carried over to the next session as the ‘starting point’. The process was repeated for the next 17 training sessions of the working memory intervention.

Figure 1: Example of ‘Odd One Out’ Card and Response Card



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. Results

3.1 Pre-test scores

Mean raw scores for pre-intervention, post- intervention and 9-month follow-up on all outcome measures are given in Table 2. All scores for working memory tasks represent total trials correct as these scores were comparable across all tasks (standardised and non-standardised) and in some cases standardised scores were unavailable due to low performance. For Sentence Comprehension, we report raw scores; and for Receptive Grammar, total blocks passed.

There were no significant differences between groups on any pre-test scores (see Table 2 for means and SDs):

- Listening Recall, $t(45) = -.23, p = .819$ ($d = -.07$, 95%CI $-.60, .51$);
- Odd One Out, $t(45) = .28, p = .782$ ($d = .08$, 95%CI $-.49, .65$);
- Sentence Comprehension, $t(45) = 1.17, p = .25$ ($d = .34$, 95%CI $-.24, .91$);
- Receptive Grammar, $t(45) = 1.62, p = .113$ ($d = .47$, 95%CI $-.11, 1.05$);
- Digit Recall, $t(45) = .82, p = .415$ ($d = .24$, 95%CI $-.34, .81$);
- Word List Recall, $t(45) = 1.90, p = .064$ ($d = .56$, 95%CI $-.03, 1.14$);
- Block Recall, $t(45) = .37, p = .711$ ($d = .11$, 95%CI $-.46, .68$);
- Pattern Span, $t(45) = .74, p = .465$ ($d = .22$, 95%CI $-.36, .79$);
- Counting Recall, $t(45) = .80, p = .427$ ($d = .23$, 95%CI $-.34, .80$);
- Backwards Digit Recall, $t(45) = .87, p = .389$ ($d = .25$, 95%CI $-.32, .83$).

Scores for Sentence Comprehension and Receptive Grammar were numerically lower in the active control group, in line with the language screening scores. Therefore, all analyses of outcome measures controlled for pre-test performance so that any differences between groups could be attributed to the intervention and not initial levels of performance [58].

3.2 Approach to analyses

Hierarchical multiple regression was undertaken to examine whether group (intervention vs active control) was a significant predictor of post-intervention and 9-month follow-up outcome scores. In all regressions, pre-intervention score for the relevant outcome measure and age in months were controlled at step one. Controlling for pre-intervention outcome scores was a conservative approach as it ensured that any initial variations in performance across participants and groups were taken into account [58]. Age was also controlled as we were using raw scores uncorrected for age. Intervention group, as a dummy variable, was entered at step two of each regression to assess the effects of the intervention.

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). In some regression analyses, one child in the working memory training group improved on an outcome measure more than the model predicted, resulting in a high standardised residual (outlier). When this occurred, regressions were re-run after removing the relevant case, but as there were never any substantive differences in findings, original analyses are reported.

Table 2: Mean raw scores (SDs) for pre-intervention, post-intervention and 9-month follow-up on all outcome measures for children in the working memory training and active control groups.

Outcome variables	Time (pre-intervention, post-intervention or 9-month follow-up)	Working memory training group (n=24)	Active control group (n=23)
Listening recall, WMTB-C	Pre-intervention	9.21 (3.62)	9.52 (5.53)
	Post-intervention	16.33 (3.77)	10.00 (5.71)
	Follow-up	16.38 (4.41)	9.78 (5.13)
Odd One Out Span	Pre-intervention	9.08 (4.10)	8.78 (3.25)
	Post- intervention	18.46 (6.16)	8.52 (3.37)
	Follow-up	18.75 (5.98)	9.48 (2.98)
Sentence Comprehension, ACE	Pre-intervention	21.00 (5.85)	18.91 (6.42)
	Post- intervention	25.79 (4.37)	20.22 (6.42)
	Follow-up	27.46 (4.16)	22.39 (6.05)
Receptive Grammar, TROG-2	Pre-intervention	8.50 (3.58)	6.83 (3.51)
	Post- intervention	11.25 (4.37)	8.35 (3.73)
	Follow-up	12.13 (3.81)	9.35 (3.46)
Digit Recall, WMTB-C	Pre-intervention	25.83 (6.06)	24.57 (4.33)
	Post- intervention	31.13 (6.02)	25.26 (4.32)

Word List Recall, WMTB-C	Follow-up	32.04 (6.19)	24.83 (4.51)
	Pre-intervention	18.83 (3.64)	17.00 (2.91)
	Post- intervention	22.67 (4.13)	17.13 (3.09)
Block Recall, WMTB-C	Follow-up	23.29 (4.18)	17.22 (2.65)
	Pre-intervention	20.67 (6.40)	20.00 (5.83)
	Post- intervention	25.87 (5.90)	21.09 (5.59)
Pattern Span	Follow-up	26.42 (5.60)	21.26 (4.89)
	Pre-intervention	17.29 (5.15)	16.13 (5.64)
	Post- intervention	21.63 (5.09)	17.13 (5.06)
Counting Recall, WMTB-C	Follow-up	22.46 (5.76)	17.74 (5.26)
	Pre-intervention	16.58 (5.49)	15.22 (6.18)
	Post- intervention	22.96 (7.03)	16.30 (6.67)
Backwards Digit Recall, WMTB-C	Follow-up	23.33 (6.83)	16.78 (6.62)
	Pre-intervention	11.21 (5.66)	9.74 (5.92)
	Post- intervention	14.08 (5.74)	9.52 (5.77)
	Follow-up	13.83 (5.65)	9.57 (4.93)

3.3 Direct effects (Listening Recall and Odd One Out)

The first analyses assessed whether training effects were present on the two trained executive working memory tasks. As four separate regression analyses were conducted (two directly trained executive working memory tasks, two time points), Bonferroni corrections were made to the significance levels for the overall models ($p < .01$). See Table 3 for summary regression information.

For post-intervention *Listening Recall*, at step one of the regression, the pre-intervention scores and age accounted for 44.3% of the variance. Examination of the standardised beta-values indicated that only the pre-intervention score was a significant individual predictor, not unexpected as pre-intervention scores are an important predictor for scores after intervention. In step two, adding group produced a significant change in R^2 (ΔR^2), accounting for a further 33.8% of the variance. The final model was significant, $F(3,43) = 51.06$, $p < .001$, predicting 78.1% of the total variance (Adj. R^2 .766). At 9-month follow-up, pre-intervention scores and age accounted for 44.9% of the variance at step one, with only pre-intervention score as a significant individual predictor. Adding group at step two produced a significant change in R^2 (ΔR^2), accounting for a further 35.8% of the variance. The final model was significant, $F(3,43) = 59.89$, $p < .001$, predicting 80.7% of the total variance (Adj. R^2 .793).

The group differences following training on Listening Recall were large, even after controlling for pre-intervention score and age. Inspection of unstandardised Beta values (B) in Table 3 shows that, holding age and pre-intervention score constant, being in the active control group rather than the trained group meant that the Listening Recall trials

correct scores were on average 6.61 points lower (effect size $f^2=.51$) at post-intervention and 6.85 points lower ($f^2=.56$) at 9-month follow-up. Note: for Cohen's f^2 [59] effect sizes are $f^2=.02$ small; $f^2=.15$ medium; $f^2=.35$ large.

For post-intervention *Odd One Out*, at step one the pre-intervention scores and age accounted for 30% of the variance, with pre-intervention score a significant individual predictor. Adding group at step two produced a significant change in R^2 (ΔR^2), accounting for a further 48.3% of the variance; here age became a significant predictor alongside pre-intervention score and group. The final model was significant, $F(3,43) = 51.58, p<.001$, predicting 78.3% of the total variance (Adj. R^2 .767). At 9-month follow-up, pre-intervention scores and age accounted for 31% of the variance, with pre-intervention score a significant individual predictor. Adding group at step two produced a significant change in R^2 (ΔR^2), accounting for a further 47.2% of the variance; here age again became a significant predictor alongside pre-intervention score and group. The final model was significant, $F(3,43) = 51.47, p<.001$, predicting 78.2% of the total variance (Adj. R^2 .767).

The group differences following training for *Odd One Out* were very large, even after controlling for pre-intervention score and age. Inspection of unstandardised Beta values in Table 3 shows that, holding age and pre-intervention score constant, being in the active control group rather than the trained group, meant that *Odd One Out* trials correct scores were on average 9.70 points lower (effect size $f^2=.93$) at post-intervention and 9.04 points lower ($f^2=.90$) at 9-month follow-up.

3.3.1 Summary

The working memory intervention led to substantial and significant gains in performance for the intervention group on two directly trained tasks, Listening Recall and *Odd One Out* span, with very large effect sizes. These findings were present at both time points and were found even after controlling for pre-intervention performance and age.

Table 3: Summary of step 2 of the regressions for the outcome variables Listening Recall and *Odd One Out* (assessing the direct effects of the training intervention).

Predictors at step 2	B (95% CIs)	SE B	β
Listening Recall post-intervention			
Pre-intervention score	.89 (.67, 1.11)	.11	.71***
Age	-.02 (-.09, .05)	.03	-.05
Group ¹	-6.61 (-8.25, -4.98)	.81	-.58***
Listening Recall 9-month follow-up			
Pre-intervention score	.83 (.62, 1.03)	.10	.66***
Age	.02 (-.04, .08)	.03	.05
Group ¹	-6.85 (-8.40, -5.31)	.77	-.60***
Odd One Out post-intervention			
Pre-intervention score	.79 (.50, 1.08)	.15	.41***
Age	.10 (.03, .18)	.04	.22**

Group ¹	-9.70 (-11.7, -7.7)	.99	-.70***
Odd One Out 9-month follow-up			
Pre-intervention score	.79 (.51, 1.07)	.14	.44***
Age	.09 (.02, .16)	.03	.20*
Group ¹	-9.04 (-10.92, -7.15)	.94	-.69***

*** $p < .001$; ** $p < .01$; * $p < .05$

¹Unstandardised betas (B) for Group indicate how much lower the outcome variable score was on average in the active control group compared to the trained group, holding the other variables constant (95% confidence intervals for these B-estimates are given in brackets).

3.4 Far Transfer effects

We next assessed whether there were far transfer effects to Sentence Comprehension and Receptive Grammar skills, either immediately post-intervention or at 9-month follow-up. As four separate regression analyses were conducted (two far transfer variables, two time points), Bonferroni corrections were made to the significance levels for the overall models ($p < 0.01$). See Table 4 for summary regression statistics.

For post-intervention *Sentence Comprehension*, at step one, pre-intervention scores and age accounted for a substantial 68.5% of the variance, with only pre-intervention score a significant individual predictor. Of particular interest was the effect of adding group in step two, which produced a significant change in R^2 (ΔR^2), accounting for a further 10.4% of the variance. The final model was significant, $F(3,43) = 53.61$, $p < .001$, predicting 78.9% of the total variance (Adj. R^2 .774). At 9-month follow-up, pre-intervention scores and age again accounted for a substantial 66.2% of the variance at step one, with only pre-intervention score a significant individual predictor. Again, in step two, adding group produced a significant change in R^2 (ΔR^2), accounting for a further 9.7% of the variance. The final model was significant, $F(3,43) = 45.08$, $p < .001$, predicting 75.9% of the total variance (Adj. R^2 .742).

Thus, even with our conservative approach of controlling for the substantial variance accounted for by pre-intervention scores and age, training group was a significant predictor of Sentence Comprehension both immediately and nine months following the intervention. Inspection of unstandardised Beta values in Table 4 shows that, when holding age and pre-intervention score constant, being in the active control group rather than the trained group meant that Sentence Comprehension scores were on average 3.98 points lower (effect size $f^2 = .12$) at post-intervention and 3.60 points lower ($f^2 = .11$) at 9-month follow-up. The gain after training approached a medium effect size. If we consider the average pre-intervention Sentence Comprehension score in our trained group (raw score = 21), together with the average age (8 years 0 months), a raw score of this value gives a scaled score of 6 according to the manual [55]. An increase of 4 points over the control group at post-intervention following training would result in a raw score of 25, with an increase in the scaled score from 6 to 9 (a jump from 9th to 37th percentile).

For post-intervention *Receptive Grammar*, pre-intervention scores and age accounted for 54% of the variance in post-intervention scores at step one, with pre-intervention score a significant individual predictor. In step two, adding group did not produce a significant change in R^2 (ΔR^2), accounting for only a further 2.5% of the variance. The final model was significant, $F(3,43) = 18.59$, $p < .001$, predicting 56.5% of the total variance (Adj. R^2 .534). At 9-month follow-up, pre-intervention scores and age accounted for 58.1% of the variance at step one, with pre-intervention score a significant individual predictor. In step two, adding group, failed to produce a significant overall change in R^2 (ΔR^2), accounting for a further 3.5% of the variance. The final model was significant, $F(3,43) = 22.93$, $p < .001$, predicting 61.5% of the total variance (Adj. R^2 .589).

Training group was not a significant predictor of Receptive Grammar at post-intervention or 9-month follow-up, although children in the intervention group did obtain numerically higher scores. Inspection of unstandardised Beta values in Table 4 shows that, holding age and pre-intervention score constant, being in the active control group rather than the trained group meant that blocks passed scores on Receptive Grammar were on average 1.40 points lower (effect size $f^2 = .03$) at post-intervention and 1.50 points lower ($f^2 = .04$) at 9-month follow-up (neither group difference was significant). In the latter case, however, there was a marginally significant effect of group ($p < .055$) with a small effect size.

3.4.1 Summary

The working memory intervention led to significant gains in the intervention group in Sentence Comprehension (three to four points in the raw score) which approached a medium effect size. These findings were evident both immediately post-intervention and

at 9-month follow-up and were found even after controlling for pre-intervention performance and age. No significant training group differences were found for Receptive Grammar.

Table 4: Summary of step 2 of the regressions for the far transfer outcome variables (ACE Sentence Comprehension and TROG-2 Receptive Grammar).

Predictors	B (95% CIs)	SE B	β
Sentence Comprehension post-intervention			
Pre-intervention score	.76 (.60, .93)	.08	.77***
Age	-.001 (-.07, .07)	.03	-.002
Group ¹	-3.98 (-5.72, -2.24)	.86	-.33***
Sentence Comprehension 9-month follow-up			
Pre-intervention score	.70 (.54, .87)	.08	.76***
Age	.00 (-.07, .07)	.03	.001
Group ¹	-3.60 (-5.34, -1.85)	.87	-.32***
Receptive Grammar pre-intervention to post-intervention			
Pre-intervention score	.89 (.57, 1.23)	.16	.76***
Age	-.03 (-.11, .05)	.04	-.11
Group ¹	-1.40 (-3.21, .41)	.90	-.17
Receptive Grammar pre-intervention to 9-month follow-up			
Pre-intervention score	.76 (.48, 1.05)	.14	.71***
Age	.002 (-.07, .07)	.03	.006
Group ¹	-1.50 (-3.04, .04)	.76	-.20 ($p=.055$)

*** $p < .001$; ** $p < .01$; * $p < .05$

¹Unstandardised betas (B) for Group indicate how much lower the outcome variable score was on average in the active control group compared to the trained group, holding the other variables constant (95% confidence intervals for these B-estimates are given in brackets).

3.5 Near Transfer effects (untrained measures of working memory)

Also of interest was the possibility of near transfer effects to related working memory tasks that were not directly trained, both immediately after the intervention and at 9-month follow-up. We assessed four short-term working memory measures (Digit Recall, Word List Recall, Block Recall, and Pattern Span) and two executive working memory measures (Counting Recall and Backwards Digit Recall). As 12 separate regression analyses were conducted (six working memory variables, two time points), Bonferroni corrections were made to the significance level for the overall models ($p < 0.004$). See Table 5 for

regression summaries for the short-term memory measures and Table 6 for regression summaries for the executive working memory measures.

For the short-term working memory measures, all regression models at post-intervention and 9-month follow-up were significant ($F_s(3,43)$ ranged between 20.24 and 54.00, all $p_s < .001$). Pre-intervention scores and age accounted for between 40.6% and 67.4% of the variance at step one in all eight models, and only pre-intervention scores were significant individual predictors in each case. At step two, adding group produced a significant change in R^2 (ΔR^2) in all eight models, accounting for additional variance of between 10.8% and 24.0%. Being in the trained group was associated with significantly higher scores on all four working memory assessments at both time points. The total variance accounted for in the models ranged from 58.5% to 79.0%.

Group differences associated with being in the working memory training group were medium to medium-large across all near transfer short-term memory measures, even after controlling for pre-intervention scores and age. Inspection of unstandardised Beta values in Table 5 shows that, holding age and pre-intervention score constant, being in the active control group rather than the trained group meant that trials correct scores on short-term working memory tasks were on average between 3.73 and 5.06 points lower at post-intervention ($f^2=.22, .24, .14$, and $.13$ for Digit, Word List, Block and Pattern) and between 3.90 and 6.29 points lower at 9-month follow-up ($f^2=.31, .30, .20$, and $.12$ for Digit, Word List, Block and Pattern).

For the two untrained executive working memory measures, all regression models at post-intervention and 9-month follow-up were significant ($F_s(3,43)$ between 36.87 and 56.15, all $p_s < .001$). Pre-intervention scores and age accounted for between 62.2% and 71.0% of the variance at step one, and only pre-intervention scores were significant individual predictors in each case. At step two, adding group produced a significant change in R^2 (ΔR^2) in all models, accounting for additional variance of between 7.5% and 12.9%. For both EWM tasks at both time points, being in the trained group was associated with significantly higher scores. The total variance accounted for in these models ranged from 72.0% to 79.6%.

Group differences associated with being in the working memory training group were medium to medium-small for the near transfer EWM measures, even after controlling for pre-intervention scores and age. Inspection of unstandardised Beta values in Table 5 shows that, holding age and pre-intervention score constant, being in the active control group rather than the trained group meant that trials correct scores on Counting Recall and Backwards Digit Recall were on average 5.44 and 3.38 points lower respectively at post-intervention ($f^2=.15$ for Counting Recall and $f^2=.08$ for Backwards Digit Recall); and 5.34 and 3.17 points lower at 9-month follow-up ($f^2=.15$ for Counting Recall and $f^2=.08$ for Backwards Digit Recall).

3.5.1 Summary

The working memory intervention led to significant gains for the intervention group in all six untrained working memory tasks, encompassing both short-term and executive working memory, usually with medium effect sizes. This provided strong evidence for near transfer to similar working memory tasks. These findings were evident both immediately post-intervention and at 9-month follow-up and were found even after controlling for pre-intervention performance and age.

Table 5: Summary of step 2 of the regressions for the short-term memory near transfer outcome variables (Digit Recall, Word List Recall, Block Recall and Pattern Span).

Predictors	B (95% CIs)	SE B	β
------------	-------------	------	---------

Digit Recall post-intervention

Pre-intervention score	.63 (.39, .88)	.12	.56***
Age	.03 (-.06, .11)	.04	.07
Group	-5.06 (-7.42, -2.69)	1.17	-.43***

Digit Recall 9-month follow-up

Pre-intervention score	.73 (.50, .96)	.11	.59***
Age	.02 (-.06, .11)	.04	.06
Group	-6.29 (-8.52, -4.06)	1.11	-.49***

Word List Recall post-intervention

Pre-intervention score	.77 (.53, 1.01)	.12	.57***
Age	.03 (-.03, .08)	.03	.08
Group	-4.12 (-5.70, -2.54)	.78	-.46***

Word List Recall 9-month follow-up

Pre-intervention score	.81 (.59, 1.02)	.11	.59***
Age	.01 (-.04, .06)	.02	.03
Group	-4.60 (-6.01, -3.19)	.70	-.50***

Block Recall post-intervention

Pre-intervention score	.74 (.55, .93)	.10	.73***
Age	.00 (-.08, .08)	.04	.00
Group	-4.30 (-6.43, -2.16)	1.06	-.35***

Block Recall 9-month follow-up

Pre-intervention score	.68 (.49, .86)	.09	.71***
Age	-.02 (-.09, .06)	.04	-.05
Group	-4.70 (-6.73, -2.68)	1.01	-.41***

Pattern Span post-intervention

Pre-intervention score	.78 (.62, .93)	.08	.76***
Age	.04 (-.02, .09)	.03	.10

Group	-3.73 (-5.27, -2.19)	.77	-.34***
Pattern Span 9-month follow-up			
Pre-intervention score	.83 (.65, 1.01)	.09	.75***
Age	.03 (-.03, .10)	.03	.08
Group	-3.90 (-5.70, -2.10)	.89	-.33***

* $p < .05$; ** $p < .01$; *** $p < .001$

¹Unstandardised betas (B) for Group indicate how much lower the outcome variable score was on average in the active control group compared to the trained group, holding the other variables constant (95% confidence intervals for these B-estimates are given in brackets).

Table 6: Summary of step 2 of the regressions for the executive working memory near transfer outcome variables (Counting Recall, Backwards Digit Recall).

Predictors at step 2	B (95% CIs)	SE B	β
Counting Recall post-intervention			
Pre-intervention score	.89 (.68, 1.10)	.10	.68***
Age	.09 (.004, .17)	.04	.17*
Group	-5.44 (-7.54, -3.34)	1.04	-.36***
Counting Recall 9-month follow-up			
Pre-intervention score	.89 (.66, 1.11)	.11	.69***
Age	.05 (-.04, .14)	.04	.10
Group	-5.34 (-7.62, -3.06)	1.13	-.36***
Backwards Digit Recall post-intervention			
Pre-intervention score	.81 (.60, 1.01)	.10	.76***
Age	.01 (-.07, .09)	.04	.02
Group	-3.38 (-5.38, -1.38)	.99	-.28**
Backwards Digit Recall 9-month follow-up			
Pre-intervention score	.75 (.58, .91)	.08	.76***
Age	.04 (-.03, .10)	.03	.09

Group	-3.17 (-4.78, -1.57)	.80	-.28***
-------	----------------------	-----	---------

* $p < .05$; ** $p < .01$; *** $p < .001$

¹Unstandardised betas (B) for Group indicate how much lower the outcome variable score was on average in the active control group compared to the trained group, holding the other variables constant (95% confidence intervals for these B-estimates are given in brackets).

4. Discussion

This study investigated whether an adaptive intervention targeting executive working memory would benefit children with developmental language disorder, as evidenced by improvements not only on the two trained tasks (direct effects) and six untrained working memory tasks (near transfer effects), but also on tests of Sentence Comprehension and Receptive Grammar (far transfer effects). Outcomes were measured immediately post-intervention and nine months later. To specifically evaluate the contribution of an executive working memory training intervention, the study randomly assigned participants to a ‘trained’ group who received the executive verbal and visuospatial tasks (Listening Recall and Odd One Out) and an ‘active control’ group who received the same tasks but without the executive element. Hierarchical multiple regression analyses controlling for pre-intervention performance and age found group to be a significant predictor of Sentence Comprehension at both time points. These findings indicated that children in the trained group obtained significantly higher scores on Sentence Comprehension than children in the active control group immediately following training, and that these gains were maintained nine months later. We estimated how the intervention affected the children’s percentile scores on Sentence Comprehension. Calculations suggested that the intervention could result in a change from the 9th to 37th percentile, an appreciable gain, as after the intervention scores would reach the low-typical range. Performance on Receptive Grammar, however, did not show these gains, with only a marginal group effect at 9-month follow-up.

In terms of direct and near transfer effects, group was a significant predictor of performance on both the trained executive working memory tasks (Listening Recall and Odd One Out) and performance on all other ‘near transfer’ working memory tasks (Digit Recall, Word List Recall, Block Recall, Pattern Span, Counting Recall and Backward Digit Recall) immediately post-intervention and nine months later. For all tasks, children in the trained group obtained significantly higher scores than children in the active control group at both time points, even after controlling for pre-intervention performance and age.

The findings of positive near and far transfer effects of the working memory training in this study run counter to earlier systematic reviews that have concluded that working memory interventions fail to produce far transfer effects. However, these reviews were mostly based on studies of children with TD [23-27]. In contrast, our findings add to more recent studies reporting far transfer to language performance, usually with children who have developmental conditions such as language difficulties [29,35-39]. It is notable that the present study achieved these positive outcomes following training on just two executive working memory tasks (one verbal and one visuospatial) administered in 18 sessions of around 10 minutes three times a week (total 3 hours). Other studies of children with DLD deployed more intervention tasks and included both short-term memory and executive working memory components (e.g., five computer-based short-term and executive working memory tasks in the Magic Memory programme [38,39] or multiple live memory tasks [35]; and total intervention times ranging from 12-18 hours across studies). The present study is also unique in conducting a longer-term follow-up and finding that gains were retained nine months after the short, targeted intervention.

How might we account for the positive near and far transfer effects of this intervention? Face-to-face delivery was used on the grounds that social engagement may increase

children's motivation and focused attention to input, providing favourable conditions for take-up of the training. This may be particularly important for the executive-load element of EWM tasks that the trained group received, differentiating this from the judgement subtasks that the active control group received. Indeed, children in the present study enjoyed taking part, all willingly completing the 18 sessions, and children in the trained group were keen to improve their previous score in both tasks, in effect competing with themselves. However, far transfer to language performance has also been observed in recent studies using computer-based training that, like the present training, appears to turn relatively pure working memory tasks into engaging activities [38,39]. This suggests that the focus and/or enjoyment of the activity and motivation rather than mode of delivery may be a key factor in positive outcomes. In addition, the adaptive nature of the executive task, ensuring children are largely successful, while encouraging them to perform just above their current level of success, may be an important factor in their engagement and motivation and is noted by many authors as a key ingredient for the success of working memory interventions [38,39].

The effect of group on performance both post-intervention and nine months later indicates that the EWM element was the active ingredient in the intervention, since the active control group experienced the same input and judgement task as the trained group but without the executive load that also required recall of the final word in each sentence (verbal intervention) or location of the 'odd one out' (visuospatial intervention). This outcome is in line with the prediction that an EWM training would benefit language comprehension on the grounds that both cognitive activities require simultaneous processing and storage of information [46,47]. In sentence comprehension, words need to be retained in short-term memory while also processing their meaning and the meaning of the sentence. Furthermore, most items in the Sentence Comprehension task used in this study required the child to select from a set of pictures the one that matched a spoken sentence (apart from the final, most difficult items that required inferencing from a sequence of sentences presented with pictorial support). To succeed, the child must not only recognise, understand and retain all elements of the verbal input, but must at the same time scan the visual input to find the picture that depicts all elements of the sentence, avoiding the distraction of pictures that share some but not all of these elements. The implication is that the intensive experience of simultaneous processing and storage in the training tasks improved executive management, as evidenced by the post-training gains on these tasks, to the benefit of the executive-loaded Sentence Comprehension.

It is notable that the other measure of far transfer to language comprehension (Receptive Grammar from the Test for the Reception of Grammar [56]), did not show significant improvement in the trained group, although there was a trend in this direction at 9-month follow-up. On our proposed interpretation of training effects, we might infer that the gains in executive management following training were not sufficient to meet the executive demands of the TROG verbal/visual material, or that the TROG includes grammatical structures that children have not acquired. These possibilities are plausible since the blocks of items testing grammatical constructs on the TROG become increasingly complex both conceptually and syntactically. Another possibility is that the very stringent scoring on the TROG, with children only receiving a point for a grammatical construct if they respond correctly to all four items in the block targeting that construct, raised the bar for achieving statistically significant changes in scores.

The attribution of gains in Sentence Comprehension to improved EWM is in line with the account of training-induced changes in working memory put forward by Gathercole et al. in a systematic review of working memory training effects [60]. According to their account, such changes arise from 'novel cognitive routines that control the sequence of cognitive processes required to perform the task', and far transfer effects only arise where the directly trained task yields a new cognitive routine that can be profitably applied to a far transfer task. However, such routines will not benefit tasks for which well-established processes are used, amongst which Gathercole et al. [60] include verbal short-term

memory (STM) tasks that rely on phonological coding. This notion of 'new cognitive routines' receives some support from the overt strategies that some children in this study developed and used in the Odd One Out task although the use of strategies was neither suggested nor encouraged by the researcher. One child used body parts, for example shoulders and head, to represent the location of left, right and middle shapes, and performed the resulting sequence of gestures in visuospatial recall; another child tapped different fingers for different locations and used the resulting sequence of taps in recall. Such strategies, like silent verbal rehearsal and chunking used in verbal recall tasks, may be construed as 'new cognitive routines' that support recall capacity.

On Gathercole's [60] account, near transfer effects to non-executive (STM) tasks would also be attributed to new cognitive routines induced by the training, rather than changes in children's basic STM capacity. The implication is that increased auditory/visual attention and/or deployment of strategies induced by the training enabled children to use their basic STM capacity more effectively (rather than directly increasing this). On this interpretation, we might expect less benefit for a nonword repetition task, which requires immediate coding and production of a novel phonological form, potentially drawing on long-term knowledge of lexical phonology, but not amenable to strategies such as rehearsal or chunking [61]. However, increases in nonword repetition have been reported following working memory training in children with language difficulties [37], indicating that this would be an interesting avenue to explore further.

What are the wider implications of the view that EWM interventions effect change in executive control of information processing that may benefit performance on language comprehension and also non-executive working memory tasks? Consider everyday verbal input that children receive, for example a sequence of verbal instructions from a teacher, or someone recounting an event or experience. Such input is rapid and transitory so must be processed and interpreted 'online'. Comprehension therefore relies on *immediate* verbal STM and language processing, both drawing on established language knowledge. If the training has enabled children to deploy their STM capacity more effectively, as suggested above, this may extend the amount of verbal input they can process and recognise, and thereby optimise the deployment of their language knowledge in online comprehension. If the training has enhanced their capacity to store and process information simultaneously, as suggested by gains on the executive working memory tasks, this may support better retention of processed input while processing subsequent input, and thereby improve their comprehension of stretches of verbal discourse. It should particularly benefit activities that do not involve rapid and transitory input, for example, reading comprehension, and school-based learning that requires close attention to and integration of new pieces of information.

Having reported the positive outcomes of this study, it is important to acknowledge its limitations. Most notably, due to resource limitations, the sample size was moderate and the researcher who delivered the training also administered all the assessments so was not blind to the child's group. This may have led to researcher bias resulting in more favourable outcomes for the trained group. It is worth pointing out that outcomes in this group varied, and that some children showed little improvement. Nevertheless, replication of the study with larger samples and blinded assessment is needed to preclude possible bias. Further research is also needed to investigate whether there are wider benefits for language performance, of the sort proposed above, in the short and longer term.

5. Conclusions

Working memory training that targeted executive working memory skills was found to have far, near and direct transfer effects in children with developmental language disorder, not only immediately post-intervention but also nine months later. This indicates its potential as an intervention to improve working memory and language comprehension in these children. The intervention programme involved short sessions over a short period, causing little disruption in the school day, and children enjoyed the activities. The

programme is relatively straightforward to administer and could be delivered by trained teaching support staff. The findings from this study invite independent evaluation with blind assessment, and further research to throw more light on which children benefit, which language and other skills show immediate benefit, and whether benefits are sustained in the even longer-term.

Author Contribution: Conceptualization, Lucy Henry, Emma Christopher, Shula Chiat and David Messer; Data curation, Lucy Henry and Emma Christopher; Formal analysis, Lucy Henry, Emma Christopher and David Messer; Funding acquisition, Lucy Henry; Investigation, Lucy Henry, Emma Christopher, Shula Chiat and David Messer; Methodology, Lucy Henry, Emma Christopher, Shula Chiat and David Messer; Project administration, Emma Christopher; Resources, Lucy Henry and Emma Christopher; Supervision, Lucy Henry, Shula Chiat and David Messer; Validation, Lucy Henry, Emma Christopher, Shula Chiat and David Messer; Visualization, Lucy Henry, Emma Christopher, Shula Chiat and David Messer; Writing – original draft, Lucy Henry, Emma Christopher, Shula Chiat and David Messer; Writing – review & editing, Lucy Henry, Emma Christopher, Shula Chiat and David Messer.

Funding: This research was funded by a PhD Studentship awarded to the second author by City, University of London.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Language and Communication Science Proportionate Review Ethics Committee, City, University of London (Ref: PR/LCS/PhD/02, date: 21.01.2016).

Informed Consent Statement: Written informed consent was sought and obtained from all parents/guardians of participating children, and verbal assent was obtained from the children themselves.

Data Availability Statement: Before publication, the original data will be uploaded to the Open Science Framework.

Acknowledgments: We would like to extend our warm gratitude to the children, teachers and families who took part in this study, and to the participating schools for so generously facilitating this research.

Conflicts of Interest: The authors declare no conflict of interest. The funder had no role in the study design; the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

References

1. Henry, L.A., Moran, A., & Messer, D. (2020). Working-memory Development. In Hupp, S., & Jewell, J.D. (Eds.). *The Encyclopedia of Child and Adolescent Development*. New York: Wiley-Blackwell.
2. Holmes, J. (2012). Working memory and learning difficulties. *Dyslexia Review Summer 2012*, 7-10.
3. Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104, 163-191. doi: [10.1037/0033-2909.104.2.163](https://doi.org/10.1037/0033-2909.104.2.163)
4. Cowan, N. (2005). *Working memory capacity*. Hove, UK: Psychology Press.
5. Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, 128, 309-331. doi: [10.1037/0096-3445.128.3.309](https://doi.org/10.1037/0096-3445.128.3.309)
6. Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211. doi: [10.1037/0033-295X.102.2.211](https://doi.org/10.1037/0033-295X.102.2.211)
7. Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In Bower, G. A. (Ed.), *Recent Advances in Learning and Motivation* (Vol 8) (pp. 47-89). New York: Academic Press. doi: [10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
8. Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417-423. doi: [10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
9. Baddeley, A. (2003a). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189-208. doi: [10.1016/S0021-9924\(03\)00019-4](https://doi.org/10.1016/S0021-9924(03)00019-4)
10. Henry, L.A. (2012). *The development of working memory in children*. London: SAGE.

11. Baddeley, A. (2003b). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829-839. doi: 10.1038/nrn1201
12. Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333, 959-964. doi: 10.1126/science.1204529
13. Archibald, L.M. (2017). Working memory and language learning: A review. *Child Language Teaching and Therapy*, 33, 5-17. doi: 10.1177/0265659016654206
14. Peng, P., & Fuchs, D. (2017). 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, 62-80. doi: 10.1177/0022219415594609
15. Cochrane, A., Simmering, V., & Shawn Green, C. (2019). Fluid intelligence is related to capacity in memory as well as attention: Evidence from middle childhood and adulthood. *PLoS ONE*, 14, e0221353. doi: 10.1371/journal.pone.0221353
16. Swanson, H.L. (2008). Working memory and intelligence in children: What develops? *Journal of Educational Psychology*, 100, 581-602. doi: 10.1037/0022-0663.100.3.581
17. Burgoyne, A.P., Hambrick, D.Z., & Altmann, E.M. (2019). Is working memory capacity a causal factor in fluid intelligence? *Psychonomic Bulletin & Review*, 26, 1333-1339. doi: 10.3758/s13423-019-01606-9
18. Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2020). Working, declarative, and procedural memory in children with Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, 63, 4162-4178. doi: 10.1044/2020_JSLHR-20-00135
19. Smith-Spark, J.H., & Fisk, J.E. (2007). Working memory functioning in developmental dyslexia. *Memory*, 15, 34-56. doi: 10.1080/09658210601043384
20. Carretti, B., Belacchi, C., & Cornoldi, C. (2010). Difficulties in working memory updating in individuals with intellectual disability. *Journal of Intellectual Disability Research*, 54, 337-345. doi: 10.1111/j.1365-2788.2010.01267.x
21. Habib, A., Harris, L., Pollick, F., & Melville, C. (2019). A meta-analysis of working memory in individuals with autism spectrum disorders. *PLoS ONE*, 14, e0216198. doi: 10.1371/journal.pone.0216198
22. Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, 24, 781-791. doi: 10.1076/jcen.24.6.781.8395
23. Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49, 270-91. doi: 10.1037/a0028228
24. Melby-Lervåg, M., Redick, T.S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of 264 "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science*, 11, 512-534. doi: 10.1177/1745691616635612
25. Sala, G., & Gobet, F. (2017). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology*, 53, 671-685. doi: 10.1037/dev0000265
26. Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist*, 50, 138-166. doi.org/10.1080/00461520.2015.1036274
27. Rowe, A., Titterton, J., Holmes, J., Henry, L., & Taggart, L. (2019). Interventions targeting working memory in 4-11 year olds within their everyday contexts: A systematic review. *Developmental Review*, 52, 1-23. doi: 10.1016/j.dr.2019.02.001
28. Peijnenborgh, J.C., Hurks, P.M., Aldenkamp, A.P., Vles, J.S., Hendriksen, J.G. (2016). Efficacy of working memory training in children and adolescents with learning disabilities: A review study and meta-analysis. *Neuropsychological Rehabilitation*, 26, 645-72. doi: 10.1080/09602011.2015.1026356
29. Pauls, L.J., & Archibald, L.M.D. (2022). Cognitive and linguistic effects of working memory training in children with corresponding deficits. *Frontiers in Education*, 6. doi: 10.3389/educ.2021.812760
30. Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., & Dahlström, K. (2005). Computerized Training of Working Memory in Children with ADHD—A Randomized, Controlled Trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44, 177-186. doi:10.1097/00004583-200502000-00010
31. Bishop, D.V.M., Snowling, M.J., Thompson, P.A., Greenhalgh, T., & the CATALISE-2 consortium (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, 58, 1068-1080. doi: 10.7287/peerj.preprints.2484v2
32. McGregor, K.K., Goffman, L., Van Horne, A.O., Hogan, T.P., & Finestack, L.H. (2020). Developmental language disorder: Applications for advocacy, research, and clinical service. *Perspectives of the ASHA Special Interest Groups*, 5, 38-46. doi: 10.1044/2019_PERSP-19-00083
33. Norbury, C.F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, 57, 1247-1257. doi: 10.1111/jcpp.12573
34. 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. doi: 10.3389/fpsyg.2015.00519
35. Acosta, V., Hernandez, S., & Ramirez, G. (2019). Effectiveness of a working memory intervention program in children with language disorders. *Applied Neuropsychology*, 8, 15-23. doi: 10.1080/21622965.2017.1374866

36. Acosta-Rodríguez, V.M., Ramírez-Santana, G.M., & Hernández-Expósito, S. (2022). Intervention for oral language comprehension skills in preschoolers with developmental language disorder. *International Journal of Language & Communication Disorders*, 57, 90-102. doi: 10.1111/1460-6984.12676
37. Maleki Shahmahmood, T., [Soleymani Z.](#), [Meysami, A.](#), [Mashhadi, A.](#), & [Nematzadeh, S.](#) (2018). Cognitive and language intervention in primary language impairment: Studying the effectiveness of working memory training and direct language intervention on expansion of grammar and working memory capacities. *Child Language Teaching and Therapy*, 34, 235-268. doi: [10.1177/0265659018793696](#)
38. Stanford, E., Durrleman, S., & Delage, H. (2019). The effect of working memory training on a clinical marker of French-speaking children with developmental language disorder. *American Journal of Speech-Language Pathology*, 28, 1388-1410. doi: 10.1044/2019_AJSLP-18-0238
39. Delage, H., Stanford, E., & Durrleman, S. (2021). Working memory training enhances complex syntax in children with Developmental Language Disorder. *Applied Psycholinguistics*, 42, 1341-1375. doi:10.1017/S0142716421000369
40. Lanfranchi, S., & Carretti, B. (2016). Improving Working Memory in Learning and Intellectual Disabilities. *Frontiers in Psychology*, 7, 725. doi: [10.3389/fpsyg.2016.00725](#)
41. Søndergaard Knudsen, H.B., & Jensen de López, K.M. (2021). Face-to-face working memory training does not enhance children's reading comprehension – a pilot study with Danish children. *Nordic Psychology*, 73, 211-225. doi: [10.1080/19012276.2020.1856001](#)
42. Henry, L.A. & Botting, N. (2017). Working memory and developmental language impairments. *Child Language Teaching and Therapy*, 33(1), 19-32. doi:10.1177/0265659016655378
43. 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. doi: [10.1044/1058-0360\(2009/09-0028\)](#)
44. Vugs, B., Cuperus, J., Hendriks, M., & Verhoeven, L. (2013). Visuospatial working memory in specific language impairment: A meta-analysis. *Research in Developmental Disabilities*, 34, 2596-2597. doi: [10.1016/j.ridd.2013.05.014](#)
45. Vugs, B., Hendriks, M., Cuperus, J., & Verhoeven, L. (2014). Working memory performance and executive function behaviors in young children with SLI. *Research in Developmental Disabilities*, 35, 62-74. doi: [10.1016/j.ridd.2013.10.022](#)
46. Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135-68. doi: 10.1146/annurev-psych-113011-143750
47. Daneman, M. & Merikle, P.M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3, 422-433. doi: [10.3758/BF03214546](#)
48. 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. doi: [10.1002/icd.1816](#)
49. Holmes, J., Gathercole, S.E., & Dunning, D.L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, 12, F9-F15. doi: 10.1111/j.1467-7687.2009.00848.x
50. Danielsson, H., Zottarel, V., Palmqvist, L., & Lafranchi, S. (2015). The effectiveness of working memory training with individuals with intellectual disabilities – a meta-analytic review. *Frontiers of Psychology*, 6, 1230. doi: 10.3389/fpsyg.2015.01230
51. Semel, E., Wiig, E.H., & Secord, W.A. (2003). *Clinical Evaluation of Language Fundamentals (CELF-4)*. San Antonio, TX: The Psychological Corporation.
52. Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI –II)*. London: Pearson Assessment.
53. Pickering, S.J., & Gathercole, S.E. (2001). *The Working Memory Test Battery for Children*. Hove: The Psychological Corporation.
54. Henry, L.A. (2001). How does the severity of a learning disability affect working memory performance? *Memory*, 9, 233-247. doi: [10.1080/09658210042000085](#)
55. Adams, C., Hesketh, A., Reeves, D., Cooke, R.E., & Crutchley, A. (2001). *Assessment of comprehension and expression (6-11)*.
56. Bishop, D. (2003). *Test for Reception of Grammar (TROG-2)*. London: Pearson Assessment.
57. Della Sala, S., Gray, C., Baddeley, A.D., Allamano, N. & Wilson, L. (1997). Pattern is span: A tool for unwelding visuo-spatial memory. *Neuropsychologia*, 37, 1189-1199.
58. Hatcher, P.J., Hulme, C., Miles, J.N., Carroll, J.M., Hatcher, J., Gibbs, S., Smith, G., [Bowyer-Crane, C.](#), & Snowling, M.J. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: A randomised controlled trial. *Journal of Child Psychology and Psychiatry*, 47, 820-827. doi: [10.1111/j.1469-7610.2005.01559.x](#)
59. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Hillsdale, NJ Lawrence Erlbaum Associates, Publishers.
60. Gathercole, S.E., Dunning, D.L., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of Memory and Language*, 105, 19-42. doi: [10.1016/j.jml.2018.10.003](#)
61. Chiat, S. (2015). Non-word repetition. In Armon-Lotem, S., de Jong, J., & Meir, N. (Eds.), *Methods for assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 125-150). Bristol, United Kingdom: Multilingual Matters.