Protocol: Interventions to Improve Mathematical Performance for Children with Mathematical Learning Difficulties (MLD)
Mairead Furlong, Fergal McLoughlin, Sinead McGilloway & David Geary

Submitted to the Coordinating Group of:

☐ Crime and Justice
☒ Education
☐ Disability
☐ International Development
☐ Nutrition
☐ Social Welfare
☐ Other:

Plans to co-register:

☐ No
☒ Yes ☒ Cochrane ☐ Other
☐ Maybe

Date Submitted:
Date Revision Submitted:
Approval Date:
Publication Date: 02 May 2016

Campbell Collaboration Systematic Review Protocol Template version date:
9 February 2014
BACKGROUND

Description of the Condition

Mathematical competency is of prime importance in everyday life, and is necessary for simple but essential tasks such as counting, reading the clock, and for budgeting time and money resources. Moreover, numeracy is related to longer-term educational, occupational, physical and mental health outcomes, as well as to the economic status of countries (Fuchs, 2009; Gross, 2009). However, approximately 20% of students have low numeracy skills, and depending on classification methods, 4% to 14% have been identified with a learning disability in mathematics (Butterworth 2010a; Shalev, 2005). Depending on the research tradition (e.g. math cognition or intervention studies), this condition is variously called developmental dyscalculia, mathematical learning disabilities, mathematical disability or mathematical learning difficulties (Szucs, 2013). For the purpose of this review, we will use the term mathematical learning difficulties (MLD) to incorporate both children who score at or below the 10th percentile in mathematical achievement (i.e., children with an identified mathematics disability) and those who score between the 11th and the 25th percentile (i.e., low achieving children at risk of a mathematics disability).

MLD are as common in girls as in boys, and are associated with psychological, neurological and genetic conditions, such as epilepsy, Turner’s Syndrome, fragile X syndrome, phenylketonuria, attention-deficit hyperactivity disorder [ADHD], dyslexia, and behavioural and emotional disorders (Shalev, 2004). MLD have an early onset (Schopman, 1996), and although some children may demonstrate improvement in later grades (i.e., are merely developmentally delayed), there is some consensus that early MLD will typically persist into late adolescence and adulthood (Bryant, 2005; Butterworth, 2011; Gerber, 2012; Shalev, 2005).

MLD involve a deficit in mastering one or many of the domains of mathematics, from the more basic number skills to the more advanced areas of algebra and geometry. To date, the study of MLD has focused predominantly on the failure to achieve competency in the more basic mathematical skills rather than in more advanced areas (Fischer, 2013). The basic skills include precursor number specific skills (or ‘number sense’) and arithmetical reasoning. Deficits in precursor number sense typically involve a lack of fluency in estimating and judging quantity and in using the mental number line; a failure to understand counting (e.g., cardinality, ordinality) and/or use of immature counting strategies; difficulties in transcoding between number words, digits and analogue quantities; understanding the base-10 number system; and borrowing from one column to the next (Gersten, 2005; Hanley, 2005). Arithmetical reasoning involves the capacity to undertake the basic computational skills of addition, subtraction, multiplication and division; and to reason from word or story problems (Butterworth, 2011; Hein, 2000).
Traditionally, MLD were defined as a specific and isolated impairment in numeracy despite normal intelligence and scholastic opportunity (APA, 2000; WHO, 1992; 2007). However, the specificity clause has recently been removed in order to better represent the heterogeneity of the condition and to improve the clinical utility of DSM diagnoses (APA 2013; Kaufmann, 2012). Therefore, MLD are currently understood to involve three subtypes: (1) those with specific and isolated mathematical difficulties (e.g., specific deficiencies in numeracy despite normal competency in non-mathematical areas and adequate schooling); (2) those with deficiencies in both numerical and non-numerical cognitive functions (e.g., attention, working memory and visuospatial skills), without reference to the child’s general intelligence level; and (3) MLD with comorbid disorders (e.g., ADHD, dyslexia, and behavioural and emotional problems) (Kaufmann, 2012).

In addition to the subtypes of MLD indicated above, there is further evidence that child MLD profiles are heterogeneous, which has implications for identifying children with MLD and presents challenges in developing appropriate interventions and in evaluating their effectiveness (Dowker, 2005). For instance, children with MLD may have specific deficits in one or more mathematical sub-domains (e.g., assessment of quantity) but may perform at grade-level or better in other areas of numeracy (e.g., fractions). Therefore, each child may have their own profile of strengths and weaknesses. However, such profile variation may be masked using standardised mathematical achievement tests, in which performance is averaged across many different types of items (Dowker, 2005). Similarly, interventions may fail to target relevant deficits if it is believed that all children with MLD share similar profiles. Indeed, many targeted mathematics programmes tend to focus on a limited component of number processing (e.g., number line problems), apparently based on the assumption that all children with MLD have difficulties with the number line. This is not necessarily the case. For child A, who has difficulties with the number line, a number line intervention may be beneficial, whereas the same intervention will not be effective for child B who has no difficulties with the number line but who has severe problems with another aspect of numeracy. Consequently, a failure to grasp the componential, non-unitary nature of MLD presents considerable challenges in developing and evaluating appropriate interventions (Dowker, 2005).

Furthermore, MLD appear to be associated with a variety of number-specific and more general mechanisms, which again may have implications for intervention. For instance, studies indicate that computational skills (e.g., addition and subtraction) are linked to specific deficits in mathematical abilities (e.g., difficulties in assessing quantity and magnitude, counting, mapping onto the mental number line) whereas difficulties in word problems are linked more to general deficits in executive working memory, language and inattentive behaviour (Fuchs, 2010; Geary, 2012). Research from cognitive science and intervention studies further indicate that motivation, attention, behaviour, mathematical self-efficacy and anxiety may all be implicated in MLD, in addition to specific deficits in numeracy (Cohen Kadosh, 2013; Diperna, 2006; Kaufmann, 2012; Krinzinger, 2006; McLean, 2014; Rosselli, 2006; Rourke, 1997). The high rate of MLD co-occurrence with
ADHD (26-42% of cases) and with dyslexia (between 17-60%) also suggests a role for more general cognitive mechanisms in the etiology and maintenance for MLD, particularly in the areas of working memory, language and attention (Fuchs, 2010; Raghubar, 2010; Ramaa, 2002; Shalev, 1997).

Although the prevalence of MLD is comparable to the incidence of dyslexia, children with MLD are often not diagnosed or treated properly due to a persistent lack of knowledge about the disorder (Dowker, 2004). Furthermore, until recently, the study of MLD was relatively neglected; for instance, between 1996 and 2005, the ratio of studies on reading disabilities to MLD was 5:1. This was a dramatic improvement over the ratio of 16:1 in the prior decade (Gersten, 2009). However, in recent years, a growing interest in numeracy has led to the development and evaluation of different interventions to treat MLD, with an emphasis on targeted mathematics interventions (Shalev, 2004). Experts also recommend that behavioural, psychological and pharmacological interventions are utilised in addressing MLD (Kaufmann, 2012). Consequently, this review will consider a range of interventions that may help in the prevention and treatment of MLD.

**Description of the Intervention**

The treatment of students with MLD is complex due to heterogeneity in its presentation, comorbidities, and putative underlying mechanisms, which appear to involve deficits in specific mathematical skills, as well as in more general domains that may mediate mathematical competency, such as attention, working memory, in-class behaviour, mathematical self-efficacy and anxiety (Fuchs, 2010; Geary, 2012; Kaufmann, 2012). Most research activity to date has involved the development and evaluation of targeted mathematics interventions that involve teaching numerical skills to students. However, given the range of non-numerical deficits also involved in MLD, this review will also assess the effectiveness of other relevant interventions in improving child MLD. These interventions include: (1) targeted mathematics interventions; (2) behavioural and psychological interventions; (3) non-invasive brain stimulation technologies; (4) pharmacological interventions; and (5) multi-component interventions that involve two or more of the preceding elements identified in 1 to 4. It is important to note that while many of these interventions might not label themselves as specifically targeting mathematics, they may still include mathematical performance as a programme objective. Therefore, we will include evaluations of virtually any programme that involved students with MLD and measured mathematical performance as an outcome variable, irrespective of whether they are called 'mathematics programmes'. In order to represent the scope of relevant research on this topic, all such programmes should be considered in a review of interventions to improve mathematical performance in students with MLD.

**Targeted Mathematics Interventions**

Targeted mathematics interventions aim to remediate skill deficits in one or more numerical areas. They may involve several components, including: direct, explicit instruction; use of
cognitive heuristics/strategies; student verbalisations of their mathematical reasoning; using visual representations while solving problems; provision of a range and sequence of examples; repeated practice; and corrective feedback (Gersten, 2009; Kroesbergen, 2003). Examples of cognitive strategies include: counting on or back from first summand (e.g., $9 + 3$ is solved by counting on 3 from 9); learning complementarity of addition and subtraction (e.g., $a + b = c \rightarrow c - a = b$); estimating set size; understanding ordinal position (e.g., 5 in 52 means 50 but 5 in 25); and translating between words and digits for quantity (Butterworth, 2010b; Cohen Kadosh, 2013; Gersten, 2009). Such interventions may differ from each other with regard to a number of factors, including: the utilisation of individualised programmes to target the specific profile deficits of children or using a broader group/classroom approach; the intensity and duration of the intervention (e.g., ranging from one session of 30 minutes to regular sessions conducted over a 6- or 12-month period); the instructional procedures employed; and and the age and grade of students (Dowker, 2005). Examples of targeted mathematics interventions that have been evaluated using a range of randomised and quasi-experimental methods include: Pirate Math, Maths Flash, Fluency and Automaticity through Systematic Teaching with Technology (FASTT Math), Number Worlds, The Number Race and Catch Up Numeracy (Cohen Kadosh, 2013; Fuchs, 2009; Holmes, 2013; Kroeger, 2012; Wilson, 2006).

**Behavioural and Psychological Interventions**

Behavioural and psychological interventions may not directly target mathematical skills. However, there is a growing body of evidence that academic skill (including mathematical competency) is mediated by student attitudes and behaviours that facilitate their participation in, and capacity to benefit from academic instruction in the classroom (Diperna, 2006). These attitudes and behaviours are called ‘academic enablers’ and involve motivation, engagement, social and behavioural skills that mediate academic competence. Thus, behavioural and psychological interventions that improve academic enablers may potentially lead to improvements in mathematical performance. For instance, tackling self efficacy, task anxiety, and school phobia are key foci of many cognitive-behavioural and counselling interventions, and such academic enablers have been shown to improve effort, persistence and achievement for a range of academic outcomes (Pintrich, 2002; Shechtman, 2005). In addition, the ability of students to actively participate and engage with academic instruction is associated with improved academic outcomes (Greenwood, 2002). Examples of engagement include arriving prepared for lessons, meeting deadlines, appropriate school behaviour, following directions and completing homework assignments. Therefore, interventions that target student classroom inattention and disruptive behaviour, such as school-based or classroom-management programmes, may be helpful in improving student MLD (Shalev, 2004).

For instance, studies by Fuchs (2010) and Geary (2012) indicated that in-class attention directly or indirectly contributes to domain-specific and general cognitive deficits associated with MLD. In addition, a meta-analysis conducted by Swanson (1998) investigated the
effects of an array of behavioural, social and academic interventions on the performance of adolescents with learning disabilities (predominantly involving those with reading disabilities but including MLD in some studies) and found beneficial impacts in relation to social, cognitive and academic domains, including mathematical performance. Therefore, it may be useful to identify studies that target academic enablers with MLD students and to compare their effectiveness with other types of relevant interventions.

**Non-invasive Brain Stimulation**

Emerging evidence suggests that non-invasive brain stimulation (NIBS) may be used to modulate brain activity and positively impact on basic numerical skills (Cohen Kadosh, 2013). NIBS technologies may include, among other methods, neurofeedback (NFB), transcranial magnetic stimulation (TMS), and transcranial electrical stimulation (TES). Transcranial direct current stimulation (TDCS) is a commonly used form of TES. NFB and TDCS have been used in the treatment of MLD and will be discussed here. TMS is not currently used due to the risk of seizures as a side effect (Cohen Kadosh, 2012).

NFB is a type of biofeedback that uses real-time displays of brain activity—most commonly electroencephalography (EEG), to teach self-regulation of brain function. Typically, sensors are placed on the scalp to measure activity, with measurements displayed using video displays or sound. Studies on the effects of NFB on MLD in children are limited but it is believed that enhancement of Beta and beta/theta ratio may improve attention-deficit cognitive disturbances associated with MLD (Gottfried, 2010). A randomised controlled trial (RCT), conducted with 28 third grade students with MLD, indicated that NFB (based on enhancing the beta/theta ratio in the CZ region to the left ear) improved mathematical results on a standardised mathematics test compared to children who received a placebo (Hashemian, 2015). The effect was also apparent at one-year follow up. Unfortunately, the study did not report which mathematical domains were improved.

TDCS involves the delivery of weak electrical currents (e.g., 1–2 mA) via electrodes, most frequently at the size of 25–35 cm², which are placed on the scalp above the brain area of interest. When the current is applied over a short duration (~20 min), it passes painlessly through the scalp and skull and alters spontaneous neural activity (Fritsch, 2010). Recent RCT studies indicate that TDCS improved basic numerical skills, arithmetic reasoning and automaticity in adults with MLD, including a transfer effect to new material, and long-term efficiency in brain functions in the stimulated brain region, i.e. the intraparietal sulcus, posterior parietal cortex, the dorsolateral prefrontal cortex and other cortical and subcortical regions that are hypothesised to contribute to numerical cognition (Cohen Kadosh, 2010; Iuculano, 2013; Snowball, 2013). Other experimental studies have found that TDCS may also improve attention, working memory, language and executive functions (Cohen Kadosh, 2013). Results, to date, have indicated that the most effective improvements are associated with stimulation being paired with a brief math-training intervention and that the timing of stimulation with respect to task performance has important effects (Stagg, 2011). Currently,
there is only a limited amount of work with pediatric populations (Cohen Kadosh, 2013). Although no side effects are apparent in adult populations, there are concerns that the atypically developing brain may potentially respond differently to stimulation than the adult brain. On the other hand, there are also concerns that children with MLD may be denied an inexpensive and promising treatment. The current consensus is that experiments in children should start with small samples and that the developing brain should be monitored for possible neurochemical, anatomical, and functional changes that are associated with TDCS (Cohen Kadosh, 2012).

Pharmacology

Pharmacological interventions may not appear to target MLD directly (Cohen Kadosh, 2013). However, when MLD are comorbid with ADHD, anxiety or depression, drugs may be appropriate. For instance, stimulants, such as methylphenidate or dextroamphetamine, enhance attention, concentration and social behaviour, thereby allowing children to respond more efficiently to instruction (Grizenko, 2006; Rubinsten, 2008). Similarly, anti-anxiety or anti-depressant medication enhance student motivation and self esteem, which may allow them to engage more effectively with mathematical instruction (Shalev, 2004; Kaufmann, 2012).

For methylphenidate or dextroamphetamine, once an optimal dosage is reached, an equivalent dosage of the same drug in a sustained-release form is often substituted to avoid the need for drug administration in school. Long-acting preparations include wax matrix slow-release tablets, biphasic capsules containing the equivalent of two doses, and osmotic release pills and transdermal patches that provide up to 12 hours of coverage. Both short-acting and long-acting liquid preparations are also available. Prodrug preparations are sometimes used because of their smoother release, longer duration of action, fewer adverse effects, and lower abuse potential. Learning is often enhanced by low doses, but improvement in behavior often requires higher doses (Sulkes, 2013).

Multiple Interventions/Multi-component Interventions

MLD are heterogeneous and what works for one person may not work for another. Therefore, experts and physicians in the field recommend that many different types of treatment/multi-component interventions may be necessary to target the range of deficits associated with MLD. These may include a combination of targeted mathematics interventions, behavioural, psychological and/or pharmacological components to address both numerical and non-numerical deficits associated with MLD (Kaufmann, 2012). Examples of combination interventions might include: targeted mathematical instruction delivered in conjunction with CBT (Shechtman, 2005), TDCS delivered in conjunction with a targeted mathematics intervention (Snowball, 2013) or behavioural programmes delivered with pharmacology for those who have MLD comorbid with ADHD (Reid, 2005; Zentall, 2007). Types of multi-component interventions will be coded during the review stage.
How the Intervention Might Work

A variety of putative mechanisms of change may mediate the effectiveness of the range of interventions included in this review. More detail follows below.

Targeted Mathematics Interventions

Targeted mathematics interventions are based on the premise that direct training and instruction in mathematics leads to better outcomes than delivering training to improve cognitive deficits in non-mathematical areas (e.g., working memory or attention), even if such deficits are implicated in MLD (Swanson, 1999). The strategies employed in targeted mathematics interventions (described in Description of the intervention) are believed to develop the quantitative (e.g., 5 objects), verbal (e.g., “five”), and visual-abstract (e.g., the numeral 5) elements of number sense (Dehaene, 2005; Butterworth, 2010b), as well as improving encoding and retrieval of basic arithmetical facts (Geary, 2012). Having a sense of the mental number line is hypothesised to be of fundamental importance for arithmetical reasoning and for mental calculation, and extends the semantic range of the concept of number to a more complex and abstract level (Kaufmann, 2012). Students are also believed to develop mathematical competency through verbalisation of mathematical reasoning, dedicated practice, and receiving detailed corrective feedback. Verbalisation is posited to anchor students’ self-regulation and to encourage them to utilise a step-by-step solution strategy rather than a more random, impulsive approach (Swanson, 2000; Gersten, 2009). Repeated practice appears to foster the automatic recall of learned content, thereby lessening the demand on working memory. This demand is especially high in multi-digit calculations with carrying, where interim results must be retained and manipulated (e.g., adding 87 and 45) (Slavin, 2009). Ongoing corrective feedback and the collation of detailed knowledge of performance may also mediate positive outcomes through positive reinforcement of accurate responses and through goal setting (Gersten, 2009). Computer-based interventions may be particularly useful in this respect as they can be designed to adapt for different performance profiles and provide intensive training in a stimulating environment. In addition, research suggests that the reward-based nature of such instruction may act on the dopaminergic system that is involved in plasticity (Lisman, 2011).

It is important to note, however, that previous reviews and meta-analyses of targeted mathematics interventions have reported inconsistent results, with effect sizes ranging from -0.44 to above 3 (Xin, 1999; Kroesbergen, 2003; Gersten, 2009; Codding, 2011; Ise, 2012; Fischer, 2013). Therefore, it is unclear to what extent targeted mathematics interventions are effective, and which variables might explain the differing findings. Variations in participant characteristics and study designs may be implicated. For instance, one review included only children with specific MLD (Kroesbergen, 2003), while other reviews pooled results across those with specific MLD and those with MLD in the context of general cognitive disabilities (Xin, 1999; Gersten, 2009; Codding, 2011). Furthermore, all but one of the reviews (Gersten, 2009) combined effect sizes across a wide range of study designs, including: RCTs, single-
case designs, pre-post (non-controlled) studies, and non-matched controlled studies. Therefore, included studies were of varying quality, with single-case designs notable for producing particularly elevated effect sizes (Busse, 1995). The review by Gersten (2009) only included RCTs and treatment versus treatment designs but effect sizes were pooled together.

In addition, although many RCTs indicate that such interventions may be more effective than comparators for most students with MLD, they do not work for all children (Fuchs 2008). As indicated earlier, the heterogeneous profile of MLD in children, along with inadequate screening of individual deficits using standardised mathematics tests, may sometimes mean that a group-based intervention is not sufficiently targeted to the needs of every child (Dowker, 2005). Thus it is important to compare the effectiveness of individual versus group-based mathematics interventions for children, as well as considering interventions that combine both individual and group elements.

Variations in effectiveness may also be related to further intervention and population characteristics. For instance, evidence for the relationship between effect size and training intensity/duration of the intervention is mixed: two reviews found that shorter interventions were more effective (Kroesbergen, 2003; Gersten, 2009), another reported no relationship between effect size and treatment duration (Fischer, 2013), whereas the meta-analysis by Codding 2011 found that more complex treatments (3+ treatment elements) implemented for less than 30 sessions produced higher effect sizes. With regard to participant characteristics, younger participants may derive more benefit from interventions than older children (Shalev, 2004; Gersten, 2009). Furthermore, preliminary evidence indicates that inequity variables (e.g., low socio-economic status [SES], gender, and place of residence) may be linked to poorer outcomes from a targeted mathematics intervention (Siegler, 2008; Wilson, 2009). Indeed, such inequity variables may reduce access to appropriate interventions or education. For instance, many less developed countries lack teachers with mathematical training at second level education, and student participation is low, particularly among females (Royer, 2007; Anderson, 2009; Price, 2013). Thus, there is a need for a review to include high quality study designs and to undertake moderator analyses that may help explain variations in intervention effectiveness.

**Behavioural and Psychological Interventions**

As indicated earlier, behavioural and psychological interventions typically target 'academic enablers' that have been shown to mediate academic competence, including mathematical performance (Swanson, 1998; Fuchs, 2010; Geary, 2012). Academic enablers include the development of student motivation, engagement, in-class attention and appropriate social and behavioural skills, as well as the utilisation of cognitive, behavioural and therapeutic techniques to address math anxiety and school phobia. It is particularly important that such enablers are addressed within children with MLD given the empirical evidence that such children differ from normally achieving children in terms of behaviour and personality variables, including low self-efficacy, locus of control, and temperament (Greenwood, 2002;
Shechtman, 2005). School-based/classroom-management interventions, cognitive-behavioural therapies, and individual counselling/therapy use a variety of behavioural and psychotherapeutic techniques that may target the academic enablers that mediate academic performance. These techniques include modelling, coaching, rehearsal, reinforcement, self-talk, relaxation, reflective listening, sharing of emotions and cognitive reframing strategies (Shechtman, 2005; Wilson, 2007; Kaufmann, 2012; McGilloway, 2012).

Non-invasive Brain Stimulation

MLD appear to be associated with domain-specific deficits in mathematics, as well as with general cognitive deficits (Geary, 2012). NFB is hypothesised to improve mathematical performance through modulating attention-deficit cognitive disturbances associated with MLD (Gottfried, 2010). TDCS stimulates brain areas hypothesised to directly impact on mathematical abilities, as well as affecting general cognitive deficits in the domains of attention, working memory, language and executive functions, which are also associated with MLD (Cohen Kadosh, 2010). More particularly, research indicates that the effects of TDCS are protein-synthesis dependent and are accompanied by several mechanisms, including the modifications of intracellular cyclic adenosine monophosphate (cAMP) and calcium levels, brain-derived neurotrophic factor, and activation of adenosine A1 receptors (Márquez-Ruiz, 2012). Magnetic resonance spectroscopy (MRS) studies also found change in the local concentration of gamma-aminobutyric acid (GABA) and glutamate (Cohen Kadosh, 2010). These mechanisms are associated with the acquisition of automatic number processing and the mapping of number into space, both important indices of numerical proficiency, as well as sharing some features with long-term potentiation (Stagg, 2009; Castillo, 2011).

Pharmacological Interventions

MLD have a high rate of comorbidity with ADHD, and to a lesser extent with anxiety and depression (Kaufmann, 2012). Therefore, it is not surprising that deficits in non-numerical cognitive functions, such as attention and working memory, are implicated in child MLD (Fuch, 2010; Geary, 2012). Furthermore, math anxiety and disruptive behaviour have been shown to mediate mathematical outcomes (Diperna, 2006; Fuchs, 2010). Pharmacology has been shown to improve attention, working memory, and internalising and externalising behaviours in children with MLD (Rubinsten, 2008), which may enhance their capacity to engage with mathematical instruction.

Multi-component Interventions

The rationale for, and examples of, multi-component interventions was outlined in Description of the intervention. MLD are multi-faceted and heterogeneous, and programmes that target both numerical and associated non-numerical deficits (e.g., attention, classroom behaviour, self efficacy) may potentially be more efficacious than a programme that targets just one of these domains (Zentall, 2007; Kaufmann, 2012).
**Why it is Important to do this Review**

Numerical competency is vital to thrive in a modern economy; yet compared to what has been achieved for reading difficulties, very little is known about how best to support students who suffer from MLD (Fuchs, 2008). In line with emerging evidence that MLD may derive from both numerical and non-numerical deficits, this review is unique in examining the comparative effectiveness of a range of interventions that may potentially improve mathematical performance within students with MLD. Furthermore, recent advances in the mathematical cognition literature indicate significant heterogeneity and componentiality in the profile of MLD; therefore, unlike previous reviews, the current review will investigate programme effectiveness in relation to the moderating influence of subtypes of MLD, severity of MLD at baseline, as well as to the type of mathematical skill targeted within the intervention (e.g., precursor skills, basic computation, word problem solving, and more advanced skills [e.g., algebra]). Other population and intervention characteristics associated with effectiveness will also be investigated; for instance, age of students, treatment modality, and programme intensity. In addition, unlike other reviews in the field (Xin 1999; Slavin 2009; Coddington 2011; Ise 2012; Fischer 2013), the current review will only include randomised controlled trials and high quality quasi-experimental studies and will incorporate an assessment of risk of bias of included studies into the analysis and interpretation of results. Lastly, this review will be the first to assess secondary outcomes associated with MLD (e.g. emotional and behavioural problems; costs data). In conclusion, it is hoped that this review will inform improved clinical decision-making and policy within the field, improve outcomes for affected children and families, while also stimulating much needed research into the development and evaluation of effective interventions for students with MLD.

**OBJECTIVES**

To examine the comparative effectiveness of a range of interventions for students with mathematical learning difficulties (MLD) in improving mathematical performance compared to a control group of standard schooling, no treatment, waiting list or placebo.

To explore moderators of intervention effectiveness in relation to characteristics of: population (subtype of MLD, severity of MLD at pretest, age and equity variables); outcome (type of mathematical skill); and intervention (duration, individual versus group-based).
METHODS

Criteria for Considering Studies for this Review

Types of Studies

Included: Randomised and cluster randomised controlled trials. Cross-over trials. Quasi-experimental studies (assigned non-randomly by alternation, date, clinician judgment, etc), involving at least one treatment and one control group, which (i) provide evidence of baseline equivalence of pretest mathematical performance (using effect size magnitude) or (ii) which statistically control for pre-test mathematical performance differences in their analyses (e.g. covariate-adjusted means).

Excluded: Single-group pretest-posttest studies or other study designs. Treatment versus treatment studies without a control group of standard schooling, no treatment, waiting list or placebo. Qualitative studies.

Types of Participants

School-aged children up to 18 years old whose mathematical competence is at a ‘clinical’ or ‘at risk level’, irrespective of cause. We will define ‘clinical’ as scoring below the 10th percentile and being ‘at risk’ as scoring below the 25th percentile on a standardised mathematics test such as the Weschsler Individual Achievement Test (WIAT-II, Wechsler 2005). We will also include studies which provide definitions of MLD according to state regulations or to the 2004 Individuals with Disabilities Education Act (Fuchs 2004). School-aged children may include children in kindergarten/preschool.

To investigate the moderating influence of MLD subtypes on intervention effectiveness (e.g. isolated MLD or MLD in context of wider cognitive deficits), general intelligence will be defined as ‘normal’ if students score 85 or above on a validated measure of intelligence (e.g. the Wechsler scales; Wechsler 2003).

Studies that include students without MLD will be included if separate outcome data are presented so that effect sizes may be calculated separately for students with MLD.

Study authors will be contacted where required information is not specified in the report.

Types of Interventions

Targeted mathematics interventions, behavioural and psychological interventions, non-invasive brain stimulation technologies, pharmacological interventions and multi-component interventions that involve two or more elements of the preceding interventions.
Types of Outcome Measures

Primary Outcomes

General mathematical skill/proficiency (as measured by performance on precursor skills, arithmetical computation, word problem solving or more advanced mathematics);

Adverse impact - (i) side effects from NIBS or pharmacological interventions; (ii) increased psychological distress from participating in the intervention.

Secondary Outcomes

Internalising problems (e.g. anxiety, depression);

Externalising problems (e.g. aggression, defiance);

Hyperactivity and/or attention symptoms;

User satisfaction; and

Costs and cost-effectiveness data; we will only costs studies conducted alongside or in association with eligible effectiveness studies.

Timing of Outcomes

We will conduct separate meta-analyses for outcomes measured at several periods of follow-up: within the first six months, between 6 to 12 months post-intervention, between 12-24 months post-intervention (and longer follow-up points, if such data are available).

Example of Eligible Study

A primary study that is likely to meet the inclusion criteria is a randomised controlled trial of 133 third-grade students with MLD assigned to one of two targeted mathematics interventions ('Math Flash' and 'Pirate Math') or a no-tutoring control group (Fuchs, 2009). Pretest and posttest outcomes were fluency of number computation, procedural calculations, word problems and algebra.

Search Methods for Identification of Studies

We will identify studies through key and text word searches of electronic databases, as well as searching grey literature (conference papers, unpublished PhD theses, reference and citation lists of included studies and other relevant reviews, handsearching relevant journals) and personal communication with experts in the field. No date, publication, geographic or language restrictions will be applied to the searches. Searches will be updated within 12 months of publication of the final review.
Electronic searches

We will search the following databases.

Cochrane Central Register of Controlled Trials (CENTRAL; current issue; part of The Cochrane Library), and which includes the Specialised Register of the Cochrane Developmental, Psychosocial and Learning Problems Group. Ovid MEDLINE (1946 to date). EMBASE (1947 to date; Elsevier). Academic Search Complete (1887 to date; EBSCOhost). ASSIA (Applied Social Sciences Index and Abstracts; 1987 to date; ProQuest). CINAHL (1937 to date; EBSCOhost). ERIC (1966 to date; ProQuest). PsycINFO (1887 to date; EBSCOhost). Science Citation Index Expanded (1945 to date; Web of Science). Social Science Citation Index (1956 to date; Web of Science). Sociological Abstracts (1952 to date; ProQuest). ProQuest Dissertations and Theses (1986 to date; ProQuest). ClinicalTrials.gov (2000 to date). ISRCTN Registry (2000 to date). World Health Organization Clinical Trials Registry Platform (WHO ICTRP; 2004 to date). Database of Abstracts of Reviews of Effects (DARE; current issue; part of The Cochrane Library).

Economic sources

We will also search the following economic databases.

Econlit (1886 to date; EBSCOhost). Health Economic Evaluations Database (HEED; 2006 to date; Wiley Interscience). Health Technology Assessment Database (HTAD; current issue; part of The Cochrane Library). NHS Economic Evaluation Database (NHSEED; current issue; part of The Cochrane Library).

We have presented an example of a search strategy conducted in Ovid MEDLINE in Appendix 1. We will amend the search strategies for other databases as appropriate.

Searching other resources

We will examine the reference and citation lists of included studies and of reviews in the field in order to find other eligible studies. We will handsearch relevant journals. We will also contact the authors of included studies, as well as experts working in the area, in order to identify unpublished and ongoing studies.

Data Collection and Analysis

Selection of studies

Two review authors (MF and FMcL) will independently assess titles and abstracts identified through searches in order to determine their potential eligibility for inclusion in the review. We will retrieve any citation deemed potentially relevant by at least one review author in full text and, again, MF and FMcL will independently assess it for inclusion. We will discuss any inter-rater differences to resolve discrepancies. If necessary, we will contact the study...
authors to obtain additional information. We will resolve disagreements by consensus with a third review author (SMcG). We will document the specific reasons for exclusion for each study that might reasonably have been expected to have been included but which did not meet the inclusion criteria.

Data extraction and management

Two review authors (MF and FMcL) will independently code/extract data from all of the included studies with a piloted data extraction form. They will extract information on study design and implementation, setting, sample characteristics, intervention characteristics and outcomes from all included studies. Regarding costs or cost-effectiveness studies, MF and FMcL will independently extract characteristics such as: year of study; details of interventions and comparators; study design; type of economic evaluation; source(s) of resource use; unit costs; decision making jurisdiction; geographical and organisational setting; analytic perspective; discount rates; and time horizon for both costs and effects (see Appendix 2 for pilot coding/data extraction form). Again, we will discuss any inter-rater differences to refine coding schemes and to resolve discrepancies. SMcG will arbitrate if the two review authors cannot reach agreement.

Where data are unavailable in the published trial reports, we will contact the study authors and ask them to supply the missing information. We will use Microsoft Excel to manage data before we enter it into Review Manager (RevMan) (RevMan, 2014) or Stata for the meta-regression analyses (Stata, 2013).

Assessment of risk of bias in included studies

Two review authors (MF and FMcL) will independently assess the risk of bias (as 'high', 'low' or 'unclear') within each included study across the seven domains listed below, using the criteria specified in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011a). We will provide relevant quotes from the study report and a justification for our judgement for each item in the 'Risk of bias' table. We will request missing data from the study authors. Where judgments differ, we will seek additional input from SMcG. We will not blind the review authors that perform the 'Risk of bias' assessments to the names of the authors, institutions, journals or results of studies.

Sequence generation (selection bias). Each participant should have an equal chance of getting into the intervention or control condition. The randomisation sequence should not be capable of being foreseen. Allocation concealment (selection bias). Adequate allocation concealment prevents those who admit participants into the trial from knowing to which group the participant will be assigned. Blinding of participants and personnel (performance bias). Adequate blinding reduces the risk that knowledge of which intervention was received, rather than the intervention itself, may affect outcomes. Blinding of outcome assessment (detection bias). If the assessor is aware of the allocation status of the participant, this may bias the recording of data. Incomplete outcome data (attrition bias). Systematic differences
between intervention and control groups in terms of attrition or exclusions may bias results. Selective outcome reporting (reporting bias). Outcomes with statistically significant results are much more likely to be reported or published, which adds bias. Other sources of bias (e.g., inappropriate influence of funders; contamination between intervention and control groups; carry-over, order or learning effects in cross-over trials).

If costs or cost effectiveness studies meet the inclusion criteria of the review, we will use the 'Drummond checklist' to critically appraise the methodological quality of included health economic studies (Shemilt, 2011). MF and FMcL will assess such quality parameters independently, using verbatim quotes, and will resolve any disagreements through discussion. Again, we will contact the study authors for missing information.

### Measures of treatment effect

#### Dichotomous data

Where outcomes are reported as dichotomous data, we will use risk ratios (RR) with a 95% confidence interval (CI) to summarise results within each study.

#### Continuous data

We will report continuous outcomes, where possible, as mean differences (MDs). Where different scales measure the same outcome across studies, we will use standardised mean differences (SMDs) to summarise results within each study. SMDs are based on Hedges’ g, which includes an adjustment for small sample bias. Again, we will use 95% CIs for individual study data and pooled estimates throughout. Where there are missing data (e.g. standard deviations (SDs) or numbers (Ns) not reported), it may sometimes be possible to calculate effect sizes through reported standard errors (SEs), t tests, F tests, exact P values and so forth, using the formulas indicated in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins, 2011b), and using the RevMan calculator (RevMan, 2014). Where this is not possible, we will contact the study authors for missing information. In cases where SDs are unavailable and we cannot calculate them, we will impute an average SD from other included studies, as this method has been found to produce approximately correct results (Higgins, 2011c). We will assess the extent to which imputing the average SD alters results in a sensitivity analysis (see the 'Sensitivity analysis’ section). We will analyse endpoint data for continuous outcomes.

Across both dichotomous and continuous data, we will use a random-effects meta-analysis that employs the inverse-variance weight function, due to expected heterogeneity. We will use forest plots to visually display the results. If reported outcomes have insufficient data, we will request additional information from the study authors.

Where studies report the same outcome construct (e.g., mathematics skill) using both dichotomous and continuous data, we will transform the effect size metric with the smaller
proportion into the metric with the larger proportion, using current Cochrane guidance for transforming odds ratios (ORs) into SMDs and vice versa (Deeks, 2011). This will allow us to analyse all effect sizes for that outcome category together. However, in reporting within the review, we will transform ORs into RRs to ease interpretation of the findings using RevMan tools (RevMan, 2014).

Economic evaluation

We will initially classify studies according to whether they are partial or full economic evaluations, that is, whether they just measure resource costs or also calculate cost-benefit ratios or an incremental cost effectiveness ratio (ICER). An ICER point estimate compares the costs and consequences of implementing an intervention relative to the costs and consequences of a specified alternative (most commonly chosen to be the status quo). In circumstances where there is evidence of little variation in resource or cost use between studies, it may be regarded as legitimate to present a pooled estimate. Otherwise, we will clearly present the distribution of costs (Shemilt, 2011). If we decide to conduct meta-analyses of resource use or cost data, we will support this with a thorough critical appraisal of the methods used to derive such estimates within the corresponding health economics studies, alongside use of 95% CIs and statistical methods to investigate and incorporate between-study heterogeneity (e.g., I² statistic, Chi² test, random-effects models). We will adjust the cost estimates collected from multiple studies to a common currency and price year before we pool these data. We will give careful consideration to the jurisdiction, analytic perspective and time horizon for both costs and effects.

If we conduct meta-analyses of resource use or cost data, we will include a narrative summary in the 'Results' section to comment on the direction and magnitude of results and their precision. Similarly, if two or more health economics studies meet the inclusion criteria of this review but we decide not to pool the resource use or cost data, or both, from these studies, we will outline the reasons for not combining the data in the 'Methods' section (Shemilt, 2011).
**Unit of analysis issues**

**Multiple reports**

As some studies may be reported in multiple reports or multiple reports reported in single studies, we will take care to ensure that the studies are reporting independent findings. If it is unclear whether reports and studies provide independent findings, we will contact the authors of the reports.

**Multiple observations for same outcome**

If studies provide more than one effect size for a particular outcome construct (e.g. a study reports two psychometrically valid measures of mathematical skill), we will combine effect sizes to create a single pairwise comparison; that is, we will obtain a mean effect size and SE for that outcome within the study using robust SEs that account for statistical dependencies (Hedges 2010).

If studies report multiple measures of the same construct at different points in time, we will conduct separate meta-analyses for outcomes measured at several periods of follow-up: within the first six months, between six and 12 months post-intervention, and between 12 and 24 months post-intervention (and longer follow-up points, if such data are available). If, within any of these periods, the included studies report measures more than once, then we will obtain a single summary effect within that time period.

**Multiple groups from one study**

In multi-arm studies, where several experimental groups are eligible for inclusion (e.g. two variants of a targeted mathematics intervention), we will combine effect sizes to create a single pair-wise comparison (Higgins 2011c). For dichotomous data, we will sum the sample sizes and events across groups. For continuous data, we will combine sample sizes, means and SDs according to the formula detailed in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011b).

**Cluster randomised trials**

If included cluster RCTs do not appropriately adjust for clustering, we will control for clustering using the procedures outlined by Higgins 2011c in the Cochrane Handbook for Systematic Reviews of Interventions. When outcome measures are dichotomous, we will divide the number of events and number of participants per trial arm by the design effect \[1 = (1 - m) * r\], where \(m\) is the average cluster size and \(r\) is the intraclass correlation coefficient (ICC). When outcome measures are continuous, we will divide the number of participants per trial arm by the design effect, with the mean values unchanged. To determine the ICC, we will use estimates from the primary trials on a study-by-study basis. If these values are not reported, we will use external estimates of the ICC that are appropriate to each trial context and average cluster size.
Cross-over trials

A unit of analysis error may occur if a paired analysis is not reported within cross-over trials. A common approach to incorporation of cross-over trials in a meta-analysis is to take all measurements from intervention E periods and all measurements from intervention C periods and analyse these as if the trial were a parallel group trial of E versus C. This method may underweight the study but this type of unit-of-analysis error is generally viewed as less serious than other types of unit-of-analysis errors (Higgins, 2011c).

Dealing with missing data

For studies with participants missing from trial analyses or incomplete outcome data (owing to attrition or exclusion), we will contact first-named study authors via email with a request for further information, that is, individual trial data or reasons for missing data, or both. If data are available, we will conduct analyses including the participants who the study authors excluded. We will note the study authors' approach to imputation, and if we deem it inappropriate, we will impute missing values assuming a worse case scenario, that is, the pre-test score when the student had MLD. If data are unavailable, we will conduct analyses using available data only and will not impute values. Within the 'Risk of bias' tables, we will report the extent of, and reasons for, missing data in the treatment and control group as a proportion of randomised participants in each study. We will conduct a sensitivity analysis (see the 'Sensitivity analysis' section) where we cannot assume that data are missing at random, attrition is higher than 20% or where the primary study did not conduct an appropriate intention-to-treat (ITT) analysis. We will use the definition provided in the Cochrane Handbook for Systematic Reviews of Interventions, where an ITT analysis aims to include all randomised participants in the trial regardless of what happened subsequently (Higgins, 2011a). We will also describe in the 'Discussion' section the extent to which the results might be biased by missing data.

Assessment of heterogeneity

We will assess clinical heterogeneity by comparing the distribution of important factors such as participant demographics, type of intervention and control comparators and outcomes measured across studies. We will assess methodological heterogeneity using the 'Risk of bias' table. We will assess statistical heterogeneity visually and by examining the I^2 statistic, a quantity which describes the approximate proportion of variation in point estimates that is due to heterogeneity rather than sampling error. We will supplement this by the Q or Chi^2 test, where a P value less than (<) 0.05 indicates heterogeneity of treatment effects (i.e. studies do not share a common effect size) (Deeks, 2011). In addition, we will estimate and present Tau^2, along with its CIs, as an estimate of the magnitude of variation between studies. This will provide an estimate of the amount of between-study variation. The three measures of heterogeneity (Q, Tau^2 and the I^2 statistic) complement each other: we will use Tau^2 and the I^2 statistic to measure the magnitude of true variation, and the P value for Q or CIs for Tau^2 or the I^2 statistic as an indicator of uncertainty regarding the genuineness of the
heterogeneity. We will also conduct sensitivity and subgroup analyses to explore possible sources of heterogeneity (see the 'Subgroup analysis and investigation of heterogeneity' and 'Sensitivity analysis' sections).

For the economic evaluations, we will give careful attention to whether the metric in question has equivalent meaning across studies before we pool data (Shemilt, 2011). We will adjust the cost estimates collected from multiple studies to a common currency using purchasing power parity and price year before we pool these data. We will assess between-study heterogeneity using the Chi² test, Tau², the I² statistic and random-effects models.

**Assessment of reporting biases**

We will assess reporting bias in the form of selective outcome reporting as one of the domains within the 'Risk of bias' assessments. Where there are more than 10 studies included in a meta-analysis, we will construct funnel plots to investigate any relationship between effect size and SE. Such a relationship could be due to publication or related biases, or due to systematic differences between small and large studies. Where we identify such a relationship, we will use Egger’s test to test for funnel plot asymmetry (Egger, 1997). We note that asymmetrical funnel plots (small study effects) are not always indicative of the presence of publication bias (Sterne, 2011).

**Data synthesis**

We will use the summary and descriptive statistics of the study-level contextual characteristics, methodological quality characteristics, and participant and intervention characteristics to describe the included studies.

We will perform meta-analyses where studies have sufficiently similar participants, comparators, outcome measures and summary points (outcomes should be measured within comparable time frames). We will use SMD effect sizes for outcomes on continuous measures and RRs for outcomes presented as dichotomous variables. We will calculate random-effects, weighted mean effect sizes for all studies using 95% CIs and display them with forest plots. We will use the estimates of Cochrane’s Q, the I² statistic and Tau² to assess variability in the effect sizes.

We will conduct main effects and moderator analysis separately on each outcome construct with the latter done as multivariate (meta-regression) analysis when there are sufficient studies (at least 10 studies; Deeks, 2011). The main objective of the analysis will be to describe the direction and magnitude of the effects of the various MLD interventions on the different outcome constructs in a manner that allows us to assess their comparative effectiveness. Additionally, the moderator analyses will attempt to identify the characteristics of participants, interventions and outcomes that are associated with larger and smaller effects on the primary and secondary outcomes (see the 'Subgroup analysis and investigation of heterogeneity' section). We will use RevMan to analyse the main effects.
(RevMan, 2014), and meta-regression in Stata to conduct the moderator analyses (Stata, 2013).

**Subgroup analysis and investigation of heterogeneity**

The primary objective of this Cochrane review is to investigate the comparative effectiveness of different interventions in improving MLD. In addition, as indicated in the 'Background' section, several other participant and intervention characteristics are of interest in understanding how and for whom interventions appear to be most effective. However, a danger of including too many predictor variables in a meta-regression is that they use up degrees of freedom and consequently the analysis may be underpowered. It is recommended that meta-regression should not be undertaken when there are fewer than 10 studies in a meta-analysis (Deeks, 2011); or indeed where there are fewer than 10 units per predictor variable. As a result, it is necessary to know how many studies are included in a meta-analysis in order to know the number of moderators that we can explore. Therefore, we have outlined a list of prespecified moderators in this section that are of interest, but we will withhold judgement on the number that can be investigated until we know how many studies are included in the meta-analysis.

Another issue in conducting moderator analyses is that predictors may often be dependent and correlated with each other. In such cases, it may be useful to group certain predictors together and create a composite indicator in order to preserve degrees of freedom. In the context of this Cochrane review, for example, we may usefully group moderators of age; socioeconomic status; living in low-, middle- or high-income countries as an inequity variable.

In addition, when several intervention groups within a study are eligible for inclusion in a review but differ according to a moderator variable (e.g., treatment duration), in order to conduct our planned moderator analysis, it will be necessary to include each pair-wise comparison separately for each intervention group within the study. We will divide out the sample size of shared intervention groups approximately evenly among the comparisons. This method only partially overcomes a unit of analysis error but is a practical means of performing approximate investigations of heterogeneity (Higgins, 2011c).

Furthermore, it should be noted that meta-regression analyses are observational in nature. If we identify a sufficient number of studies or if information is available, we will examine the following moderators for their influence on effect sizes.

Type of intervention: targeted mathematics interventions, behavioural and psychological interventions, NIBS technologies, pharmacological interventions and multi-component interventions. Mathematical skill targeted: precursor number skills (e.g. counting, magnitude, spatial number representation, number conservation, classification, seriation and so forth), number computation/arithmetic reasoning (addition, subtraction,
multiplication, division), word problem-solving, and more advanced mathematics (algebra, calculus). Subtype of MLD. Severity of MLD at pre-test: clinical or at risk. Age of child. Inequity composite variable, including gender, socioeconomic status, country of residence. Treatment duration: interventions comprising zero to 10 hours versus those with 10 to 20 hours, 20 to 30 hours or more than 30 hours. Treatment modality: individual tailored versus group-based programme (small group or classroom) or combination of individual and group content.

We will perform meta-regression using the 'metareg' macro available for the Stata statistical package (Stata, 2013). We will use a random-effects meta-regression. We will report the number of predictors and studies in the model, Q model sum of squares, Q residual sum of squares, along with the regression coefficients, CIs and test of significance.

Sensitivity analysis

We will perform the following sensitivity analyses to examine the robustness of decisions we make in the review by excluding the following.

Quasi-RCTs, cluster-RCTs if not adjusted for clustering, and carry-over RCTs with likely order or learning effects. Studies where we cannot assume missing participants to be missing at random, or with attrition rates higher than 20%. Studies where authors imputed SDs or ICCs for missing data. Studies with outlier effect sizes. Unpublished studies.

If necessary, we will conduct additional sensitivity analyses if other issues arise that may impede our confidence in the estimated pooled effect sizes.

'Summary of findings' table

We will construct a 'Summary of findings' table in RevMan (RevMan, 2014) and will present the magnitude of the interventions examined for each outcome, the number of participants and the quality of evidence for outcomes using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) criteria (Schünemann, 2011a; Schünemann, 2011b). We will present results for data reported at zero to six months post-intervention and after six months post-intervention. Within the constraints of the review, we will only include eligible RCTs and quasi-RCTs that report adverse effects; we will report on the limitations of this approach within the review.

Using the GRADE approach we will assess the quality of the evidence for each outcome using the factors outlined in the Cochrane Handbook for Systematic Reviews of Interventions: risk of bias, inconsistency, imprecision, indirectness and magnitude of effect (Higgins, 2011a). Two review authors (MF and FMcL) will independently judge the quality of the evidence for each outcome as one of the following.

'High' - further research is very unlikely to change our confidence in the estimate of effect. 'Moderate' - further research is likely to have an important impact on our confidence in the
estimate of effect and may change the estimate. 'Low' - further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. 'Very low' - we are very uncertain about the estimate of the effect.

REFERENCES


Stata (2013). Stata Statistical Software [Computer program]. Version Release 13. College Station, TX: StataCorp LP.


APPENDICES

1 Ovid MEDLINE search strategy

1 Dyscalculia/2 Learning Disorders/3 Developmental Disabilities/4 Intellectual Disability/5 aculculia.ab.6 development* dyscalculi*.ab.7 (learn* adj3 (disabilit* or difficult* or problem*)).ab.8 (math* adj3 (learn* disabilit* or learn* difficult*).ab.9 mld.ab.10 (math* adj3 (disabilit* or difficult* or disorder* or perform* or competenc* or underachiev* or low perform* or achiev* or outcome* or operation* or reason* or skill* or educat*)).ab.11 (numer* adj3 (disability* or difficult* or disorder* or perform* or competenc* or underachiev* or low perform* or achiev* or outcome* or operation* or reason* or skill* or educat*)).ab.12 (computati* adj3 fluen*).ab.13 (math* adj3 fluen*).ab.14 (numer* adj3 fluen*).ab.15 Low* numer*.ab.16 mental* retard*.ab.17 (number* adj3 concept*).ab.18 or/1-1719 exp Child/20 Adolescent/21 child*.ab.22 adolescen*.ab.23 or/19-2224 (math* adj3 (train* or intervention* or teach* or instruct* or educat*)).ab.25 (arithmetic* adj3 (train* or intervention* or teach* or instruct* or educat*)).ab.26 (numer* adj3 (train* or intervention* or teach* or instruct* or educat*)).ab.27 (cognit* adj3 (train* or intervention* or teach* or instruct* or educat*)).ab.28 (educat* adj3 intervention*).ab.29 addition intervention*.ab.30 subtraction intervention*.ab.31 multiplication intervention*.ab.32 division intervention*.ab.33 element* educat*.ab.34 primary educat*.ab.35 second* educat*.ab.36 (comput* adj3 (intervention* or program* or instruct*)).ab.37 digital* adj3 intervention*.ab.38 (school* adj3 (intervention* or program* or instruct* or train*)).ab.39 (classroom* adj3 (intervention* or instruct* or train*)).ab.40 Electric Stimulation/41 transcranial electrical stimulation.ab.42 tes.ab.43 non invasive brain stimulation.ab.44 tcds.ab.45 transcranial direct current stimulation.ab.46 neurofeedback.ab.47 Psychotherapy/48 psychotherap*.ab.49 behavior therapy/ or cognitive therapy/50 (behavior* adj3 therap*).ab.51 (cognitive adj3 (therap* or train* or intervention* or program*)).ab.52 Pharmacy/53 Pharmacology/54 pharmac*.ab.55 methylphenidate.ab.56 Clinical Medicine/57 or/24-5658 18 and 23 and 57

2 Pilot data extraction/coding form

Interventions for children with mathematical learning difficulties (MLD) review

Study ID#: ____________ Coder: ____________ Date: ____________

APA Citation:____________________________________________________

Section A: Eligibility screening

Study name

Author
Year

Study linked to multiple reports?
If yes, name all reports

Email details of first author

Country

Is the report a:
P = Primary evaluation study RE = Review (Effect/meta-analysis) RD = Review (Descriptive)
D = Descriptive T = Theoretical paper O = Other

Are the participants receiving an eligible intervention?
Targeted mathematics intervention
Behavioural/psychological intervention
NIBSPharmacology
Multiple components intervention
(Yes, no, unclear)
Is the study a:
Parallel group RCT with a control group? (Yes, no, unclear)
Cluster-RCT? (Yes, no, unclear)
Cross-over RCT? (Yes, no, unclear)
Non-randomized controlled study with a control group? (Yes, no, unclear)

Was baseline equivalence between groups established in relation to pretest mathematics scores? (Yes, no, unclear)

If groups were not equivalent, were appropriate statistical controls used?
(Yes, no, unclear)

Does the study measure the primary outcome of mathematical skill?
(Yes, no, unclear)

Does the study involve students with a defined level of MLD, either clinical or at risk?
(Yes, no, unclear)
Are the participants in school?
(Yes, no, unclear)

Are the participants less than 18 years old?
(Yes, no, unclear)

State questions to be sent to study authors if unclear

Abbreviations: APA: American Psychological Association; ID: identifier; MLD: mathematical learning difficulties; NIBS: non-invasive brain stimulation; RCT: randomised controlled trial.

Section B: Study design

Study name

Is the study published or unpublished?

Unit of assignment to conditions:
Individual student
Group/cluster (specify)
Other

How many sites in study?

If applicable, was the method of assignment the same in each site?
(Yes, no, unclear)

Number randomised/assigned to each condition

Section C: Participant descriptors

Study name

Intervention(s)

Severity of MLD at pretest:
Clinical (10% or below)
At risk (11% to 25%)Unclear
Subtype of MLD at pretest:
MLD alone
MLD in context of other cognitive and IQ deficits
MLD comorbid with ADHD, dyslexia or emotional and behavioural problems
Unclear

Was IQ level reported? (Yes, no, unclear)

Did the study include students without MLD? (Yes, no, unclear)

Did the study report separate effect sizes for students with MLD?
(Yes, no, unclear)

Age of child

Gender (e.g. % male and female)
Socioeconomic status

Race/ethnicity

High, middle or low income country

Comparison group

Severity of MLD at pretest:
Clinical (10% or below)At risk (11% to 25%)Unclear

Subtype of MLD at pretest:
MLD alone
MLD in context of other cognitive and IQ deficits
MLD comorbid with ADHD, dyslexia or emotional and behavioural problems
Unclear
Was IQ level reported? (Yes, no, unclear)

Did the study include students without MLD? (Yes, no, unclear)

Did the study report separate effect sizes for students with MLD? (Yes, no, unclear)

Age of child (mean, SD)

Grade of child

Gender (e.g. % male and female)

Socioeconomic status
Race/ethnicity

High, middle or low income country

Abbreviations: ADHD: attention deficit hyperactivity disorder; IQ: intelligence quotient; MLD: mathematical learning difficulty; SD: standard deviation.

Section D: Intervention descriptors

Study name

List the type of intervention(s):
Targeted mathematics intervention (list variants also) Behavioural/psychological intervention NIBS Pharmacology Multiple components intervention Unclear

Briefly describe content of intervention

Location of intervention
Duration of intervention (sessions/hours)

Frequency of intervention

Intensity of intervention

Intervention modality:
Individual/Group/class/school/Combination/Unclear

Who delivers the intervention?

Evidence of implementation fidelity?
(Yes, no, unclear)
Other important information

Abbreviations: NIBS: non-invasive brain stimulation.

Section E: Control group descriptors

Study name

Name the control/comparison condition
Wait list No treatment Standard schooling Placebo

Briefly describe content of comparator

Location of comparator

Duration of comparator (sessions/hours),
if applicable

Frequency of comparator (if applicable)
Intensity of comparator (if applicable)

Comparator modality (if applicable)
Individual Group/class/school Combination Unclear N/A

Who delivers the comparator? (if applicable)

Evidence of implementation fidelity, if applicable?
(Yes, no, unclear, NA)

Other important information

Abbreviations: NA: not applicable.

Section F: Dependent measures and effect sizes
Primary outcome: mathematical skill

Identify type(s) of mathematical skill measured:
Precursor skill (specify)
Computational skill (specify)
Word problem-solving
Advanced mathematics (specify)
Unclear

Reliability of outcome measure used

Validity of outcome measure used

Number of follow-up data collection points

Time frame of data collection points

Statistical test type
Appropriateness of statistics
(Yes, no, unclear)

Means adjusted
(Yes, no, unclear)

Statistical value (t test, F value)

Degrees of freedom

Reported effect size and standard error (Note: may be multiple effect sizes depending on number of measurements and number of eligible experimental groups included - see rows in bold below)

Outlier effect size
(Yes, no)

P value

Number and effect size for follow-up 1 for intervention and control conditions

Number and effect size for follow-up 2 for intervention and control conditions

Number and effect size for follow-up 3 for intervention and control conditions

As treated or intention-to-treat analysis?
Number randomised and number analysed in intervention and control groups
Per cent missing data in intervention and control groups

*Repeat questions if there are a number of eligible experimental groups in the study or if there are a number of subscales reported, or both, at each time point
*Calculate single pairwise comparison effect size if there are a number of eligible experimental and control conditions in study. Separate effect sizes for each time period (less than 6 months, 6 to 12 months, 12 to 24 months, later)

Section F: Dependent measures and effect sizes

Study name

Primary outcome: adverse effects

Identify type(s) of adverse effect measured

Reliability of outcome measure used

Validity of outcome measure used

Number of follow-up data collection points
Time frame of data collection points

Statistical test type

Appropriateness of statistics
(Yes, no, unclear)

Means adjusted
(Yes, no, unclear)

Statistical value (t test, F value)

Degrees of freedom
Reported effect size and standard error (Note: may be multiple effect sizes depending on number of measurements and number of eligible experimental groups included - see rows in bold below)

Outlier effect size
(Yes, no)

P value

Number and effect size for follow-up 1 for intervention and control conditions

Number and effect size for follow-up 2 for intervention and control conditions

Number and effect size for follow-up 3 for intervention and control conditions

As treated or intention-to-treat analysis?
Number randomised and number analysed in intervention and control groups
Per cent missing data in intervention and control groups

*Repeat questions if there are a number of eligible experimental groups in the study or if there are a number of subscales reported, or both, at each time point

*Calculate single pairwise comparison effect size if there are a number of eligible experimental and control conditions in the study. Separate effect sizes for each time period (less than 6 months, 6 to 12 months, 12 to 24 months, later)

Section G: Dependent measures and effect sizes

Study name

Secondary outcome: internalising problems (anxiety, depression)

Identify type(s) of internalising problem measured

Reliability of outcome measure used
Validity of outcome measure used

Number of follow-up data collection points

Time frame of data collection points

Statistical test type

Appropriateness of statistics
(Yes, no, unclear)

Means adjusted
(Yes, no, unclear)

Statistical value (t test, F value)
Degrees of freedom

Reported effect size and standard error (Note: may be multiple effect sizes depending on number of measurements and number of eligible experimental groups included - see rows in bold below)

Outlier effect size
(Yes, no)

P value

Number and effect size for follow-up 1 for intervention and control conditions

Number and effect size for follow-up 2 for intervention and control conditions
Number and effect size for follow-up 3 for intervention and control conditions

As treated or intention-to-treat analysis?
Number randomised and number analysed in intervention and control groups
Per cent missing data in intervention and control groups

*Repeat questions if there are a number of eligible experimental groups in the study or if there are a number of subscales reported, or both, at each time point

*Calculate single pairwise comparison effect size if there are a number of eligible experimental and control conditions in the study. Separate effect sizes for each time period (less than 6 months, 6 to 12 months, 12 to 24 months, later)

Section G. Dependent measures and effect sizes

Study name

Secondary outcome: externalising problems (aggression, defiance)
Identify type(s) of externalising problem measured

Reliability of outcome measure used

Validity of outcome measure used

Number of follow-up data collection points

Time frame of data collection points

Statistical test type

Appropriateness of statistics
(Yes, no, unclear)
Means adjusted
(Yes, no, unclear)

Statistical value (t test, F value)

Degrees of freedom

Reported effect size and standard error (Note: may be multiple effect sizes depending on number of measurements and number of eligible experimental groups included - see rows in bold below)

Outlier effect size
(Yes, no)

P value
Number and effect size for follow-up 1 for intervention and control conditions

Number and effect size for follow-up 2 for intervention and control conditions

Number and effect size for follow-up 3 for intervention and control conditions

As treated or intention-to-treat analysis?
Number randomised and number analysed in intervention and control groups
Per cent missing data in intervention and control groups

*Repeat questions if there are a number of eligible experimental groups in the study or if there are a number of subscales reported, or both, at each time point

*Calculate single pairwise comparison effect size if there are a number of eligible experimental and control conditions in the study. Separate effect sizes for each time period (less than 6 months, 6 to 12 months, 12 to 24 months, later)
Section G: Dependent measures and effect sizes

Study name

Secondary outcome: hyperactivity/attention problems

Identify type(s) of problem measured

Reliability of outcome measure used

Validity of outcome measure used

Number of follow-up data collection points

Time frame of data collection points
Statistical test type

Appropriateness of statistics
(Yes, no, unclear)

Means adjusted
(Yes, no, unclear)

Statistical value (t test, F value)

Degrees of freedom

Reported effect size and standard error (Note: may be multiple effect sizes depending on number of measurements and number of eligible experimental groups included - see rows in bold below)
Outlier effect size

(Yes, no)

P value

Number and effect size for follow-up 1 for intervention and control conditions

Number and effect size for follow-up 2 for intervention and control conditions

Number and effect size for follow-up 3 for intervention and control conditions

As treated or intention-to-treat analysis?
Number randomised and number analysed in intervention and control groups
Per cent missing data in intervention and control groups
*Repeat questions if there are a number of eligible experimental groups in the study or if there are a number of subscales reported, or both, at each time point

*Calculate single pairwise comparison effect size if there are a number of eligible experimental and control conditions in the study. Separate effect sizes for each time period (less than 6 months, 6 to 12 months, 12 to 24 months, later)

Section G: Dependent measures and effect sizes

Study name

Secondary outcome: user satisfaction

Identify measure

Reliability of outcome measure used

Validity of outcome measure used
<table>
<thead>
<tr>
<th>Number of follow-up data collection points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time frame of data collection points</td>
</tr>
<tr>
<td>Statistical test type</td>
</tr>
<tr>
<td>Appropriateness of statistics</td>
</tr>
<tr>
<td>(Yes, no, unclear)</td>
</tr>
<tr>
<td>Means adjusted</td>
</tr>
<tr>
<td>(Yes, no, unclear)</td>
</tr>
<tr>
<td>Statistical value (t test, F value)</td>
</tr>
</tbody>
</table>
Degrees of freedom

Reported effect size and standard error (Note: may be multiple effect sizes depending on number of measurements and number of eligible experimental groups included - see rows in bold below)

Outlier effect size

(Yes, no)

P value

Number and effect size for follow-up 1 for intervention and control conditions

Number and effect size for follow-up 2 for intervention and control conditions

Number and effect size for follow-up 3 for intervention and control conditions
As treated or intention-to-treat analysis?
Number randomised and number analysed in intervention and control groups
Per cent missing data in intervention and control group

*Repeat questions if there are a number of eligible experimental groups in the study or if there are a number of subscales reported, or both, at each time point

*Calculate single pairwise comparison effect size if there are a number of eligible experimental and control conditions in the study. Separate effect sizes for each time period (less than 6 months, 6 to 12 months, 12 to 24 months, later)

Footnotes

Section H: 'Risk of bias' domain

Authors' judgement

(Low risk, high risk, unclear risk)

Support for judgement

Random sequence generation (selection bias)
Quote:

Email communication:

Comment:

Allocation concealment (selection bias)

Blinding of participants and personnel (performance bias)

Blinding of outcome assessment (detection bias)

Incomplete outcome data (attrition bias)
Selective outcome reporting (reporting bias)

Other sources of bias
Inappropriate influence from funders
Contamination between intervention and control groups
Evidence of carry-over, order or learning effects in cross-over RCTs
Appropriate adjusting for cluster trials

Abbreviations: RCTs: randomised controlled trials.

Section I: Costs study detail and 'Risk of bias' assessment - Drummond checklist

Issue addressed
(Yes, no, unclear, not applicable)

Explanation
(provide quote and comment)
Study design

1. The research question is stated

2. The economic importance of the research question is stated

3. The viewpoint(s) of the analysis are clearly stated and justified

4. The rationale for choosing alternative programmes or interventions compared is stated

5. The alternatives being compared are clearly described
6. The form of economic evaluation used is stated

7. The choice of form of economic evaluation is justified in relation to the questions addressed

Data collection

8. The source(s) of effectiveness estimates used are stated

9. Details of the design and results of effectiveness study are given
10. Details of the methods of synthesis or meta-analysis of estimates are given

11. The primary outcome measure(s) for the economic evaluation are clearly stated

12. Methods to value benefits are stated

13. Details of the participants from whom valuations were obtained were given
14. Productivity changes are reported separately

15. The relevance of productivity changes to the study question is discussed

16. Quantities of resource use are reported separately from their unit costs

17. Methods for estimation of quantities and unit costs are described

18. Currency and price data are recorded
19. Details of currency of price adjustments for inflation or currency conversion are given

20. Details of any model used are given

21. The choice of model used and the key parameters on which it is based are justified

Analysis and interpretation of results

22. Time horizon of costs and benefits is stated
23. The discount rate(s) is stated

24. The choice of discount rate(s) is justified

25. An explanation is given if costs and benefits are not discounted

26. Details of statistical tests and confidence intervals are given for stochastic data

27. The approach to sensitivity analysis is given
28. The choice of variables for sensitivity analysis justified

29. The ranges over which the variables are varied is justified

30. Relevant alternatives are compared

31. Incremental analysis is reported
32. Major outcomes are presented in a disaggregated as well as aggregated form

33. The answer to the study question is given

34. Conclusions follow from the data reported

35. Conclusions are accompanied by the appropriate caveats
SOURCES OF SUPPORT

Internal sources

Department of Psychology, National University of Ireland Maynooth, Ireland

Salary for MF. NUI Maynooth provides the time for MF to conduct the review

External sources

None, Other

DECLARATIONS OF INTEREST

Mairead Furlong has no known conflicts of interest. Fergal McLoughlin has no known conflicts of interest. Sinead McGilloway has no known conflicts of interest. David Geary is the coauthor of several RCTs that may potentially be eligible for inclusion in the review. He will not be involved in assessing study eligibility, data extraction or 'Risk of bias' assessments of studies that he has been involved in.

ACKNOWLEDGEMENTS

We thank Professor Sandra Jo Wilson (Editor of Campbell’s Education Co-ordinating Group) for her helpful suggestions and advice in writing this Cochrane protocol. We are also grateful to Dr Carlton Fong (Managing Editor) and Dr Sean Grant (former Managing Editor) for their support. In addition, we thank the Co-ordinating Editor (Professor Geraldine MacDonald) and Managing Editor (Dr Joanne Wilson) of the Cochrane Developmental, Psychosocial and Learning Problems Group for the co-registration of this review with Cochrane.

ROLES AND RESPONSIBILITIES

Mairead Furlong (MF) wrote the text of the protocol, with feedback from the rest of the review author team (Fergal McLoughlin (FMcL), Sinead McGilloway (SMcG) and David Geary (DG). All review authors were involved in development of the search strategy, and will liaise with the Information Retrieval Specialist at Campbell regarding the literature searches.

ACKNOWLEDGEMENTS

Acknowledge all the individuals and organizations contributing to the preparation of the protocol that are not identified in prior sections.
EXPECTED TIMEFRAME

A timetable with target dates for accomplishing the key tasks required to complete the review. If the review is not submitted within 24 months of protocol approval, the review area may be opened up for other authors.

PLANS FOR UPDATING THE REVIEW

Describe any plan for updating the review once it is completed.