

A classroom intervention targeting working memory, attention and language skills: a cluster randomised feasibility trial

Anita Rowe (✉ harron-a@ulster.ac.uk)

University of Ulster <https://orcid.org/0000-0002-2758-5585>

Jill Titterington

University of Ulster at Jordanstown

Joni Holmes

MRC Cognition and Brain Sciences Unit

Lucy Henry

City University

Laurence Taggart

University of Ulster

Research

Keywords: working memory, classroom interventions, dosage, working memory, attention, language, feasibility

Posted Date: February 8th, 2020

DOI: <https://doi.org/10.21203/rs.2.22975/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published on February 6th, 2021. See the published version at <https://doi.org/10.1186/s40814-021-00771-w>.

Abstract

Background: International debate around the best models of speech and language therapy provision for children with language disorders has highlighted the need for research into classroom-based approaches and intervention dosage. Working memory (WM) is a cognitive skill that underpins attention and language. 'Recall to Enhance Children's Attention, Language and Learning' (RECALL) is a novel, six-week, classroom-based intervention targeting WM to enhance attention and language skills in 4-5 year olds.

Methods : A cluster randomised feasibility trial was conducted to address uncertainties around: the feasibility of a definitive trial to evaluate RECALL; the acceptability of RECALL to the health professionals and teachers who would deliver it; and the factors that may impact on the fidelity of its delivery in the classroom. Six classes of 4-5 year olds participated: 2 received RECALL; 2 received an existing intervention targeting attention skills (not underpinned by WM theory); and 2 received education as usual (no intervention). Ten children in each class (n= 60) were sampled to assess the appropriateness of the outcome measures that may be used in a definitive trial. A process evaluation included: observations of the fidelity of the intervention delivery and semi-structured interviews with the health professionals (HPs) and teachers who delivered RECALL.

Results: The recruitment targets were achieved and all six schools completed the trial. For the experimental RECALL intervention, 95% of sessions were delivered but fidelity to the intervention protocol varied between the two schools (76% versus 45%). The interview data revealed mixed findings regarding the acceptability of the intervention tasks and the outcome measures. A greater understanding of the theory underpinning RECALL would have enhanced the teachers' fidelity to its delivery.

Conclusions: The trial processes (e.g., recruitment and consent procedures) could be easily scaled-up into a future definitive trial but the RECALL intervention requires modification to enhance its acceptability. Large class sizes, child and facilitator factors impacted on the dose (number of practice items) accessed by individual children, particularly those most at risk. This study highlights the need for thorough training for teachers and HPs engaged in the delivery of classroom interventions for children with language disorders. Trial registration: ISRCTN13633886 . Registered 7 Sept 2018 . Keywords: working memory, classroom interventions, dosage, working memory, attention, language, feasibility

Key Messages Regarding Feasibility:

This study addressed the acceptability of a novel intervention for children with language disorders, and the feasibility of delivering it to whole classes of 4–5 year old children.

The fidelity of the intervention delivery, and its potential effectiveness, could be optimised if it were repackaged as a small group intervention.

The methods used to record children's progress during the intervention period should be modified prior to a full trial.

Background

The use of classroom-based interventions

Worldwide, there has been debate around the best models of Speech and Language (SLT) provision for school-aged children who are at risk of language disorders, particularly those from areas of social disadvantage (SD) where high proportions of children present with impoverished language skills on school entry [1-2]. SLT services are increasingly providing collaborative, classroom-based interventions but there is a lack of research-based evidence for this approach [3]. This raises important questions about whether valuable and limited resources are being used in the most efficient way [3-5]. Due to the role SLTs have in early intervention and prevention for language disorders, there is a need for ecologically valid research (conducted in real-life contexts) to provide an evidence-based practice approach [6-7].

Context of the current study

The current study was conducted in the real-life context of health and education services in one region of the UK, Northern Ireland (NI), where there are high rates of SD associated with educational underachievement [8-9]. Extending the role of health professionals within early intervention and integrated service provision across the health and education sectors is a key strategy [10] that aims to harness health professionals' (HPs) specialist knowledge of the developmental skills that form the foundation for learning (e.g., language and motor skills) to enhance educational practice. The Regional Integrated Support for Education (RISE) teams are based in five Health and Social Care Trusts (HSCTs) that provide integrated health and social care services across NI. The RISE teams include: speech and language therapists (SLTs), occupational therapists (OTs), physiotherapists (PTs) and social, emotional and behavioural specialists (SEBs). They provide: individualised (specialist) support for children referred by their teachers; and whole class (targeted and universal) interventions that aim to prevent potential future difficulties for at risk children (i.e., non-referred children) [3]. The teams work in a transdisciplinary model in which professionals jointly plan and deliver interventions. Within the classroom context this also involves teachers and classroom assistants. Overall, this approach aims

to maximise clinical and cost effectiveness by enhancing the holistic nature of interventions and streamline the clinical pathway for children by professionals sharing their expertise [11].

In NI children commence formal education at four years of age, and the mainstream school population includes a wide range of children including those with undiagnosed and diagnosed intellectual and / or developmental difficulties. The majority of children referred to the RISE teams are 4-5 year olds (year 1 pupils in the UK) from schools in areas of SD. Attention and language difficulties are most frequently cited as the reason for referral [12]. Current support, developed and provided by the RISE teams, is a whole-class intervention targeting attention skills: the Attention and Listening Programme (ALP). This intervention has not been evaluated robustly and, unlike the new intervention developed in this study, it is not underpinned by WM theory.

Rationale for developing an intervention that targets WM

Working memory (the ability to hold in mind and mentally manipulate information over short periods of time in the face of distraction) is a cognitive skill linked to both everyday attentional skills and language development [13-14]. Interventions aimed at improving WM may therefore enhance these closely related real-world skills [15]. However, the potential for WM interventions, and in particular computer-based training programmes, to improve untrained WM tasks and real-world skills (transfer effects) has been widely debated [16].

Overall, the evidence indicates that existing computer-based training programmes consistently produce gains on the trained tasks and closely related memory tasks [17-18]. It has been suggested that to improve the therapeutic value of WM training, it may be necessary to embed it within typical classroom activities that are ecologically valid [19]. To test this, a new intervention was developed in the current study: 'The Recall to Enhance Children's Attention Language and Learning' (RECALL) programme. This is a theoretically underpinned, evidence-based intervention that targets WM in 4–5-year-old children through group and whole-class activities over a 6-week period. It is designed to be delivered by HPs from the RISE teams and teachers, and was co-produced with a group of these practitioners through a series of interactive workshops [20-22].

Rationale for conducting a feasibility trial

Existing collaborative practice between the RISE teams and schools in NI provided the optimal setting to develop and evaluate the RECALL intervention. Prior to conducting a definitive trial of RECALL, it was crucial to conduct a feasibility trial to deepen the understanding of the intervention [23], and to test whether it could be run with children as young as 4 years in a group setting. Previous evidence for WM classroom-based WM training interventions have been run with older children (6-7 year olds) who were trained on a one-to-one basis [24]. Little is currently known about the optimal intervention dosage (amount and intensity) required in classroom-based interventions for children with language disorders, or the factors that might impact on this [25-26]. A goal of this study was to evaluate these issues in a feasibility trial of RECALL.

Study aims and objectives

The specific aims of the study were:

- 1) To determine the feasibility of conducting a definitive cluster randomised trial (CRT) evaluating whether RECALL is more effective than an existing intervention (ALP), and 'education as usual', in 4-5 year olds from areas of SD;
 1. a) To understand trial processes including: recruitment, consent and sampling; blinding; attendance, and loss to follow-up.
 2. b) To determine the appropriateness of the outcome measures for the children, teachers and HPs.
- To explore the acceptability of RECALL to HPs and teachers who deliver classroom-based interventions in mainstream schools.
- To measure the compliance and fidelity of the intervention delivery.

Methods/design

This section provides a summary of the study design, methods of the feasibility trial, and the interventions implemented (RECALL and RISE). The study was designed and is reported according to the CONSORT 2010 extension to cluster randomised pilot and feasibility trials [27]. The trial was registered with the International Standard Randomised Controlled Trial Registry (ISRCTN13633886). Comprehensive details can be accessed in the study protocol associated with this trial [28].

Study design

This was a three-arm, cluster randomised feasibility trial with a parallel group design that took place in two HSCT areas in NI. Two classes of 4-5 year olds were randomly allocated to each arm of the trial: i) RECALL (experimental condition); ii) the existing ALP intervention developed by the RISE teams (active control condition); and iii) education as usual (no intervention condition). The experimental RECALL and active control interventions were delivered by HPs from the RISE teams once per week, followed up by two practice sessions delivered by teachers. Children's outcomes were measured at baseline and 1-week post-intervention by Research Assistants (RAs).

Participants

The target population were: HPs from the RISE teams (SLTs, OTs, PTs and SEBs) who had experience of delivering classroom-based interventions and were not involved in the co-production of RECALL; mainstream primary schools situated in areas of SD based on data from the NI Multiple Deprivation Measure [29]; teachers who had not previously received the ALP intervention; and children in year 1 classes (4-5 year olds).

Sample size and procedure

The recruitment targets were: eight HPs from the RISE teams and six schools. The aim was to recruit one class of 4-5 year olds in each school ($n \sim 30$) with a stratified sample of 10 children in each class ($n = 60$ in total) for outcome measurement. The aim was to recruit a sample representing the typical range of ability in mainstream schools: i) children about whom teachers have concerns around listening and communication skills but do not have a diagnosed developmental or intellectual difficulty ($n = 5$ per school, $n = 30$ in total); ii) children with diagnosed developmental or intellectual difficulties ($n = 2$ per class, $n = 12$ in total); and iii) typically developing children who do not have any identified listening and communication problems as recognised by the teachers ($n = 3$ per class, $n = 18$ in total).

Randomisation and blinding

Randomisation took place at the school level after baseline data collection. The school names were placed in opaque envelopes that were randomly selected and allocated by the third author (overseen by the second author). The HPs were not blinded to the schools' allocation as they inevitably knew which intervention they were delivering to which school. The school participants (principals, teachers and parents), and the RAs who conducted the outcome measurement with the children were blind to intervention groupings.

Interventions

The experimental RECALL and active control (ALP) interventions were both 6-week interventions consisting of 40-minute sessions repeated 3 times per week. The HPs delivered the first session each week and demonstrated the activities for the teachers who provided two further practice sessions during the week (18 sessions in total). Details of these interventions are reported here according to the Template for Intervention Description and Replication (TIDieR) Checklist [30].

1. **Experimental intervention: RECALL** This novel intervention targets WM explicitly and is based on a systematic review of evidence suggesting that repeated practice on certain (non-computerised) activities can improve WM and have the potential to produce effects on untrained WM skills (near-transfer) and real-world skills such as attention and language (far-transfer) [31]. The common ingredient across the effective interventions was the executive- loaded nature of the trained task i.e., training on a task that taps into attentional and processing resources under executive control and not just the storage of information.

RECALL includes 3 executive-loaded tasks with specified dosage and task progression (Table 1). Each session starts with a whole-class activity in which a fantastical theme is introduced for that week using a puppet e.g., space. This is based on evidence that fantastical play supports children's WM [32]. The class is then divided into 3 groups (of 9-10 children) that rotate around the three tasks, namely, listening recall, odd one out and phoneme awareness tasks (described in Table 1). The HPs attended a 2-day training course prior to delivering RECALL (provided by the first author). The training aimed to enable the HPs to deliver the first intervention session each week in the classroom, thereby modelling the activities for the teachers who were to provide two further practice sessions per week. This method of cascading training is carried out routinely in the context of the RISE teams and schools. The decision to only provide direct training for the HPs (and not for the teachers) was based on evidence from a qualitative study conducted prior to the intervention development which highlighted that, due to resource constraints, it was highly unlikely that teachers would be released from their everyday duties to attend training (in preparation). The HPs and teachers were provided with a detailed manual including the theory underpinning the intervention.

Executive-loaded task	Dosage	Task progression
Listening recall [24] <ul style="list-style-type: none"> - Targets verbal ELWM. - The children listen to a short sentence, judge whether it is true or false, then recall the last word of the sentence 	11 trials (practice items) per session.	The number of to-be-remembered words increases from one word in week one to two words by week 6.
Odd one out [24] <ul style="list-style-type: none"> - Targets verbal ELWM - The children look at three pictures in a grid, decide where one is the odd one out is (left, middle or right), then recall the location of the odd one out picture 	11 trials per session.	The number of to-be-remembered locations increases from one in week one, to three or four by week 6.
Phoneme awareness [33-34] <ul style="list-style-type: none"> - Targets the ability to isolate and manipulate sounds in spoken words e.g., identifying the first sound in a word 	10-15 minutes per session.	Difficulty increases from alliterative matching to blending onset and rime. Each task progresses from early to late developing phonemes based on typical speech sound development.

Table 1. RECALL components, dosage and task progression

1. **Active control intervention: ALP** This pre-existing programme was informally developed by the RISE teams and aims to improve attention and listening skills through: repeated practice of listening tasks and teaching children the importance of listening through visual and verbal cues. It is not underpinned by WM theory and does not require the children to recall verbal or visuospatial information.

1. **No intervention control: Education as usual** These schools did not receive any classroom-based interventions such as RECALL or ALP during the 6-week trial period.

Outcome measures

The primary outcomes relate to the feasibility of the trial processes and the acceptability of RECALL. The main feasibility outcomes were the rates of recruitment, consent and retention in terms of the number and proportion of participants at each stage in the study. The acceptability of RECALL was explored through semi-structured interviews with the HPs and teachers who delivered it (see Appendix 2). These were audio-recorded, transcribed verbatim and analysed using Braun and Clarke's (2006) [35] approach to thematic analysis.

The secondary outcomes were children's skills ($n=60$) at baseline and one-week post-intervention. These determined the acceptability of the outcome measures. Following good practice in WM research [36], this included standardised assessment of: i) the trained tasks (listening recall, odd one out and phoneme awareness); ii) the untrained WM tasks (near-transfer); and iii) attention and language skills (far-transfer effects). Table 2 details the assessments used.

During the baseline assessment phase, the RAs felt that the children appeared to be performing close to ceiling level on two of the tasks, leading to concerns regarding the sensitivity of the measures to detect change as a result of treatment. This raised questions about the appropriateness of both the phoneme isolation subtest of the Preschool and Primary Inventory of Phonological Awareness (PIPA) [37] for phoneme awareness and the comprehension scale of the New Reynell Developmental Language Scales (NRDLS) [38] for language. Consequently, two alternative measures were trialled in one randomly selected school ($n= 10$ children): the phoneme segmentation subtest of the PIPA for phoneme awareness; and the Clinical Evaluation of Language Fundamentals- Preschool (CELF-P) [39] for language. These two subtests were not included in the measures administered in the other 5 participating schools.

The children's performance from week to week on the trained tasks was also monitored in order to explore the appropriateness of the difficulty level of the tasks (Table 2). The children completed the odd one out and phoneme awareness tasks in individual booklets using crayons and stampers to indicate their responses. For verbal tasks (listening recall and some phoneme awareness tasks) the use of individual digital voice recorders was trialled with five randomly selected children in one of the RECALL classes.

Outcome measured	Skill	Standardised assessment
Trained task	Trained WM tasks	<p>Automated Working Memory Assessment (AWMA) [40]</p> <ul style="list-style-type: none"> · A computerised assessment administered using a laptop · 2 subtests administered in all 6 schools ($n=60$ children): - Listening recall - Odd one out
Trained task	Phoneme awareness	<p>The Preschool and Primary Inventory of Phonological Awareness (PIPA) [37]</p> <ul style="list-style-type: none"> · A standardised assessment consisting of 6 subtests for children aged 3 years to 6 years 11 months · 2 subtests trialled: - Phoneme isolation subtest (administered in 5 schools, $n=50$ children) - Phoneme segmentation subtest (administered in 1 school, $n=10$ children)
Near-transfer	Untrained WM tasks	<p>Automated Working Memory Assessment (detailed above) [40]</p> <ul style="list-style-type: none"> · 4 further subtests administered in all 6 schools ($n=60$ children): - digit recall - block recall - counting recall - non-word recall
Far-transfer	Attention	<p>NEPSY-II – A Developmental Neuropsychological Assessment (NEPSY) [41]</p> <ul style="list-style-type: none"> · Includes standardised performance-based measures of attention for children under 6 years · 2 subtests administered in all 6 schools ($n=60$ children) - Auditory attention - Statue
	Language	<p>The New Reynell Developmental Language Scales (NRDLS) [38]</p> <ul style="list-style-type: none"> · A standardised assessment for children aged between 3 years and 7 years 6 months. · Comprehension scale administered in 5 schools ($n=50$ children) <p>Clinical Evaluation of Language Fundamentals- Preschool (CELF-P) [39]</p> <ul style="list-style-type: none"> · A standardised assessment for 3 – 6 year olds that examines children’s: understanding and use of syntax (grammar/sentence structure), semantics (word meanings) and grammatical morphology (markers of grammatical relationships) · Core language subtests ($n=10$) conducted in 1 school ($n=10$)
		Behaviour in the classroom
	Communication skills at home	<p>The Focus on Communication Outcomes Under Six – 34 (FOCUS-34) [43] ($n=60$)</p> <ul style="list-style-type: none"> · A checklist of children’s communication skills at home completed by parents to measure change over time · Completed by parents in all 6 schools ($n=60$ children)

Table 2. Standardised assessments trialled at baseline and post-intervention

Process Evaluation

A process evaluation was conducted in parallel to the feasibility trial [44]. This was based primarily on the framework for the design and reporting of process evaluations of cluster randomised trials [45]. It also included elements of Steckler and Linnan's (2002) model [46] that are relevant to the delivery of classroom-based interventions. These included: i) the consideration of context (local factors that influence implementation); ii) fidelity (the extent to which the intervention is delivered as conceived); iii) the dose delivered (the amount of intervention offered to participants); and iv) the dose accessed by individuals (the extent of participants' engagement in the intervention).

Compliance and fidelity to the intervention protocol were measured through observations of three RECALL sessions in each school (one delivered by the HPs and two by the teacher), each carried out by the first author. One session in each school was observed simultaneously and rated independently by the second or third authors. Fidelity was scored using a structured checklist based on Carroll et al.'s 2007 framework (see appendix 1) [47]. These data were integrated with the findings from the semi-structured interviews with the HPs and teachers who delivered RECALL to determine the overall feasibility and acceptability of intervention.

Ethical approval

Ethical approval was granted by the Ulster University Research Ethics Committee (REC/18/0036) and approval was obtained from the relevant HSCT research offices.

Results

This section presents the participant characteristics and results in relation to the three key aims: the feasibility of conducting a definitive CRT; the acceptability of RECALL; and the compliance and fidelity of the intervention delivery.

Participant characteristics

Table 3 provides details of the number and characteristics of the schools (clusters) and the individual participants recruited to the study compared to the recruitment targets.

Participants	Recruitment targets	Number recruited	Characteristics
Health professionals	<i>n</i> = 8	<i>n</i> = 8	Professional background: SLT (<i>n</i> = 4) OT (<i>n</i> = 2) PT (<i>n</i> = 1) SEB (<i>n</i> = 1)
Schools (clusters)	<i>n</i> = 6	<i>n</i> = 6	Social disadvantage ranking (based on data from the NIMDM 2017 [29]): Within lowest decile for their HSCT area (<i>n</i> = 3) Within lowest quintile for their HSCT area (<i>n</i> = 3)
Children recruited for outcome measurement	<i>n</i> = 60	<i>n</i> = 60	Gender: girls (<i>n</i> = 26, 43%); boys (<i>n</i> = 34, 57%) Age at baseline: 56 months to 67 months (mean = 61 months)
	<i>n</i> = 30 (50% of sample)	<i>n</i> = 22 (37%)	1) children about whom teachers had concerns around listening and communication skills
	<i>n</i> = 12 (20%)	<i>n</i> = 12 (20%)	2) children with diagnosed developmental or learning difficulties
	<i>n</i> = 18 (30%)	<i>n</i> = 26 (43%)	3) typically developing children who did not have any identified listening and communication problems as recognised by the teachers

Table 3 Participant characteristics

Feasibility of conducting a definitive CRT

-

Recruitment, consent and sampling

Figure 1 shows the study flow chart including the response, recruitment and retention rates throughout the study. The recruitment targets were met in terms of: HPs (*n* = 8); schools (*n* = 6); and the total number of children (*n* = 60) (Table 3). Due to staff absence (maternity leave/sick leave) the RISE teams could only facilitate the study in particular geographical sectors within their HSCT areas. Consequently, from the list of schools identified in areas of SD (*n* = 43), a considerable number (*n* = 17) had to be excluded on the basis of their location. As a result, the criteria in respect of SD was widened to include schools ranked within the lowest quintile within the HSCT (rather than the lowest decile). The overall rate of parental consent (72%) was good. However, some parents of children about whom teachers had concerns did not consent and the desired proportion of children in this sub-group was not achieved (*n* = 22, 37% compared to the target of *n* = 30, 50%). It was also apparent during the sampling process that teachers did not always know whether children did/did not have a diagnosis.

Blinding

The outcome assessors (RAs) remained blind to the intervention groups but, due to the nature of the intervention and materials provided, the teachers in RECALL became aware of their allocation.

Attendance and loss to follow up

No schools or individual participants dropped out of the study. Two children did not complete post-intervention assessments as they were absent from school, indicating minimal loss to follow up (3%).

Figure 1. RECALL cluster randomised feasibility trial flow char (following CONSORT guidance, 2010) [27]

Appropriateness of outcome measures

The acceptability of the outcome measures used with the children was considered from two perspectives: i) the ease of administration and scoring; and ii) the appropriateness of the tests in terms of their psychometric properties for assessing WM, attention and language in the population of interest (4-5 year olds in areas of SD).

Ease of administration

Two methods were used to monitor the children's progress from week to week. For the odd one out task and some of the phoneme awareness tasks, the children each had an individual booklet and they marked their response using stampers. The data gathered through the observations of RECALL in the classroom and via the semi-structured interviews with the HPs and teachers indicated that this approach was not acceptable. The children needed help to turn the pages of the booklets so the HPs or teachers had to repeatedly pause the task to ensure all of the children were on the right page. The children were distracted by the stampers and tended to stamp ad hoc in their booklets. Hence, not only was the data collected unreliable, this method interfered with the task delivery.

For tasks that required a verbal response (listening recall and some of the phoneme awareness tasks) individual digital voice recorders were trialled with five children. The devices were small and unobtrusive and did not interfere with the delivery of the task. However, this method did not yield usable data. The microphones picked up too much background noise from the classroom, meaning the child's voice could not be distinguished. It was also difficult to hear the facilitator's voice when presenting the trial items so the accuracy of the child's response, could not be judged.

Regarding the pre-and post-intervention outcome measures, administering the full battery of assessments with each child was time-consuming and this may have impacted negatively on the children's motivation and performance. In particular, the NRDLs took a considerable amount of time to complete, whereas the CELF-P (trialled in one school) was much quicker to administer. With regards to the NEPSY-II's auditory attention and statue subtests, all of the RAs found it difficult to observe and simultaneously record the children's performance. Therefore, they doubted the accuracy of their scoring. If this test were used in a full trial, thorough training and practice should be provided to those administering it and inter-rater reliability must be measured.

Regarding the proxy measures of children's functional skills, all of the teacher rating scales (BRIEF-P) of attention in the classroom were completed. This suggests that the checklist was acceptable to teachers. Children's communication skills at home were measured using the FOCUS-34. This tool looks at change/improvement in the child's communication skills over time (rather than providing a direct measure of their ability). It should be completed by the same parent at each time point with support from a SLT [43]. Due to the classroom-based nature of this trial, the forms were sent home for parents to complete and return to the school. Therefore, the parents completed this measure without support. Completed checklists were returned at both time points for 35 children (58% of the sample) but examination of the raw data indicated that for 8 children the forms were not completed by the same parent at the two time points. This raises questions about the reliability of the data. Furthermore, two outlying scores were apparent indicating possible misunderstanding of scoring (a Likert scale) by the parents. In a future trial, greater support would need to be provided to parents (as outlined in the protocol for the FOCUS-34) to avoid these potential issues.

Psychometric properties

One of the key aims of this feasibility study was to determine whether the outcome measures are sensitive enough to detect change as a result of intervention. To examine the sensitivity of the outcome measures, we looked at the distribution of the children's scores at baseline for each assessment. If a lot of children scored at floor levels (low scores), a measure may be sensitive to change, though only if the floor effect is not so low that it masks future improvement. Conversely, if a lot of children score at ceiling levels (high scores) at baseline, the measure may not detect change at the post-intervention time point.

To identify potential floor or ceiling effects, the direction of the scores were explored in terms of skewness: negative skewness values indicate a clustering of scores at the high end that could indicate ceiling effects; and positive skewness values indicate a clustering of scores at the low end that may indicate floor effects [48]. First, the skewness of the children's scores for the full sample ($n=60$) at baseline were explored. This sample included children considered to be typically developing ($n=26$), as well as those with identified difficulties ($n=20$) and those about whom teachers had concerns around listening and communication skills ($n=22$). This means that any clustering of scores to the high end

(negative skewness) could be attributed to high performance on the part of the typically developing children, which could mask the true sensitivity of a measure to detect improvements in children with poorer baseline skills. Consequently, further assessments of normality were conducted by splitting the sample into two groups: typically developing children (TD group) ($n=26$); and children about whom there are concerns/ recognised difficulties with listening and communication skills (concerns group) ($n=34$).

Table 4 presents the descriptive statistics for the baseline data (mean, standard deviation and skewness) for all of the outcome measures for the total and the split sample, along with a brief interpretation of these results. Degrees of skewness were interpreted as follows: less than -1 or greater than $+1$ = highly skewed; between -1 and -0.5 or between $+0.5$ and $+1$ = moderately skewed; and between 0.5 and $+0.5$ = approximately symmetric distribution [49]. Moderate skewness values were considered to be acceptable, but high skewness values (denoted by shaded cells in Table 4) were taken as an indication that a test may not be appropriate for a full trial. The table clearly shows the difference in the distribution of scores between the TD group and the concerns group. For the TD group, potential ceiling effects were found for the phoneme isolation subtest of the PIPA (skewness -1.99), and the BRIEF-P global executive and WM scales (skewness 1.67 and 1.32 respectively[11]). The overall direction of the scores for the children with concerns was the same for these measures (skewed towards better performance) but only to moderate levels indicating that these tools should be appropriate for use in a full trial of RECALL.

The pattern of results presented in Table 4 highlighted some issues that required further investigation. For the listening recall subtest of the AWMA, scores were clustered at the lower end for both groups (TD skewness 1.08 , concerns group skewness 2.21) indicating potential floor effects that may mask children's improvement in a large scale trial. In addition, there was a need to clarify the optimal measures for phoneme awareness (the phoneme isolation or segmentation subtest) and language skills (the NRDLS or the CELF-P). Given the issues highlighted previously about the parents' scoring of the FOCUS-34 the results of this measure also needed further examination.

Baseline and post-intervention scores (mean and standard deviations) for the three intervention groups (RECALL, RISE and no-intervention control) for the full sample ($n=60$) were compared (Table 5) to provide an indication as to whether the outcome measures detected change. As the difference score between pre and post intervention assessment on the FOCUS-34 is the measure of interest for that assessment, only the change score was analysed. The data were not tested for statistical significance of treatment effects because this study aimed to assess the feasibility of a future full trial and consequently, the sample obtained was not statistically powered to support this type of analysis [27]. The findings are summarised below.

Listening recall: Table 5 shows that this test detected differences between the means for the three intervention groups at baseline and at the post-intervention time points. Therefore, despite children's scores being highly skewed towards the lower end, this test should be appropriate for a large-scale trial of RECALL.

Phoneme awareness: The phoneme isolation subtest was originally favoured for this study because it relates directly to the tasks trained in RECALL (identifying the first sound in a word). Due to concerns about potential ceiling effects at baseline, the phoneme segmentation subtest was trialled as an alternative in one school. Post-intervention results suggest that the phoneme isolation subtest is sensitive to change in the population of interest because differences across the intervention groups were apparent. Table 5 shows that the RECALL and RISE (active control) groups improved on the phoneme isolation task, but the no-intervention control group did not. Since the RAs reported that these tasks were quick and easy to administer, it may be acceptable to include both the phoneme isolation and segmentation subtests of the PIPA in a full trial.

Language: The direction of the distribution of scores for the NRDLS and the CELF-P (Table 4) indicated that the NRDLS scores were moderately skewed towards high scores ($-.57$) and the cumulative raw scores for the three subtests of the CELF-P were moderately skewed towards low scores ($.56$). Taken together with the RAs' report that the NRDLS was time-consuming to complete, this suggests that the CELF-P may be a more appropriate language measure for a definitive trial of RECALL.

-

FOCUS-34: For the purpose of the analysis of the parent rating scale two outliers (where it was apparent that the parents had misinterpreted the form) were removed. The FOCUS-34 measures change over time, with a difference of more than 11 points indicating significant clinical change [43]. Table 5 shows a clear difference between the mean change score (the post-intervention score minus the baseline measure) for the no-intervention control group ($\bar{x}=2.12$, $SD=10.23$) in comparison to the RECALL ($\bar{x}=13.46$, $SD=21.70$) and RISE groups ($\bar{x}=12.58$, $SD=18.38$). This suggests that the measure would be sensitive to change over time.

Outcome Measure			Full sample (n= 60)		Split sample				Interpretation of results
Outcome	Task	Test	Mean (SD)	Skewness[2]	Typically developing group (n= 26)	Concerns group (n= 34)	Mean (SD)	Skewness	
Trained task	Listening recall	AWMA[3]	1.16 (1.68)	1.57	1.58 (1.98)	1.08	.81 (1.33)	2.21	Both groups: scores highly skewed towards the low end-potential floor effects
	Odd one out	AWMA	7.16 (3.54)	.28	7.88 (3.98)	-.13	6.56 (3.07)	.61	Both groups - distribution approximates normality
	Phoneme awareness	PIPA Phoneme isolation*	8.90 (4.04)	-1.08	10.08 (3.75)	-1.99	7.81 (4.05)	-.60	Full sample and TD group: highly skewed towards high scores - potential ceiling effects. Children with concerns-moderately skewed
		PIPA Phoneme segmentation†	.30 (.675)	2.28	.00	-	.38 (.74)	1.95	Both groups: highly skewed towards the low end - potential floor effects.
Near-transfer (untrained WM)	Digit recall	AWMA	18.24 (4.96)	-.46	18.69 (6.41)	-.80	17.88 (3.42)	.66	Children with concerns: moderate skewness towards high scores for digit recall and counting recall.
	Block recall	AWMA	10.74 (3.24)	-.28	11.65 (3.90)	-.73	10.00 (2.41)	-.39	
	Counting recall	AWMA	6.21 (3.19)	-.42	7.00 (3.60)	-.53	5.56 (2.71)	-.90	
	Nonword recall	AWMA	4.52 (3.56)	.15	3.96 (3.14)	.03	4.97 (3.86)	.08	
Far-transfer	Auditory Attention	NEPSY-II	19.54 (6.22)	-.69	20.62 (6.47)	-.94	18.70 (5.98)	-.62	Both groups: moderate skewness towards high scores
	Statue	NEPSY-II	22.64 (5.58)	-.81	25.69 (2.95)	-.61	20.24 (6.02)	-.23	Full sample: moderate skewness towards high end. Concerns group-approximates normality.
	Language	NRDLS*	61.06 (4.58)	-.74	62.75 (3.25)	.29	59.50 (5.11)	-.57	Both groups: NRDLS scores moderately skewed towards high performance; CELF-P scores moderately
CELF-P† (Cumulative Raw Scores)		55.40 (8.53)	.44	61.5 (10.61)	-	53.80 (8.01)	.56		

Behaviour in the classroom	BRIEF-P[4] Global Executive Composite	99.57 (30.21)	.90	88.73 (32.13)	1.67	107.85 (26.20)	.76	skewed towards lower end
	BRIEF-P WM scale	62.20 (15.56)	.52	25.27 (10.48)	1.32	31.7 (8.34)	.26	For both scales of this measure: scores are highly skewed to lower end (indicating better performance) for the TD group but not for the concerns group.
Communication skills at home	FOCUS 34 baseline	189.39 (39.76)	-1.52	204.05 (36.76)	-2.51	179.13 (39.11)	-1.35	Highly skewed for full sample and both groups but to a greater degree for TD

Table 4. Descriptive statistics for raw scores at baseline for the full and stratified samples

Outcome	Task	Test used	Time point	RECALL	RISE Active Control	No Intervention	
				(n= 20)	(n= 20)	(n= 20)	
				Mean (SD)	Mean (SD)	Mean (SD)	
Trained task	Listening recall (ELWM)	AWMA	Baseline	.47 (.77)	1.22 (1.83)	1.41 (1.66)	
			Post-intervention	4.11(3.12)	5.28 (4.51)	2.35 (3.74)	
	Odd one out (ELWM)	AWMA	Baseline	7.00 (3.13)	5.94 (3.11)	8.06 (4.13)	
			Post-intervention	8.42 (3.16)	10.44 (3.09)	9.24 (4.49)	
	Phoneme awareness	PIPA	Baseline	6.33 (5.32)	9.47 (2.97)	9.05 (4.21)	
			Post-intervention	7.56 (3.64)	9.63 (3.40)	7.16 (3.85)	
		Phoneme isolation subtest*	Baseline	.30 (.21)	-	-	
			Post-intervention	2.10 (.31)	-	-	
Near-transfer (untrained WM)	Digit recall	AWMA	Baseline	16.58 (5.78)	19.78 (4.25)	17.59 (4.32)	
			Post-intervention	19.37 (4.04)	18.61 (4.64)	18.29 (4.95)	
	Block recall	AWMA	Baseline	11.05 (2.80)	10.28 (3.48)	10.76 (3.7)	
			Post-intervention	11.05 (2.55)	10.56 (5.22)	9.41 (3.97)	
	Counting recall	AWMA	Baseline	16.58 (5.78)	19.78 (4.25)	17.59 (4.32)	
			Post-intervention	19.37 (4.04)	18.61 (4.64)	18.29 (4.95)	
	Nonword recall	AWMA	Baseline	3.58 (3.61)	6.39 (2.97)	3.35 (3.37)	
			Post-intervention	7.26 (2.88)	8.61(4.13)	6.65 (3.23)	
	Far-transfer	Auditory Attention	NEPSY-II	Baseline	18.00 (6.29)	20.05 (6.69)	21.59 (5.83)
				Post-intervention	17.47 (6.78)	21.68 (5.45)	19.82 (5.33)
		Statue	NEPSY-II	Baseline	21.32 (5.82)	22.47 (6.01)	23.72 (5.13)
				Post-intervention	26.37 (4.14)	26.47 (6.60)	23.72 (5.10)
Language		NRDLS Comprehension Scale*	Baseline	60.56 (5.72)	61.47 (3.79)	60.35 (4.76)	
			Post-intervention	62.33 (2.74)	62.95 (2.70)	61.35 (5.15)	
		CELF-P [†] (Cumulative Raw Scores)	Baseline	55.4 (8.53)	-	-	
			Post-intervention	57.00 (7.24)	-	-	
Behaviour in the classroom		BRIEF-P[5] Global Executive Composite	Baseline	60.20 (12.61)	57.55 (14.40)	68.85 (17.66)	
			Post-intervention	57.45 (11.68)	50.70 (9.75)	63.45 (15.40)	
		BRIEF-P	Baseline	27.9 (7.37)	25.85	33.00	

		Working memory scale		(9.63)	(11.11)
			Post-intervention	25.55 (6.97)	22.95 (6.02)
Communication skills at home	FOCUS-34 (Change score)		Post-intervention minus baseline	13.46 (21.70)	2.12 (10.23)

Table 5. Baseline and post-intervention mean and standard deviations for raw scores at baseline and post-intervention (per group) for full sample (n= 60): rows are shaded to ease reading

Acceptability of RECALL

Figure 2 shows that three major themes were identified in the qualitative data gathered through the semi-structured interviews with the HPs and teachers who delivered RECALL.

Some RECALL components are acceptable

All of the HPs and teachers liked the fantastical play component of RECALL, reporting that the puppet, fantastical themes and props were appropriate and fun for year 1 children. The phoneme awareness tasks were easy to administer due to their similarity to usual classroom practice. The listening recall task was also quick and easy to administer. It was at an appropriate level of difficulty (with both the teachers and the HPs reporting that the children seemed to improve across the 6-week intervention period) and engaged the class. The fact that the sentences tied in with the fantastical themes and were funny seemed to appeal to even the most inattentive children. One of the teachers reported:

"I think, the listening recall one benefitted and involved every child....It was actually boys I noticed who probably stick out with the listening recall and the boys who like imaginative play and who like a giggle. So, I actually found that really related to boys. It related to everybody, but they stood out. It surprised me that they were interested. It was just because they thought it was funny, so it just hooked them in and they wanted to be part of it."

Odd one out is challenging

None of the HPs or teachers liked the odd one out task in its current format. The teachers were uncertain about the nature of this specific task and how to deliver it e.g., whether it was acceptable for children to place their fingers on the location of the odd one out picture in their booklets. As outlined earlier, the booklets were hard to manage and the stampers were distracting, meaning that the facilitators had to pause frequently to help the children, thereby disrupting the flow of the activity and elongating it. The participants all reported that there were too many trial items per session so the children became unmotivated, especially those with existing attention difficulties who tended to copy their peers' responses. The HPs and teachers all stated that the difficulty level increased too quickly and the children would have benefitted from additional practice at the 2-to-be-remembered item level. One of the teachers stated:

"I found it was a very big challenge for a lot of them [the children]. At the start it wasn't too bad, but then as it progressed and maybe you were at three odd one out on the one page, then four- it was really, really difficult. Again, those few [children] in the top group would have been trying to focus really well but so many just lost it and a lot of them were randomly stamping. The wee weaker groups, they just weren't focused at all."

Figure 2. Qualitative data themes identified in semi-structured interviews with the HPs and teachers who delivered RECALL

Groups are too big

Whilst the use of booklets and stampers to record children's responses impacted on the acceptability of RECALL, the size of the groups was also identified as a barrier to the intervention delivery. The number of children in the class (divided into groups of 9 or 10) made it difficult to deliver the tasks and to monitor children's progress. This was noted by all of the HPs and teachers during the semi-structured interviews (even for the listening recall task which was universally liked by the participants).

This was summed up by one of the HPs:

"....if there was less children it would be so much easier to guide and judge how they were doing. Because you were only getting a general idea [of how they were doing]."

Compliance and fidelity to the intervention delivery

There was good compliance with the implementation of RECALL regarding the total number of sessions completed (95%) and the number of trials delivered (11 practice items of listening recall and odd one out, and 10-15 minutes of phoneme awareness training). In terms of the quality of delivery, fidelity to the intervention protocol varied between the 2 schools (there was a high degree of inter-rater consistency on the fidelity measure across the research team): school 1 (76%) and school 2 (45%).

This discrepancy related to the delivery of the odd one out task during the teacher-delivered RECALL sessions. In school 1, the teacher divided the class into three groups, as specified in the intervention protocol. In school 2, the teacher presented the task to all of the children at the same time, holding up the picture stimuli and walking around the classroom until each child had seen them. Then the children all stamped the location of the odd one out picture in their booklets. This lengthened the time that the children had to hold the information in their WM, both changing the nature of the task and making it too difficult. The overall duration of the session also increased and the children, especially those who were inattentive, became unmotivated and restless.

[1] On the BRIEF-P higher scores indicate higher executive dysfunction meaning that the child presents with greater difficulty coping in the classroom so positive skewness values indicate potential ceiling effects.

[2] Skewness: 0=perfect normality; negative skewness values indicate a clustering of scores at the high end; positive skewness values indicate clustering at the low end (except on the BRIEF-P (Gioia et al., 2003) where lower scores indicate greater degrees of executive dysfunction so positive skewness = clustering of scores at the high end. Shaded cells =highly skewed values (>1 or <-1)

[3] Raw scores on AWMA subtests represent the number of trials correct (rather than memory span)

[5] Note: higher scores on the BRIEF-P [42] indicate greater degrees of executive dysfunction. A reduction in scores over time indicates improvement. For tests marked* sample ($n= 50$); for tests marked† sample ($n= 10$). For the FOCUS-34 [43] change scores of >11 points indicate significant clinical change.

Discussion

This study responds to calls for globally significant, rigorous and ecologically-valid research into collaborative, classroom-based approaches for children at risk of language disorders and the factors that may impact on their delivery [3-7]. To our knowledge, RECALL is the first

theoretically-underpinned, evidence-based, classroom-based intervention that specifically targets WM to enhance attention and language skills in 4-5-year-old children from areas of SD.

The first research aim was to determine whether it is possible to conduct a definitive CRT to evaluate whether RECALL is more effective than an existing intervention (ALP) and education as usual. The successful recruitment of HPs, schools and children from areas of SD, high completion rates and minimal loss to follow-up suggest that the trial processes could be scaled-up into a definitive trial. However, because staffing levels within the RISE teams may fluctuate, consultation with the service managers will be essential for the successful roll out of a large-scale study.

The stratified sampling method employed in the current study should be modified because this was affected by ambiguity around whether some children had a diagnosis or not. The three strata could be collapsed into two: typically developing children and those about whom teachers have concerns and may/may not have a diagnosis. Including children considered to be typically developing would be valuable in a large trial since high proportions of children in areas of SD are at risk of language disorders and little is known about the individual differences that moderate the effects of WM training [50] and language interventions [51]. A full trial will require a large sample with sufficient power to detect differences between subgroups of children as well as intervention groups [52].

Regarding the appropriateness of the pre-and post-intervention outcome measures, the descriptive statistics suggest that the following measures could be used in a full trial: the AWMA for working memory; the phoneme isolation subtest of the PIPA for phoneme awareness; the BRIEF-P and FOCUS-34 (completed according to the protocol) as proxy ratings of attention in the classroom and communication skills at home; and the NEPSY-II for attention (provided thorough training is provided for those administering it and the inter-rater reliability is assessed). The phoneme segmentation subtest of the PIPA could be used in addition to the phoneme isolation subtest. This would add minimal time to the assessment process and its inclusion would mitigate against any risk of reduced sensitivity of the phoneme isolation subtest. To assess language, the NRDLS should be replaced by the CELF-P since this takes less time to administer and should be more acceptable for both the child and RA administering it. The use of digital voice recorders to monitor verbal responses is not feasible in the classroom setting and the use of booklets impeded the completion of the odd one out task. Therefore, alternative methods of monitoring children's performance on a weekly basis, perhaps by a trained observer, will be required in a definitive trial.

Exploration of the acceptability of RECALL produced mixed findings. The HPs and teachers liked the listening recall and fantastical play components. These were considered to be fun and at an appropriate level for 4–5 year olds. The phoneme awareness component was also acceptable. The fact that the teachers reported that these tasks are similar to their usual practice may suggest they are not required in RECALL. However, since the descriptive analysis of the post-intervention scores showed a trend towards improvement for the intervention groups, further investigation of the effectiveness of these tasks could be valuable in a full-scale trial.

The odd one out task is not acceptable to HPs or teachers in its current form for two key reasons. First, they found it difficult to manage the materials, which included picture stimuli, booklets and stampers. Second, many of the children became inattentive during this task. These factors relate to the task itself, but also to the classroom setting and the children's characteristics. The task delivery could be simplified by enabling the children to indicate the odd one out location by pointing. The dose delivered per session (11 practice items) and dose frequency (three times per week) may also have been too intense for 4-5 year olds. In addition, from week three onwards the task became too difficult i.e., there were too many items to-be-remembered. These findings are consistent with emerging evidence regarding the effects of dosage on the outcomes of language interventions, suggesting that if treatment is too intense it can be detrimental to children's learning [5,25,51]. The current study underscores the need for robust investigation of dosage in both WM and language interventions. Modifications to RECALL, including the task delivery and its dosage could be explored through further co-production work and small group work with 4-5 year olds prior to a full-scale trial.

In relation to the classroom setting and the children's characteristics, the potential effectiveness of RECALL was impeded by the size of the groups set up for the task (9-10 children) and their composition, where the weaker children were observed copying their more-able peers. Many of the children, particularly the most at risk children (already presenting with inattentive behaviour in the classroom) became unmotivated by a task that was too challenging for them.

The third research aim was to measure the compliance and fidelity to the intervention protocol. The inconsistency between the high level of compliance observed and the variation in fidelity between the teacher-delivered sessions in the two schools highlights the importance of facilitator factors in intervention delivery. The teachers' uncertainty about how to deliver the tasks demonstrates that the detailed intervention manual and demonstration provided by the trained HPs were not sufficient. This supports the provision of direct training on the theoretical underpinning and delivery of RECALL for all teachers involved in a future definitive trial of the intervention. This should include a minimum of eight hours' direct instruction, as well as coaching and feedback [3]. These measures would better ensure fidelity of task delivery which is essential for the therapeutic effectiveness of the intervention, particularly the odd-one-out task (and was lost for half of the participants in the RECALL arm of this trial).

The evidence discussed so far has illuminated a dynamic interplay between the way a therapeutic task is presented and its difficulty level (dose form); the setting within which it is delivered such as group size; children's characteristics including their motivation and attention; and facilitator characteristics e.g., theoretical knowledge of the task and how to deliver it. This complex blend impacts on, and can dilute the number of trials accessed by individual children, particularly those who are most at risk of language disorder. Figure 3 graphically represents this.

Figure 3. The factors impacting on dosage in classroom-based interventions

Limitations of the present study

This was a small scale study with just two schools in each arm of the trial. However, the findings are strengthened by the study design through the inclusion of an active control group receiving an intervention of comparable structure and dosage to the experimental RECALL intervention. The need to widen the school eligibility criteria to include those in the lowest quintile of SD means that half the sample were in less disadvantaged areas than originally anticipated. This raises questions about whether the findings are generalisable to schools in more disadvantaged areas of NI or beyond. The lack of blinding of the HPs and the fact that the teachers in the RECALL schools became unblinded to their group allocation may have introduced bias regarding their perspectives on the intervention. In addition, compliance and fidelity to the delivery of the active control intervention was not measured and this should be addressed in a definitive trial.

Conclusions

RECALL is a novel, multi-component intervention that targets WM to enhance attention and language skills. To our knowledge, this is the first paper to report on the feasibility of implementing a WM intervention in real-life contexts. With the exception of the methods used to monitor children's progress from week to week, the trial processes could be scaled-up into a future definitive trial to evaluate the effectiveness of RECALL.

In relation to the intervention components, this study has provided unique evidence of the potential effectiveness of the two directly trained WM tasks (listening recall and odd one out) for children as young as 4-5 years. Listening recall was implemented successfully and was acceptable to the HPs and teachers who delivered the intervention; and odd one out could be modified to enhance its acceptability and the fidelity of its delivery.

Overall, the potential effectiveness of RECALL for the children who may benefit most from it (i.e., those presenting as inattentive in the classroom and are at risk of low WM) could be optimised if it were implemented in small group settings. This would enhance its acceptability to HPs and teachers and improve its potential effectiveness by maximising the dosage accessed by individual children. RECALL could be modified through further co-production work and feasibility testing involving small group work with 4-5 year olds. This study has highlighted the challenges of balancing empirically-evidenced dosage with the feasibility and acceptability of what can be delivered in real-life contexts. Furthermore, it emphasises the need for teachers to have thorough training on the theoretical underpinning to interventions for children with language disorders in the mainstream classroom.

Abbreviations

ALP: Attention and listening programme; AWMA: Automated Working Memory Assessment; BRIEF-P: Behaviour Rating of Executive Functions-Preschool; CA: Classroom Assistant; CELF-P: Clinical Evaluation of Language Fundamentals-Preschool; CRT: Cluster Randomised Trial; ELWM: Executive-loaded working memory; FOCUS-34: Focus on Communication Skills Under Six-34; HSCT: Health and Social Care Trust; NEPSY-II: A Developmental Neuropsychological Assessment; NI: Northern Ireland; NRDLS: New Reynell Developmental Language; OT: Occupational Therapist; PIPA: Preschool Inventory of Phonological Awareness; PT: Physiotherapist; RA: Research Assistant; RA: Research Assistant; RISE: Regional Integrated Support for Education; RECALL: Recall to Enhance Children's Attention, Language and Learning programme; SD: social disadvantage; SEB: Social Emotional and Behaviour Specialist; SLT: Speech and Language Therapist; STM: short-term memory; WM: working memory.

Declarations

Ethics approval: Ethical approval for this study has been granted by the Ulster University Research Ethics Committee (REC/18/0036). The trial has been registered with the International Standard Randomised Controlled Trial Registry (ISRCTN13633886)

Consent for publication: Not applicable.

Availability of data and material: As this is a feasibility trial, sharing of the dataset is not anticipated. However, any requests for data or material should be made to the corresponding author. Requests will be reviewed by the Trial Steering Committee.

Competing interests: The authors declare that they have no competing interests.

Funding: This trial constitutes part of a doctoral research study funded by the Research and Development Division of the Public Health Agency, Northern Ireland. The funder was not involved in the design or conduct of any aspect of the trial.

Authors' contributions: AR, JT and LT conceived of the study and applied for the funding of the doctoral project within which this study took place. LH and JH contributed thinking and support to the study design for outcome measurement and intervention in particular. The first author wrote the first draft of the manuscript to which the co-authors then contributed. All of the authors have read and approved the final manuscript.

Acknowledgements: The authors thank Professor Mike Clarke (Director of MRC Methodology Hub, Queen's University Belfast) for his input to the methodological design of this feasibility trial and Professor Brendan Bunting (Ulster University) for advice in relation to the data analysis.

References

1. Law J, McBean K, Rush R. Communication skills in a population of primary school-aged children raised in an area of pronounced social disadvantage *International Journal of Language and Communication Disorders*. 2011;46(6):657-664.
2. Locke A, Ginsborg J, Peers I. Development and disadvantage: implications for the early years and beyond. *International Journal of Language and Communication Disorders*. 2002;37(1):3-15.
3. Ebbels SH, McCartney E, Slonims V, Dockrell J, Norbury CF. Evidence based pathways to intervention for children with Language Disorders. *International Journal of Language and Communication Disorders*. 2019;54(1):3-19.
4. ICAN/RCSLT. (2018). *Bercow: Ten Years on an Independent Review of Provision for Children and Young People With Speech, Language and Communication Needs in England*. London: ICAN/RCSLT.
5. Schmitt MB, Justice LM, Logan AR. Intensity of language treatment: contribution to children's language outcomes. *International Journal of Language and Communication Disorders*. 2017;52(2):155-167.
6. Law J. Population woods and clinical trees. A commentary on 'Evidence-based pathways to intervention for children with language disorders'. *International Journal of Language and Communication Disorders*. 2019;54(1):26-27.
7. Murphy C. The limits of evidence and the implications of context: considerations when implementing pathways to intervention for children with language disorders. *International Journal of Language and Communication Disorders*. 2019;54(1):20-23.

8. Department of Health, Social Services and Public Safety (DHSSPS). Making Life Better. A whole system strategic framework for public health. 2014. https://www.health-ni.gov.uk/sites/default/files/publications/dhssps/making-life-better-strategic-framework-2013-2023_0.pdf Accessed 30 Sep 2016.
9. Health and Social Care Board (HSCB) (2011) *Transforming Your Care. A review of Health and Social Care in Northern Ireland*. Belfast: HSCB;2011.
10. Bengoa R. Systems, not structures. Changing health and social care – full report. <https://www.healthni.gov.uk/sites/default/files/publications/health/expert-panel-full-report.pdf> Accessed 30 Nov 2016.
11. Gascoigne Supporting children with speech, language and communication needs within integrated children's services. RCSLT Position Paper. London: RCSLT;2006.
12. Harron A, Dickson F. Perspectives on trans-disciplinary working for children in mainstream education. Oral presentation: Rehabilitation and Therapy Research Society conference;2013.
13. Cowan N, Fristoe NM, Elliott EM, Brunner RP, Sauls JS. Scope of attention, control of attention and intelligence in children and adults. *Memory and Cognition*. 2006;34(8):1754-1768.
14. Baddeley AD, Gathercole SE, Papagno C. The phonological loop as a language learning device. *Psychological Review*. 1998;105:158-173.
15. Holmes J, Butterfield S, Cormack F, Loenhoud AV, Ruggero L, Kashikar L, et al. Improving working memory in children with low language abilities. *Frontiers in psychology*. 2015;6:519.
16. Melby-Lervåg M, Hulme C. Is working memory training effective? A meta-analytic review. *Developmental Psychology*. 2013;49(2):270-291.
17. Gathercole SE, Holmes J, Dunning DL, Norris D. Working memory training involves learning new skills. 2019: *Journal of Memory and Language*. 2019;105:19–42.
18. Holmes J, Woolgar F, Hampshire A, Gathercole S Are working memory training effects paradigm-specific? *Frontiers in Psychology*. 2019;10:1103.
19. Dunning D, Holmes J. Does working memory training promote the use of strategies on untrained working memory tasks? *Memory and Cognition*. 2014;2; 854-862.
20. Funnell SC, Rogers PJ. Purposeful program theories effective use of theories of change and logic models. San Francisco: John Wiley & Sons;2011.
21. Hawkins J, Madden K, Fletcher A, Midgley L, Grant A, Cox G. et al. Development of a framework for the co-production and prototyping of public health interventions. *BMC Public Health*. 2017;17:1-11.
22. Wight D, Wimbush E, Jepson R, Doi L. Six steps in quality intervention development (6SQuID). *Journal of Epidemiology and Community Health*. 2016;70:520-525.
23. Craig P, Dieppe P, McIntyre S, Mitchie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 2008;337:a1655.
24. Henry LA, Messer DJ, Nash G. 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*. 2014;23(1):84-103.
25. McGinty AS, Breit-Smith A, Fan X, Justice LM, Kaderavek JN. Early Childhood Research Quarterly Does intensity matter? Preschoolers' print knowledge development within a classroom-based intervention *Early Childhood Research Quarterly*. 2011;26:255-267.
26. Justice LM. Conceptualising "dose" in paediatric language interventions: Current findings and future directions. *International Journal of Speech-Language Pathology*, 2018;20(3):318-323.
27. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ*. 2016;355:i5239.
28. Authors (blinded for review)
29. Northern Ireland Multiple Deprivation Measure (NIMDM) (2017) <https://www.nisra.gov.uk/statistics/deprivation/northern-ireland-multiple-deprivation-measure-2017-nimdm2017>. Accessed 28 Mar 2018.
30. Hoffmann TC, Glasziou PP, Broutron I, Perera R, Moher D, Altman DG. et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *British Medical Journal*. 2014;348:g1687.
31. Authors (blinded for review)
32. Thibodeau RB, Gilpin AT, Brown MM, Meyer BA. The effects of fantastical pretend-play on the development of executive functions: An intervention study. *Journal of Experimental Child Psychology*. 2016;145:120-138.
33. Melby-Lervåg M, Hulme C. Serial and free recall in children can be improved by training: evidence for the importance of phonological and semantic representations in immediate memory. *Psychological Science*. 2010;21(11)1694-1700.

34. Van Kleeck A, Gillam RB, Hoffman LM. Training in phonological awareness generalizes to phonological working memory: a preliminary investigation. *The Journal of Speech and Language Pathology – Applied Behavior Analysis*. 2006; 1(3):228-243.
35. Braun V, Clarke V, Using thematic analysis in psychology. *Qualitative Research in Psychology*. 2006;3(2):77-101.
36. Shipstead Z, Hicks KL, Engle RW. Working memory training remains a work in progress. *Journal of Applied Research in Memory and Cognition*. 2012;1:217-219.
37. Dodd B, Crosbie S, Macintosh B, Teitzel T, Ozanne A. *Primary and Preschool Battery of Phonological Awareness (PIPA)*. London:Psychological Corporation; 2000.
38. Letts C, Edwards S, Sinka I, Schaefer B, Gibbons W. Socio-economic status and language acquisition: children 's performance on the new Reynell Developmental Language Scales. *International Journal of Language and Communication Disorders*. 2013;48:2:131-143.
39. Wiig EH, Secord WA, Semel E. *Clinical evaluation of language fundamentals—Preschool, 2nd (CELF Preschool-2)*. Toronto: The Psychological Corporation/A Harcourt Assessment Company; 2004.
40. Alloway TP. *Automated working memory assessment*. London: Pearson Assessment;2008.
41. Korkman M, Kirk U, Kemp S. *NEPSY-II: a developmental neuropsychological assessment*. San Antonio: The Psychological Corporation;2007.
42. Gioia GA, Espy KA, Isquith PK. *The Behavior Rating Inventory of Executive Function-Preschool version (BRIEF-P)*. Odessa: Psychological Assessment Resources;2003.
43. Thomas-Stonell N, Oddson B, Robertson B, Walker J, Rosenbaum P. *The FOCUS©-34: Focus on the Outcomes of Communication Under Six*. Toronto: Holland Bloorview Kids Rehabilitation Hospital; 2015.
44. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2012;350:h1258
45. Grant A, Treweek S, Dreischulte T, Foy R, Guthrie B. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials*. 2012;14:15.
46. Steckler A, Linnan L. eds. *Process evaluation for public health interventions and research*. San Franscisco: Jossey-Bass;2002.
47. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implementation Science*. 2007;2(40).
48. Pallant J. *SPSS Survival Manual*. Buckingham Philadelphia: Open University Press;2016.
49. Justice LM, Logan JA, Schmitt MB, Jiang H. Designing effective speech-language interventions for children in the public schools: Leverage the spacing effect. *Policy Insights from the Behavioral and Brain Sciences*. 2016;3(1):85-91.
50. Eldridge SM, Costelloe CE, Kahan BC, Lancaster GA, Kerry SM. How big should the pilot study for my cluster randomised trial be? *Statistical Methods in Medical Research*. 2015;25(3):1039-1056.

Figures

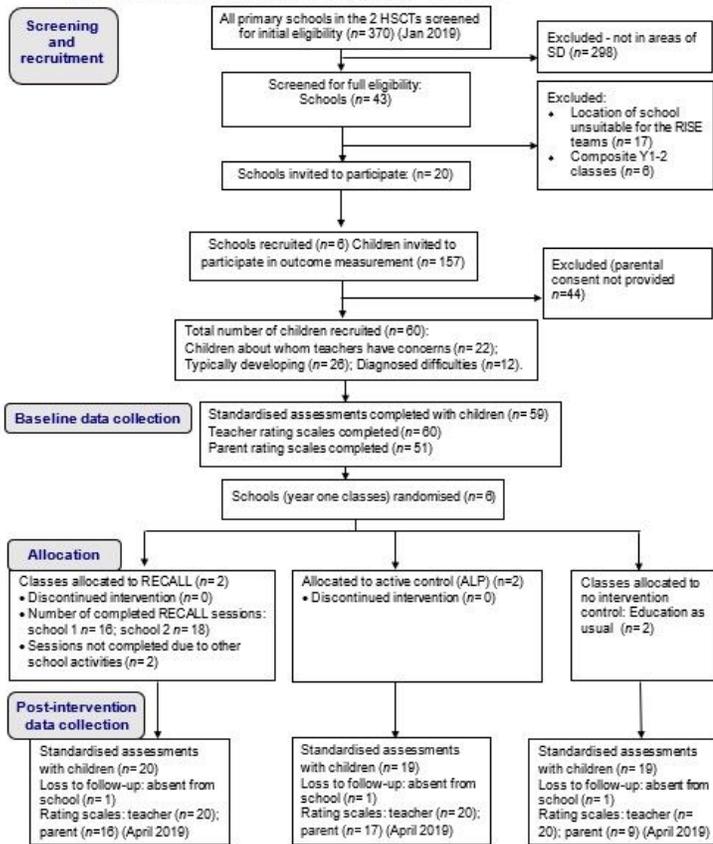


Figure 1. RECALL cluster randomised feasibility trial flow char (following CONSORT guidance, 2010) [27]

Figure 1

RECALL cluster randomised feasibility trial flow char (following CONSORT guidance, 2010) [27]

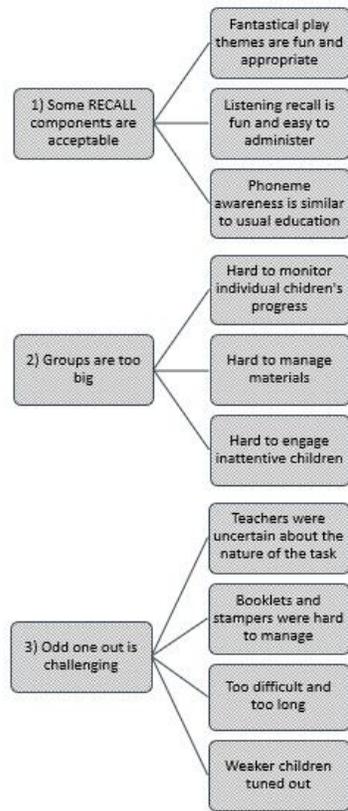


Figure 2

Qualitative data themes identified in semi-structured interviews with the HPs and teachers who delivered RECALL

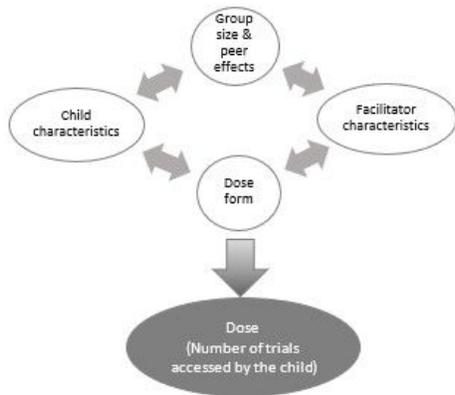


Figure 3

The factors impacting on dosage in classroom-based interventions

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [RECALLFeasibilitystudyCONSORTChecklist28thJan2020.doc](#)
- [A1.docx](#)