# Cognitive Science Approaches in the Classroom: Protocol for a systematic review

Education Endowment Foundation

**Principal investigator:** Dr Thomas Perry

## 1. Introduction to the Review and this Protocol

The ***Cognitive Science Approaches in the Classroom*** Systematic Review will investigate approaches to teaching and learning inspired by cognitive science that are commonly used in the classroom, with a particular focus on acquiring and retaining knowledge.

The review comprises a core systematic review, specified in this protocol, as well as two sub-strands which support the main review as well as produce outputs in their own right.

- **Core Strand – 'Systematic Review'** – Systematic review of the available evidence reporting cognitive science-informed interventions inside classrooms. In addition to the impact evidence, this also includes a review of the 'translation' of the basic science during the analysis, looking across other strands (see below) and the evidence therein, making an assessment of the ecological validity of interventions reviewed in the core strand and their fidelity to a) the cited underpinning cognitive neuroscience research and b) the broader state-of-the-art in cognitive neuroscience (as per the 'underpinning cognitive science review').
- **Sub-strand 1 – 'Underpinning Cognitive Science Review'** – Review of literature in contemporary cognitive neuroscience demonstrating the underpinning research that is mobilised in interventions.
- **Sub-strand 2 – 'Practice Review'** – Review of practice documents and school data collection to identify and illustrate the applications of cognitive science in the classroom.

This protocol's core focus is to specify the aims and methods for the overall review, with a focus on the core review strand. It will also state the objectives and give an overview of the methods for the review sub-strands and describe how these inform and build on the core review, thereby giving greater transparency around how the Cognitive Science interventions and techniques in focus for the core review are identified, their basis in cognitive science examined, and their classroom operationalisation illustrated.

### *Cognitive Science and Education Advisory Group*

This protocol has been developed in collaboration with a diverse advisory group, which includes headteachers, cognitive neuroscientists, and experts in education research, policy, and practice. All members have a strong interest in the applications of cognitive science to educational contexts. The purpose of the AG is to contribute expertise relating to cognitive science, applied research, policy, and classroom practice. Furthermore, involving a wider team reduces bias when conducting systematic reviews (Utterly & Montgomery, 2017). The Advisory Group met prior to the submission of this protocol to discuss the aims and focus of the review (July 2020 – see Appendix 1). The group will meet approximately two more times during the project, at key points in the timeline, supplemented by ongoing opportunities for input via email communication. During these meetings, panel members provide their expertise and guidance on specific aspects of the project. The remaining two meetings that have been planned are focused on 1) discussing emerging findings from the review, and identify any areas where further investigation or clarity is needed, and 2) discussing the overall findings, their implications and the dissemination strategy once the review is largely complete.

# Cognitive Science Approaches in the Classroom: Protocol for a systematic review

**Principal investigator:** Dr Thomas Perry

Education Endowment Foundation

**Table of contents**

## 2. Background and review rationale

### a. *Scientific, Policy and Practical Background*

**Cognitive Science, Policy and Practice**

Learning sciences are self-evidently of great relevance to education. Learning sciences form an interdisciplinary field which draws on cognitive science, educational psychology, computer science, anthropology, sociology, information sciences, neurosciences, education, instructional design and numerous other areas (Sawyer, 2008). The central role of cognition in learning and the brain's role in processing and storing information arguably places cognitive science at the heart of the learning sciences, with many influential publications aimed at practitioners (see below) taking areas of cognitive science as their focus.

It is helpful to conceive of cognitive science as encompassing two areas of research:

- **cognitive psychology** underpinned by behavioural theory and observation, and
- **cognitive neuroscience** underpinned by brain imaging technologies such as electrophysiology (EEG) and functional imaging (fMRI)

For many decades, the dominant science for informing education practice has been cognitive psychology. Multiple publications directed at educators and a lay public aim to make accessible lessons for learning drawn from cognitive and education psychology (for recent examples see Weinstein, Sumeracki and Caviglioli, 2018; Kirschner and Hendrick, 2020; also see Deans for Impact, 2015; Pashler et al., 2007). These have proved highly influential for educators looking for a scientific basis to inform and improve their practice. More recently, neuroscience research has gathered a great deal of momentum, not least because of the advent of new imaging technologies which enable much finer-grained research; strong claims are made for the application of educational neuroscience (Goswami, 2006; Howard-Jones, 2014) or 'neuro-education' (Arwood and Meridith, 2017). Again, there is a body of publications focused on the implications of the science for classroom practice (e.g. Tibke, 2019; Jensen and McConchie, 2020).

Both areas of cognitive science are currently and increasingly informing interventions, practice and policy in education. Of particular interest to education has been basic cognitive psychology and cognitive neuroscience research in the areas of brain structure and function, motivation and reward, short- (i.e. working-) and long-term memory, and cognitive load. Cognitive load theory (CLT) (Sweller, 1994) merits consideration: it has been of such policy interest that it is now embedded in the school inspection framework as an official criterion that requires school practices to be informed by it and makes recommendations for practice. Also, across the current Teaching Standards for early career teachers there are clear expectations that newly qualified teachers will be well informed about cognitive science, in particular concerning cognition, memory and retrieval practice, and that they will develop skills to apply this knowledge in the classroom.

Other areas of cognitive science have potential implications for education, although are yet to enter widespread educational attention or practice. Not yet of interest to policy makers but with translational potential, for example, is research in cognitive neuroscience exploring synchrony of brain oscillations and memory formation (Clouter, Shapiro *et al.,* 2017). Brain oscillations refer to electrical activity generated by the brain in response to stimuli. There is growing interest in how these oscillations can synchronise between individuals during social interactions (i.e., brain-to-brain synchrony). The significance of brain-to-brain synchrony has recently been tested in education interactions, where it has been found in pair and group interactions, is shown to predict memory retention, and has been connected to forms of instruction, teacher-student relations, and learning outcomes (Bevilacqua *et al.* 2018; Davidesco *et al.* 2019; Dikker *et al.* 2017).

Learning concepts derived from cognitive science vary in the extent to which they enjoy consensus in the scientific community regarding the science and its educational implications. The basic research into cognitive load, for example, is well-established – drawing on behavioural psychology and a conceptual amalgamation of cognitive load and attention applied to educational principles of learning (for overviews, see de Jong, 2010; Sweller, 2016). It is yet, however, to be tested by the latest imaging neuroscience techniques. Furthermore, whether understanding of cognitive load in brain function and activity over very short (sub-second) units of time can be mobilised effectively to inform planning and delivery of teaching and learning that takes place over substantially larger units of time, and in a way that will have a measurable beneficial effect, remains to be demonstrated.

**Cognitive Science-Informed Practice (CSIP)**

Insights from cognitive science have the potential to both displace, complement, and add to, complex and varied understandings of effective classroom practice. Within schools, many techniques which are currently described as inspired by cognitive science may have been practised previously using a different rationale, including ones drawing on cognitive science. As discussed in the first project advisory group meeting (see Appendix 1 for a summary), techniques currently being described as inspired by cognitive science may have been practised previously by “hunch” (e.g. quizzes), rather than overtly inspired by cognitive science-informed practices. Many of these practices will be known to teachers as they resonate with established understandings of effective pedagogy. Cognitive science, as well as potentially identifying new or improved practices, can also provide a shared understanding and common language about existing techniques and furthermore help establish why some commonly used practices work, or not.

Like classroom practice in general, CSIP must be applied to specific subjects, phases, students and learning contexts. Another point emerging from the first advisory group meeting was that, for the review to realise its potential benefit it must, to some extent, connect cognitive science-informed teaching and learning strategies to contexts and subjects, and consider – where possible – the influence of these on the results. To some extent, mediating and moderating factors will be reported in classroom trial evidence and therefore amenable to analysis within a systematic review (see methods, below, for further detail). Scoping work (see 2b, below) however suggests that the classroom trial literature is yet to reach a point of maturity to enable impact evidence to be comprehensively assessed across subjects and contexts. Moreover, as CSIP is emergent within practice and in many areas still in a nascent state (e.g. subjects, phases, practitioner groups), going beyond trial-based evidence by collecting empirical evidence and reviewing practice-focused documents – as per our practice review sub-strand – is valuable to lay the groundwork for future practitioner-facing accounts of the CSIP evidence as well as better understand its relation to pre-existing practice.

**Evidence-Informed Practice (EIP)**

In relation to CSIP and EIP more generally, it is worth distinguishing evidence in terms of its a) robustness as a basis for causal inferences (i.e. internal validity) and b) its level of application and generalisability to specific classroom contexts (i.e. external and ecological validity). In other words, evidence varies in strength, is generated in a particular context and will need – to a greater or lesser extent – translation/adaptation for application in another. It is also worth distinguishing evidence from basic science and evidence from classroom trials as a basis for evidence-informed practice. As we have noted above, there is a varying level of evidential support across areas of cognitive science and ostensibly[1] the context in which the basic scientific evidence is produced is further removed from classroom teaching and learning and therefore requires a greater degree of translation. There remains much to understand with regard to the translation of cognitive science evidence to education and its application in the classroom, as well as with regard to the underpinning basic science itself (Fischer *et al.*, 2010; Aronsson and Lenz Taguchi, 2017; Rose and Rose, 2013).

Where cognitive science has been translated and applied within classroom interventions and techniques, it cannot safely be assumed that this will have the expected impact on pupil learning, however strong the evidence base in the basic science: Selected translational work provides a mixed picture. For instance, classroom-based manipulation of reward based on neuroscientific work on memory has shown null results in an EEF trial (Mason *et al.*, 2017) and neuroscience-based language learning interventions in the classroom which have elsewhere been effective (Goswami, 2015; Kyle *et al.,* 2013) have recently shown no significant benefit (Worth *et al*., 2018). These results do not necessarily (but may) bring the underpinning science and its evidence base into question. They do however highlight the value of seeking evidence with both high internal and ecological validity as a basis for practice, and distinguishing and evaluating the evidence base in these terms. Studies of classroom applications of cognitive science principles have yet to be systematically reviewed, as per the core focus of this review; as a result, we hold that at present practice in this area can only be said to be evidence-informed in a weak sense (i.e. of having some consonance with basic cognitive science). What is required as a basis for a stronger form of EIP is a systematic review of classroom interventions, appraised in relation to both internal and external/ecological validity.

The EEF is at the forefront of promoting evidence-informed practice in education in England and this review is part of their work summarizing the best available evidence for teaching and learning to support teachers and leaders to raise the attainment of 3-18-year-olds, particularly those facing disadvantage. We return to this point about types of evidence and the current state of evidence for cognitive science in the classroom in connection with the review problem statement, below, after briefly discussing relevant narrative and systematic reviews already in this area.

[1] One might argue that robust evidence which reveals something fundamental about how the brain learns will be highly applicable across contexts, and perhaps even have greater external validity than evidence from a particular implementation of a classroom intervention taking place in a given subject, school and learning context.

### *b. Existing narrative and systematic reviews in this area*

**Cognitive Science-Informed Teaching and Learning Strategies**

The previous section cites numerous practice-focused books and reports which provide narrative reviews relating to cognitive science in the classroom. These sources will be incorporated into this review through: mining of their underpinning studies during database searching, synthesising and summarising their implications for practice during the practice review strand, and synthesising and summarising their characterisation of the cognitive science literature as part of the underpinning science review strand. For context, it is worth noting that our scoping searches for published work regarding the educational applications of cognitive science (without being limited to specific strategies), yielded only 3 systematic reviews/meta-analyses published since 2000 (RCTs in education research: Connolly, Keenan, & Urbanska, 2018; teacher-led neuroscience-based RCTs: Churches et al., 2020; mathematics interventions: Kroeger, Douglas-Brown, & O'Brien, 2012). There are also several recent and influential non-systematic reviews, which influence the focus and design of the present review; these are summarised and their implications for this review discussed in this section.

Connolly et al. (2018) produced the first systematic review to evaluate all RCTs conducted in education from 1980-2016. While the study did not examine the evidence of effects for cognitive science-inspired strategies specifically, the yield of 1017 unique RCTS disputes the claims that it is not possible to undertake quality RCTs in education. However, as a rapidly growing field, there is now a sufficient threshold to conduct a systematic review on strategies that specifically focus on information acquisition and retrieval, which, as alluded to above, is an area of substantial interest to science, policy, and practice. Indeed, there have been recent reviews that specifically evaluate the effectiveness of *cognitive-informed* strategies, specifically. A recent meta-analysis conducted by Churches et al. (2020), analysed the findings of 34 teacher-led randomized controlled trials (RCTs), with the aim of bridging the gap between neuroscience and educational practice. In essence, teachers designed and implemented cognitive science-informed interventions, with each trial focussing on a single strategy (e.g. attention, retrieval practice, spaced practice, and interleaving). Overall, this seminal work demonstrated that cognitive-informed strategies can indeed benefit pupil outcomes.

In the initial searching and screening stages, our review has a somewhat broader scope: our approach involves rigorous probing of the wider literature, including classroom trials that are not exclusively designed by teachers, and are not necessarily RCTs (as this may not have been achieved yet with newer or more novel strategies). A key message from both cognitive scientists and practice and policy experts in our first advisory group meeting (Appendix 1) was to ensure that our treatment of cognitive science for knowledge acquisition and retention takes care not to be reductive; we therefore also aim to interrogate wider and related concepts that influence information acquisition and retrieval in addition to the core strategies (e.g. spaced practice) identified which mostly focus on how curricula are delivered and designed. This approach also reflects methodological recommendations that to meaningfully connect to practice, we need "more nuanced and sophisticated trials" that "are acutely aware of the contingent and context-specific nature of educational interventions" (Connolly et al., 2018, p. 14).

With respect to the utility of specific cognitive science-informed strategies, other reviews are also of note (Dunlosky et al., 2013; Weinstein et al., 2018). A 2013 monograph (Dunlosky et al., 2013) reviewed the literature to ascertain the relative utility of 10 learning techniques, with some consisting of the 'core' strategies outlined above (e.g. interleaving, spaced practice), but also including ones that students commonly self-report using (e.g. highlighting, rereading). Similarly, Weinstein et al.'s *Teaching the Science of Learning* (2018) highlights six important cognitive strategies: spaced practice, interleaving, retrieval practice, elaboration, concrete examples, and dual coding. Such reviews (also see scoping bibliography, Section 11) provide useful and accessible overviews of the most popular techniques; however these reviews are generally not systematic, and the scientific bases are not explored in depth, one of the objectives of the present review. Interestingly, Dunlosky et al. (2013) also considered the role of a number of mediators and moderators, such as state and trait-based student characteristics (e.g. age, intelligence), learning conditions (e.g. group vs. Individual learning), and task type. Similarly, our review will also code for moderators that may influence the effectiveness of cognitive-influenced interventions.

Several reviews with relevance to the application of cognitive science to the classroom have already been commissioned by the EEF, and it is important to distinguish how the present review will build upon, but remain distinct from, those reviews. Of particular note is the 2014 EEF review of educational interventions and approaches informed by neuroscience (Howard-Jones, 2014). This review provides a helpful overview of the neuroscientific approaches of relevance to the classroom and an appraisal of the evidence base for these as it stood in 2014. The present review does not aim to replicate this review, but to re-examine cognitive science-informed approaches and the studies underpinning them within the context of a systematic review, integrating the evidence from those studies with more recent ones and supporting the development of the EEF Education Database (see objectives below). A systematic approach, along with review sub-strands focused on the underpinning science and on practice, is needed to gain a more current, holistic and coherent picture of a rapidly growing field. The present review, encompasses the whole of the process from cognitive science (i.e. which strategies show potential, from perspective of basic science?), to impact assessment (i.e. which practices have been subjected to classroom trails and what does this evidence reveal about their effectiveness?), to practice (i.e. which strategies do teachers *actually* use?). Through its sub-strands our review also aims to connect the core evidence base to a wider range of strategies, including those that may be commonly used in educational contexts, but perhaps without an overt basis in cognitive science, whilst also identifying and examining advances in cognitive science that offer promising applications for education, but have yet to be translated into practice.

There have also been reviews concerning specific classroom strategies inspired by cognitive science. For example, reviews have examined how game-based learning relates to learning outcomes (Jabbar & Felicia, 2015), how spaced repetition promotes effective learning (Kang, 2016), and the role of instructional explanations in example-based learning (Wittwer & Renkl, 2010). However, even where systematic reviews exist for particular strategies, an overarching systematic review is needed to update the evidence and simultaneously capture, analyse, and evaluate all relevant teaching and learning strategies with respect to both their underpinning cognitive science, *and* their applicability to classroom contexts, by restricting studies to those with a classroom trial design. In summary, to our knowledge, there are no reviews which systematically review evidence of impact on pupils from classroom interventions, are recent, and have broad coverage of cognitive science-informed strategies for acquiring and retaining knowledge.

**Wider Cognitive Science-Informed Strategies**

The current review takes as its focus (see next section), cognitive science-informed teaching and learning strategies for acquiring and retaining knowledge. Cognitive science, however, is a broad and interdisciplinary field, which (as reflected in the 2014 EEF review) includes research on the physical, emotional and social conditions which support the processing, acquisition and use of knowledge. Understanding, for example, how emotions affect cognition and the implications of this for classroom practice is a question which falls within the purview of applied cognitive science more generally. These wider considerations fall at the boundary of this review. We recognise that a) clearly delineating cognitive science-informed strategies for acquiring and retaining knowledge from other cognitive science-informed classroom strategies is not always clear cut and, moreover, b) that taking into account as contextual factors the physical, emotional state of the brain and the social conditions in which cognition occurs – even if considered only insofar as these inform our assessment of strategies for acquiring and retaining knowledge – remains of value. How this boundary is navigated in terms of review methodology (e.g. inclusion and exclusion criteria, data extraction, analysis and the sub-review strands) is set out in detail below and the next section is devoted to defining and justifying the review focus in more detail. What remains to outline in this section are reviews which sit on this boundary and thereby inform and define the present review.

First are cognitive science topics relating to social and emotional aspects of cognition. The EEF has already published an evidence review on Social and Emotional Learning Strategies[2], aspects of which (e.g. mindfulness, stress reduction) are present in the cognitive science literature and relevant to the present review. School-based SEL strategies aim to help pupils acquire and effectively apply the knowledge, attitudes, and skills necessary to understand and manage emotions, set and achieve positive goals, establish and maintain positive relationships, and make responsible decisions[3]. Such trials demonstrate positive effects on wellbeing, behaviour, and the more social aspects of school life; some studies also suggest that emotional competence predicts academic achievement (e.g. for meta-analysis, see MacCann et al., 2020), indicating potential cognitive benefits. The relationship between the more emotional factors of cognition and test-based outcomes has yet to be systematically investigated from a cognitive science perspective. A second key area adjacent to this review is that of metacognition. Of particular note, a review on metacognition and self-regulated learning[4], with an accompanying Toolkit strand, explored the evidence base regarding teaching and learning strategies that relate to pupils' metacognition (in which pupils are encouraged to explicitly think about their own learning). To avoid the duplication of research efforts, the present review will not include metacognition as one of its strategies, using this EEF review to define its boundaries. Finally, we are aware that there is a concurrent EEF systematic review taking place focusing on feedback. Similarly, the present

---

[2] https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/social-and-emotional-learning/

[3] See http://casel.org

[4] https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/metacognition-and-self-regulation-review/

review will not cover feedback processes as cognitive-informed classroom strategies and liaise with the feedback review team and examine any possible duplication of effort with respect to the EEF Evidence Database if this is beneficial.

To return to the point made above, we hold that there is value in the present review probing these boundary areas to: a) avoid duplication of effort or content within the EEF's evidence and resource base, b) to identify contextual and moderating factors for cognitive science-informed teaching and learning strategies for acquiring and retaining knowledge, and c) identify boundary cases which meet the review inclusion criteria and thereby avoid an overly reductive account of strategies within the core focus. We discuss the navigation of this boundary at greater length in the next section.

## c. *Focus and Problem Statement*

**Problem Statement**

As described above (Section 1a), over the last decade and in particular over the last few years, there has been growing interest and practical application of findings from cognitive science to classroom practice. There are now a substantial number of recent and influential reviews and impacts on practice and policy. In large part, reviews identify and advance implications for classroom practice based on lab-based research and plausible interpretation of the basic science.

Several problems are apparent with the current state of applied cognitive science literature and practice:

- many techniques require considerable translation for application to the classroom;
- the research and practice of translation is emergent and variegated, with particularly rapid change in the last few years;
- it cannot be taken for granted that techniques with firm foundations in the basic science will be or are being successfully applied in effective interventions and practices in the classroom: initial evidence demonstrating successful application is mixed;
- the extent to which emerging practice has fidelity to the underlying cognitive science is unclear;
- education research offers incomplete understandings of the influences on and mechanisms of learning (Youdell & Lindley 2018); how cognitive science-informed practice relates to this research and existing practices, including other evidence-informed practices, is unclear.

As the basic science has been applied, there has been a growing literature reporting classroom trials of interventions informed by/based on cognitive science, but this literature has yet to have subjected to a systematic impact review, and what reviews there are dated, focused on specific strategies and/or have not used rigorous systematic review methods. Crucially, the distinction between classroom trials and other (e.g. lab-based) trials is often not made: as we note in the previous section, there is no current, comprehensive, systematic review of cognitive science-informed teaching and learning strategies trialled in the classroom. As a result, the current and potential impact on pupil outcomes of various cognitive science-informed interventions and techniques remains uncertain. This uncertainty and incompleteness contrasts with a widespread positive view about the potential benefits of applied cognitive science including education policy makers' keen attention to advances in cognitive psychology and neuroscience and efforts to use these to inform teacher education, guidance for classroom practice, and school inspection. This is therefore the ideal time for this systematic review of cognitive science in the classroom.

These gaps in understanding discussed above bring into focus the three strands of this review: Our core strand – and central focus for the review – concerns the evidence of impact in the classroom of cognitive science-informed interventions and techniques. The sub-strands examine, first, the underpinning cognitive science and, second, what cognitive science-informed practice and interventions 'look like' in the classroom and how they vary according to subject, context and pupil characteristics. In section 2d, below, we provide an overview of and background information pertaining to these strands and the links between them. Before doing so, we return to the question of the focus and boundaries of the review raised in the previous section.

**Review Focus and Scope**

The Systematic Review will investigate approaches to teaching and learning informed by cognitive science that are commonly used in the classroom, with a particular focus on acquiring and retaining knowledge. This focus reflects the areas of cognitive science which have to date been the most influential for classroom practice and ostensibly have the most general application across the education sector.

While there are numerous formulations and accounts of key techniques informed by cognitive science relating to acquiring and retaining knowledge, the following concepts are the most widely known and included in the EEF's initial review specification:

- Spaced practice
- Interleaving
- Retrieval practice
- Dual coding and
- (strategies to manage) cognitive load.

These concepts are briefly outlined in our Concept Map, along with examples of how these might be used in the classroom, as informed by our scoping work to date (Appendix 2). Our core focus also includes a wider and related set of strategies (see also Appendix 2). Indicative examples include the use of concrete examples, highlighting during study (Dunlosky et al., 2013), elaboration and interrogation techniques and techniques from 'brain training'. Inclusion of these strategies beyond those identified above will be based on the specified inclusion and exclusion criteria set out in this protocol, below, rather than through *a priori* specification of concepts for inclusion. We aim for a broad and comprehensive inclusion of cognitive science concepts focusing on acquiring and retaining knowledge. A clear and practicable categorisation of concepts and account of the relationships between them will be developed alongside coding work undertaken during the review as the evidence emerges. This is a form of framework synthesis, where an *a priori* initial conceptual framework is modified during systematic review to organise the evidence as well as an outcome of the review (Gough, Oliver and Thomas, 2017). A tentative, working concept map produced from early scoping work is provided in Appendix 2 and discussed further in Section 2e, below. The first section on this framework, 'Approaches Informing the Design of Teaching and Learning Techniques', captures our core focus on knowledge acquisition and retention strategies.

We have also identified several other areas relating to physical, emotional areas of cognitive science and areas involving direct measurement or manipulation of neural activity. As discussed above, these are at the boundary of our review. Below we describe the rationale for inclusion and exclusion which feeds into the review methodology in the later sections.

One clear message from the first advisory group (see Appendix 1) was that for the review to realise its potential benefits, there was a need to avoid a reductive view, not only of cognitive science, but also of the concept of learning and consequently, of links between the two. There are several issues connected to successfully navigating this boundary:

- First, focusing exclusively on the five core concepts (as above) risks an overly narrow focus on the concepts currently popularised in policy and practice in a particular location (England). In effect, this sets the bounds of the review around cognitive science strategies (and in particular those from cognitive psychology) which are included in previous influential (but typically not systematic) reviews, as opposed to a defined (conceptual and practical) set of inclusion and exclusion criteria.
- Second, a focus on 'cognitive science-informed strategies for acquiring and retaining knowledge' might conceivably include classroom strategies relating to social and emotional aspects of learning, or to physical or social factors supporting cognition. Previous research (e.g. Howard-Jones, 2014) has for example highlighted that maths anxiety leads to greater activity in the amygdala region of the brain and reduced working memory. If a simple classroom strategy was identified which organised a sequence of maths problems in a way which reduced anxiety, arguably this would promote the acquisition and retention of knowledge, be based on a strategy informed by cognitive science and fall within the scope. Similarly, there may be interventions which employ multiple strategies which cut across the broad groups identified. One example of this might be motivational or emotional strategies which employ quizzes (i.e. retrieval practice), or concrete examples to engage learners.
- Third, and related to the previous point, many concepts identified within the wider boundary concepts of the working concept map may be considered to be contextual, moderating and/or mediating factors for those in the core area. The boundaries between 'active ingredients' from cognitive science and the contextual or moderating factors may not be entirely clear. Potential overlaps and links with other concepts are highlighted within our concept map (see Appendix 2 and Section 2e); many of these overlaps are between concepts from the core focus area and the wider areas in the map.
- Fourth, wider concepts which meet the review criteria may not be clearly manifested in practice and amenable to systematic review of classroom trials. The field is yet to be systematically mapped and it is unclear at this point what is feasible within the resource envelope of the review. How tightly around the core focus of the review and what scope there is to include broader conceptions of cognitive science-informed strategies for acquiring and retaining knowledge into the systematic strand of the review will depend on the – as yet unknown – weight of evidence from the more narrowly-defined core focus areas.
- Finally, the definitions of cognitive science and the boundaries between related concepts such as meta-cognition vary between reviews. As discussed above, the EEF reviews on meta-cognition and self-regulation, on social and emotional learning, and the currently ongoing review on feedback, and the focus and scope of these, form important considerations for the boundaries of the present review. One

challenge in defining and managing boundary concepts will be to ensure duplication of effort with these reviews is avoided while highlighting important areas of overlap and linkage.

These issues inform the methodology for this review, as detailed below. Navigation of these boundary issues will involve clear application of the review specification set out in this protocol and in particular:

- the use and application of clear inclusion and exclusion criteria,
- the application of a clear and specified search strategy,
- the use of coding frameworks to identify points of contact between included studies and wider cognitive science and educational moderating and contextual factors, and
- the exploration of issues which fall outside of the core systematic review criteria but nonetheless have value in illustrating and interpreting the results within the review's sub-strands (e.g. variations in classroom practices which have not been trialled or linkages between wider cognitive science concepts and those associated with specific strategies for acquiring and retaining knowledge).

### *d. Overview of Review Strands*

There are still many questions as to how exactly cognitive science can be translated into classroom practice, taking into account the complexity of both cognitive science and educational settings, and the impact of doing so. The core review strand evaluates the impact of cognitive science which has already been translated and trialed in the classroom; we maintain, however, that making and examining links between this impact evidence and a) the underpinning science, and b) cognitive science-informed classroom practice is of value for making sense of the results and their implications, and for supporting further research and practice in applied cognitive science. This section justifies this point further and provides an overview of the review strands.

The design of this review is underpinned by an appreciation that in translating cutting-edge research for application in education we are looking for a golden thread that runs through a) the basic science, b) the plausibility of its applications, c) evidence of efficacy, d) evidence of effectiveness, and e) the evidence of differential effectiveness by context/student groups.

Recognising this, our systematic review will comprise a core strand and two further sub-strands:

- **Core Strand – 'Systematic Review'** – Systematic review of published literature reporting cognitive science-informed interventions inside classrooms. In addition to the impact evidence, this also includes a review of the 'translation' of the basic science during the analysis, looking across other strands (see below) and the evidence therein, making an assessment of the ecological validity of interventions reviewed in the core strand and their fidelity to a) the cited underpinning cognitive neuroscience research and b) the broader state-of-the-art in cognitive neuroscience (as per the 'underpinning research review').
- **Sub-strand 1 – 'Underpinning Cognitive Science Review'** – Review of literature in contemporary cognitive neuroscience demonstrating the underpinning research that is mobilised in interventions
- **Sub-strand 2 – 'Practice Review'** – Review of practice documents, and school data collection to identify and illustrate the applications of cognitive science in the classroom

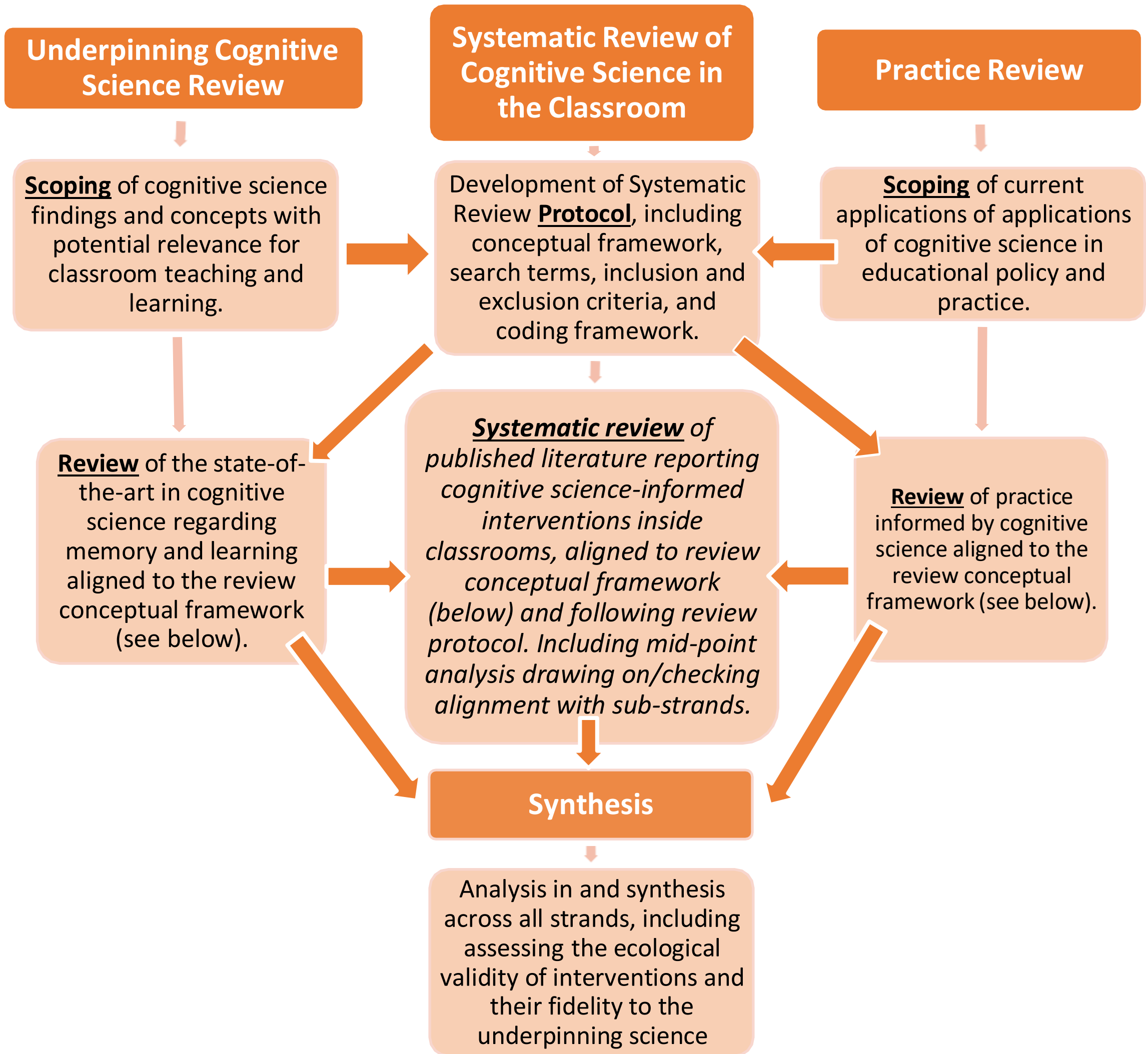
An overview of how these strands inform one another to produce robust, relevant findings and data is provided in Figure 1, below.

As depicted in Figure 1, there are three main points of contact between the review sub-strands and the core review. First, the **scoping work,** as part of the practice and underpinning cognitive science sub-strands, has already taken place to inform this protocol. Setting out the results of the scoping work as a **definitional and conceptual framework** (see next section) and providing details of searches and decisions around framing and focus, makes transparent the basis for the systematic review methodology provided below.

The scoping work has informed this protocol in the following areas:

- Identification and definition of core concepts
- Specification of Interventions / Practices / Phenomena of Interest
- Search terms
- Identification of pertinent modifiers associated with phenomena of interest (e.g. differential effects for educational phases or subjects, pupil subgroups, or implementation variants)
- Inclusion/Exclusion Criteria
- Coding Framework for Data Extraction Tool

**Figure 1 – Overview of Strands for the Cognitive Science in the Classroom Review**



The definitional and conceptual framework also lays the groundwork for continued work within the review sub-strands and ensures a degree of alignment of the cognitive science, educational principles and practices, and the appraisal of intervention evidence throughout the review. In particular, the definitional and conceptual framework will create an initial definition of core terms and the relationships between them and help to identify and map moderating factors which will be tested and refined as the review unfolds. It will also help ensure that the strands of the review form a coherent whole and allow the 'golden thread' from the basic science, through the translation to evidence of impact to be investigated and evaluated consistently and systematically.

The second point of contact between the sub-strands and the core systematic review is the **mid-point analysis**. The mid-point analysis will be an opportunity to examine the weight and quality of the evidence for each of the cognitive science areas and, in that context, pull-together and review side-by-side emerging findings from across the review strands. Assessing the weight and quality of evidence is done after full application of and coding for our core eligibility criteria (including quality appraisal criteria), as part of our interative inclusion stage, but before detailed extraction of further data required for the final analysis and synthesis. We detail our staged and iterative approach to applying eligibility criteria and study inclusion in the methodology section below. At the mid-point, a picture will emerge about the amount, focus and quality of evidence in the respective areas of cognitive science investigated before detailed extraction and analysis has been completed; our mid-point analysis allows us to draw on this emerging picture to identify studies by cognitive science area for more detailed data extraction, and inform and align the core review to the other review strands.

The final point of contact between the review strands is the data extraction, analysis and synthesis. In this phase we will analyse the findings from each of the review strands and use the results and the review questions to shape the synthesis, initially individually, then carrying out a cross-strand analysis to assess, for each cognitive science area, the extent to which the evidence supports, disproves and develops our understanding of the 'golden thread' of the successful translation of cognitive science into effective classroom theory and practice.

### *e. Definitional and Conceptual Frameworks*

***Translation, Implementation and Knowledge Transfer*** – Overarching concepts informing the design of this review include translation, implementation and knowledge transfer. We hold that there are rarely singular, unambiguous implications of findings from cognitive science for the classroom; rather, a process whereby the basic science has been interpreted and translated into education interventions and techniques, operationalized and implemented across subjects and learning contexts by teachers has necessarily taken place. Important considerations therefore when considering the implications of the results of the trial-based evidence include considering the translation of the cognitive science, the ecological validity of the intervention and the fidelity of implementation to the intervention design. Drawing implications and conclusions from the evidence-base produced within the core systematic review strand is likely to encounter the general scientific problem of underdetermination of theory by evidence. This review has therefore been designed so that consideration of translation and implementation within the core review and sub-strands can inform, for example, questions about whether interventions are likely to work in other settings or whether inconsistent or null findings are likely to stem from weaknesses in the underpinning science, its translation or implementation. While the core focus remains on what the evidence of impact is (see previous section), these wider considerations enable the results to be contextualised, support the process of drawing implications for research and practice from the core findings and informing further research and practice in applied cognitive science.

Table 1, below, summarises our theory of translation from cognitive science to classroom practice, to the production of an evidence base on impact. Below, we indicate the scope of the review strands in relation to this over-arching theory.

**Table 1 – Theory of Translation**

| ***Cognitive Science*** | | ***Changes in Classroom Practice*** | | | ***Evidence of Impact*** | |
|---|---|---|---|---|---|---|
| *Cognitive Science Evidence Base* | *Cognitive Science Theory* | *Classroom Teaching and Learning Theory/Theories* | *Classroom Teaching and Learning Practice(s)* | *Contextual and moderating factors* | *Impact evidence strength* | *Validity- including ecological* |
| Cognitive Science research, including both neuroscience and cognitive psychology, develops bodies of theory and evidence about cognition and learning. Some of this ostensibly has implications for classroom teaching and learning. | | Applied researchers and practitioners interpret, translate and operationalise findings from basic science into theories and/or practices of classroom teaching and learning. A concept from the basic science may generate more than one educational theory or practice. In many cases, particular contexts and/or modifying factors are identified that are thought to influence the effectiveness of a given practice. | | | Some applications of cognitive science to classroom teaching and learning have been trialled. The impact of these trials will depend on the validity of the underpinning science, its translation and its implementation. Trials will vary in both internal and ecological validity. | |
| | | ***Core Strand – Systematic Review*** | | | | |
| ***Underpinning Cognitive Science Review*** | | | | | | |
| | | ***Practice Review*** | | | | |

Finally, it is also valuable to note that CSIP can be supported by knowledge transfer in both directions (science-practice and practice-science): cognitive science techniques already part of established practice can be translated (or traced) back to the cognitive science, both in this review's underpinning science sub-strand and in future research; basic science may shed light on why and in which contexts some commonly used practices work and others do not.

**Conceptual Framework**

The strands of the review are held together by our conceptual framework which will be tested and refined throughout the systematic review process (Gough, Oliver and Thomas, 2017). It will a) specify our focus in terms of the theories from cognitive science (including related concepts or synonyms), b) provide definitions for and an overview of each theory, related practices, and variations in these, c) identify relationships and boundaries between them, and d) identify any known modifiers or contextual factors which are thought to influence the effectiveness or applicability of the theory in practice.

The conceptual map (see Appendix 2) presents the initial concepts of interest. These have been broadly divided into categories:

- Approaches Informing the Design of Classroom Teaching and Learning (Core): spaced practice, interleaving, dual coding, retrieval practice, cognitive load (also potentially concrete examples; brain training)
- Approaches that involve physical factors (Wider): exercise, nutrition/hydration, sleep
- Approaches that involve the motivational or emotional state of the learner (Wider): mindfulness, stress/anxiety reduction, social and emotional learning, reward/game-based learning
- Approaches that involve direct manipulation or measurement of neural activity (Wider): transcranial electrical stimulation, brain-to-brain synchrony

The scientific basis of these theories and techniques will be explored in the underpinning science review. It is important to note that this concept map is not exhaustive and will evolve through the iterative review process. For example, the underpinning cognitive science review may reveal important concepts/techniques that display potential application to the classroom that are not yet used. Similarly, the practice review may reveal classroom practices that could potentially have a basis in cognitive science. Some of the more contemporary concepts may not yield classroom trials and may be restricted to underpinning science review only. It is important to acknowledge that these concepts overlap. A preliminary outline of some of the important conceptual connections is indicated in conceptual map.

## 3. Objectives

### *a. Objectives: Systematic Review of Cognitive Science Interventions in the Classroom*

**Core Strand – 'Systematic Review'** – Systematic review of published literature reporting cognitive science-informed interventions inside classrooms

This core strand of work will respond to the EEF Research Questions:

1. What is the impact within the classroom on pupil outcomes of approaches rooted in or inspired by cognitive science and have strong evidential underpinnings from cognitive science regarding memory and learning?
2. What are the key features of classroom approaches based on cognitive science that have been successful in improving pupil outcomes, and teachers' and learners' contributions to them?
3. Do approaches rooted in or inspired by cognitive science have differential effects on outcomes for significant groups of pupils (for example, Key Stage 2 pupils, or pupils eligible for free school meals? Or, in certain subjects?). If so, what are the key features of successful approaches?
4. What does this review tell us about how the 5 core strategies (see above) relate to the underlying cognitive science research, and to each other?

We define cognitive science interventions through our conceptual framework, identifying the a) Classroom Teaching and Learning Theory/Theories and b) Classroom Teaching and Learning Practice(s) which are under review. These are all informed by or based on cognitive science. We recognise however that the nature and strength of the connection between cognitive science, theory and practice varies, something we examine in our sub-strands (see below). We define effectiveness as a difference in attainment between pupils receiving a cognitive science intervention and either a) a business as usual condition or b) a specified alternative condition that is causally attributable to the treatment conditions. However, it will also be important to check that no harms are being done (I.e. that the interventions do not lead to worse outcomes). The control condition will be identified in data extraction and accounted for in the analysis and reporting. The security of causal attribution will be judged in line with EEF evidence strength guidance[5] and our quality assessment criteria, specified below.

We will also assess the ecological validity of interventions reviewed in the core strand and their fidelity to a) the cited underpinning cognitive neuroscience research and b) the broader state-of-the-art in cognitive neuroscience (as per the 'underpinning research review'). Again, criteria for this are specified below. This part of the analysis brings together the emerging findings from the core review with the practice review and the underpinning science review to investigate the 'golden thread' from basic science to classroom practice. For each approach it will assess: 1) The veracity of the translation from the science to current classroom interventions and 2) The recency/quality/reliability of the underpinning research cited by interventions.

Finally, we note that the database produced as part of the core systematic review strand is designed in line with the coding frameworks and criteria of the EEF Evidence Database work, supporting the Teaching and Learning Toolkit. An additional objective for this review is therefore to support the database project and conduct this review in accordance with its methodology and aims.

### *b. Objectives: Underpinning Cognitive Science Review*

**Sub-strand 1 – 'Underpinning Cognitive Science Review'** – Review of literature in contemporary cognitive science (including cognitive psychology and cognitive neuroscience) demonstrating the underpinning research that is mobilised in interventions

This sub-strand offers significant value-added, adding to robustness of the review and deepening understanding in education of the underpinning cognitive science. This will ask:

1. What is the state-of-the-art in cognitive science regarding memory and learning?
2. What is the current state of evidence about mechanisms for memory and learning and the effects of these mechanisms on learners, and how confident is the field about this?
3. What are the links between the best evidence about cognitive science and about teaching and learning and what translation of this evidence for education, if any, has been tested or recommend by the field?
4. What does this review tell us about how the 5 core strategies relate to the basic science and to each other?

Examining the underpinning cognitive science and what educational implications have been recommended by the field allows this review to describe and assess how and the extent to which classroom interventions located in the systematic review stand are informed by specific areas of cognitive science. It also enables the identification of educational applications of cognitive science beyond the focus areas that a) can be included in the core systematic review and/or b) show promise for further research or translational work.

5 https://educationendowmentfoundation.org.uk/evidence-summaries/about-the-toolkits/evidence-strength/

### *c. Objectives: Practice Review*

**Sub-strand 2 – 'Practice Review'** – Review of policy and practice documents, and school data collection to identify and illustrate the applications of cognitive science in the classroom.

***During scoping*** – We will conduct scoping work to identify applications of cognitive science in the classroom from policy and practice documents (e.g. reports, frameworks, guidance, popular-scientific texts) to support protocol development guided by the question:

1. What applications of cognitive science in the classroom are currently prominent in policy and practice documents (e.g. reports, frameworks, guidance, popular-scientific texts)?
2. What form(s) do applications of cognitive science take when manifested in practice?
3. Do these support the development of protocols capable of assessing fidelity to the underpinning science and, if so, in what ways?

These questions will continue throughout the practice review, which will compile and review a body of policy- and practice-focused literature and documents. Details of search strategies and locations, bibliographies, and methods of analysis will be recorded in a search and methods record. This document and literature review will be developed through primary data collection:

***During data gathering, screening and extraction*** – we will survey practitioners in England to identify common practices within the parameters of the review. This will ensure that our searches and coding are focused on the approaches of most interest to teachers and provide data and illustration for final project reports guided by the question:

4. What do a stratified (for primary and secondary and classroom teachers and school leaders) sample of approximately 3000+ practitioners in England identify as common approaches that are based on cognitive science via a mainly quantitative survey?
5. What does this tell us about framing our extraction and analysis so that our results are of most interest to teachers and provide data and illustration for our final project report?

Our survey will be distributed via teacher and school organisations and social media. Questions will be developed as part of the practice review based on questions 9-11, above, and refined following mid-point analysis from the core systematic review.

***During analysis and synthesis*** – We will also identify centres of excellence and/or cognitive science themes for case study data collection (depending on what is possible because of Covid19). We will identify potential case studies from an item in our survey which asks whether respondents would consider participating in a follow up interview or visit (sampling purposively based on survey responses) and/or approaching known school contacts such as EEF research schools. We will collect data from at least 4 schools/colleges to observe and record relevant practice, targeting practice which will best illustrate review findings and approaches inspired by cognitive science in different contexts and phases.

1. How do practices in at least 4 schools/colleges captured through a mix of documentary analysis, observation and or remote recording best illustrate review findings and approaches inspired by cognitive science in different contexts and phases?

The multitude of existing and developing practices described as being inspired by cognitive science and the many questions pertaining to the translation of cognitive science into the classroom and learning outcomes for different groups of students form the background for the present review. A clear account of what constitutes CSIP and how it varies helps define the focus and scope of the core systematic review and ultimately illustrate its findings; moreover – like the examination of the underpinning science – this also enables the identification of educational applications of cognitive science and variations on them which are beyond the focus areas set out below that a) can be included in the core systematic review and/or b) show promise for further research or translational work.

## 4. Methodology – Core Systematic Review

### *a. Inclusion and exclusion criteria for the review*

Our approach to searching and screening is iterative, with two overall groups of eligibility criteria applied. The initial eligibility criteria for studies to include in the core systematic review are the following:

1. **Population:** Children and young people between 3 and 18 years of age in classroom settings (i.e. excluding University based studies with first year undergraduate students). This can include mainstream or special education settings, in any country. While schools are generally attended by children aged 5 upwards, a growing body of evidence suggests that cognitive science is useful for helping understand how play contributes to learning processes in very young children (e.g. see Schlesinger et al., 2020). Although the 'classroom' may not necessarily be in a school, learning can still take place in other Early Years settings (e.g. playgroup). Furthermore, recent published EEF reviews on other aspects of learning, such as metacognition and self-regulated learning (Muijs & Bokhove, 2020), and social and emotional learning (Wigelsworth et al., 2020), have also included studies with pre-school age children (I.e. 3-4 years) and adolescents up to age 18 (I.e. students in Further Education). Including a broad age range will ensure our review aligns with those other EEF reviews.
2. **Interventions/Practices of interest**:
    i. An evaluation of a classroom trials and/or interventions
    ii. Uses approaches derived from cognitive science relating to the acquisition and retention of knowledge. Studies need to demonstrate that the approach is derived from or inspired by cognitive science theory/principles. Scoping for each of the key areas outlined in the concept map (Appendix 2) will help outline the key terminology, theories and underpinning science for that specific area, which will then be used as a basis to judge whether a strategy is sufficiently based on cognitive science. For example, a classroom study that includes reference to 'germane load' or cites the work of Sweller, has clearly been informed by cognitive load theory. If a study is excluded as not overtly being derived from cognitive science, there may still be rationale for it to be included in the practice review to explore practices related to those that are.
3. **Study design and outcomes:**
    i. Initially we will include all studies reporting empirical evidence of any type or quality about pupil impact

ii. At the initial screening stage, we will include all studies which have *any form* of counterfactual (i.e. within- or between-group comparison of outcomes between conditions), even where serious threats to internal validity, such as selection bias, contamination or group inbalance are evident or probable.

Following an appraisal of the coverage and quality of evidence (see below) across the various cognitive science areas, we expect to tighten this criterion to include only experimental and quasi-experimental studies, which we will code using the EEF data extraction framework (v1.0, October 2019) (please see below for details of our code sets/extraction frameworks). In the former, we include randomised experiments (randomised at any level). In the latter we will include (in addition to the EEF framework) well-controlled observational/correlational designs such as instrumental variables regression approaches, prospective and retrospective propensity score matching, difference-in-difference methods and regression discontinuity designs (or similar, such as interrupted time series from natural experiments) (see Shadish, Cook and Campbell, 2002). We also expect to tighten this criterion to include only studies reporting test-based outcomes (measured in either grades and/or cognitive outcomes). We have included this initial 'low bar' for eligibility screening to cover the possibility that the amount and quality of evidence in a given area of cognitive science (as per our concept map) is limited, in which case we would retain studies providing only low internal validity evidence (which we will report as such). The approach to quality appraisal is set out in detail below.

iii. Systematic reviews and meta-analyses will also be included in initial searches from which empirical studies will be mined and assessed against these criteria. Results and conclusions from reviews will not be included in the core systematic review but will inform the review sub-strands and thereby be a point of comparison during review analysis and synthesis. We will flag (but exclude) all reviews as such in initial screening using an eligibility code.

iv. Studies that do not meet the eligibility criteria for the core review may still be included in either the practice review or the underpinning science review where they address the questions and objectives of these sub-strands, as specified above. These will be flagged (but excluded)

4. **Language:** written in English and peer reviewed (for journal articles).
5. **Date:** All
6. **Bodies of Literature, included in the review:**
   i. All peer reviewed journal articles. To limit search results, we will filter out non-peer reviewed journal articles at the search stage, where possible. Where databases do not include a filter for peer reviewed articles, we will refer to the journals themselves for information.
   ii. Reports based on research commissioned by policy makers, charitable or other non-commercial organisations.
   iii. Due to the time and cost constraints for the review, we will exclude conference proceedings, working papers and master's and doctoral dissertations/theses *that were published before January 2017* on the grounds that high-quality studies are likely to have been subsequently published.

It should be emphasised that the above describes the initial screen. We detail the iterative process through which further criteria are applied for inclusion and extraction in more detail below.

## *b. Search strategy for identification of studies*

**Search Databases**

For the review we will search the following databases, which include generic and specialist material:

**Table 2 – Search Databases**

| Databases Included from the *Web of Science Core Collection* | Databases Included from the *ProQuest Collection* |
|---|---|
| • Science Citation Index Expanded (SCI-EXPANDED)<br>• Social Sciences Citation Index (SSCI)<br>• Arts & Humanities Citation Index (A&HCI)<br>• Book Citation Index– Science (BKCI-S)<br>• Book Citation Index– Social Sciences & Humanities (BKCI-SSH)<br>• Emerging Sources Citation Index (ESCI) | • Ebook Central<br>• Education Database<br>• Psychology Database<br>• Social Science Database<br>• Sociology Database<br>• ERIC<br>• International Bibliography of the Social Sciences (IBSS)<br>• Applied Social Sciences Index & Abstracts (ASSIA)<br>• Sociological Abstracts |
| ***Science Direct*** | ***JSTOR*** |
| All publications | For titles in the following areas:<br>• Education<br>• General Science<br>• Psychology<br>• Science and technology studies |
| ***EBSCO*** | |
| Education databases | |

We will also search in the following databases:

- **Campbell Collaboration Library of Systematic reviews:**
  - https://onlinelibrary.wiley.com/journal/18911803
- **Cochrane Library of Systematic Reviews:**
  - https://www.cochranelibrary.com/search?cookiesEnabled
- **EPPI Centre library of reviews:**
  - https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=62
- **Open Science Framework:**
  - https://osf.io/registries?view_only=
- **Google Scholar**
  - To capture commissioned research reports, we will also search Google Scholar and review the first 1000 returns

Additional searches:

- Two high impact and relevant journals selected for handsearching: ***Trends in Neuroscience*** and ***Education and Educational Psychology Review***. The reference lists for included studies and relevant systematic reviews will also be manually searched.

We **do not** plan to search in the following databases as we feel they are of lower relevant to our focus and/or will have too much overlap with the above to be effective use of the review resources:

- PubMed
- FirstSearch

Following searches, we will also liaise with EEF colleagues to identify relevant studies already in the EEF Education Database. An application will be submitted to the EEF to access these studies. We will add these to the main results and retaining records for which data are already extracted when removing duplicates.

Where databases allow screening by category (e.g. Web of Science, JSTOR), this will be used to remove records which are not relevant at the initial search stage.

We will be using the EPPI-Reviewer software for the review (see below for further details).

**Search Terms**

Our search strings have been developed through preliminary database searches to assess search term sensitivity and precision[6]. We also have considered feedback from advisory group members (see Appendix 1) about where to prioritise and how to define cognitive science concepts. The search terms will be based on the interventions outlined in our conceptual map (Appendix 2).

For each concept, the string will contain terms related to a) methodology, b) education (outcomes and classroom specific), and c) terms and synonyms related to the specific cognitive science area (including a general cognitive science search). These search terms will be entered into each search database with the minimum of adaptation needed to use the search syntax and functionality and ensure comparability across databases.

**Table 3 – General Search Terms (all searches)**

| ***Search Term Group*** | ***Search String (Fragment)*** | ***Search Location[1]*** |
|---|---|---|
| Group 1 – Methodology | intervention OR trial OR evaluat* OR experiment* OR quasi-experiment* OR pilot OR test* | Title, abstract or key words |
| Group 2 – Education Outcomes | AND<br>learning OR attainment OR achievement OR "test scores" OR outcomes OR exam* OR impact OR effect OR performance | |
| Group 3 – Classroom setting | AND<br>classroom OR teach* OR school OR “further education” OR nursery OR "early years" OR kindergarten OR pre-primary OR lesson | |
| Group 4 – Focus Concept | AND, one of the general or concept-specific search term fragments in Table 4, below. | |

[1] Subject to search database functionality

[6] https://handbook-5-1.cochrane.org/chapter_6/6_4_4_sensitivity_versus_precision.htm

The general search terms above will be combined with one of the search strings related to cognitive science in general and specific cognitive science concepts, below.

**Table 4 – Cognitive Science Concept-specific Search Terms – Core Concepts**

| ***Cognitive Science Concept*** | ***Search String (Fragment – to be combined with the general search terms, above)*** | ***Search Location[1]*** |
|---|---|---|
| Cognitive Science General | cog* OR brain* OR neuro* OR "learning science" | Title, abstract or key words |
| Spaced practice | spac* OR distributed | |
| Interleaving | interleav* OR interweav* | |
| Retrieval practice | retriev* OR "testing effect" | |
| Dual coding | dual | |
| Strategies to manage cognitive load | "working memory" OR "short-term memory" OR (load AND (Cognitive OR intrinsic OR extraneous OR germane)) | |

[1] Subject to search database functionality

At this protocol stage, we have tested and specified search terms (as per Table 3 and 4 above). As described in the sections above, we also plan to investigate the wider cognitive science concepts identified within the concept map (Appendix 2). We have included a general cognitive science search string above, and hope that this is sufficient to capture wider cognitive science classroom interventions meeting our eligibility criteria. As we describe below, a degree of iteration may be needed and beneficial. Early results and mid-point findings may reveal further cognitive science concepts for which additional searches and search strings (as above) may be required. Any further development of search terms and full search records from all search databases will be recorded and appended to a revised version of this protocol on completion, including the basis for all decisions which tighten or broaden the scope of the review. Once all records located through searches are imported within the EPPI-Reviewer software, all subsequent review methods will be recorded using this specialist software (see below for further details).

The bases for the decision on concept selection and additional searches are set out in the next section on the approach to iteration, below.

**Selection of studies**

The initial eligibility criteria for the review are set out in Section 4a, above. The review will work iteratively though two rounds of activity applying these initial eligibility criteria as a first screen. Rounds of increasingly detailed and rigorous criteria are applied at each stage. We map these activity rounds to the eligibility criteria applied in Table 5, below, which we follow with a further detail of each round.

**Table 5 – Staged application of initial eligibility criteria.**

| | **Round 1** *Screen Titles and Abstracts* | **Round 2** *Screen full reports* |
|---|---|---|
| 1. **Population:** Children and young people between 3 and 18 years of age in classroom settings | ✓ | ✓ |
| 2. **Interventions/Practices of interest**: | | |
| i. Evaluation of a classroom trials and/or intervention | ✓ | ✓ |
| ii. Uses approaches derived from cognitive science relating to the acquisition and retention of knowledge | ( ✓ )[1] | ✓ |
| 3. **Study design and outcomes:** | | |
| i. Initially we will include all studies reporting empirical evidence of any type or quality about pupil impact, including reviews, from which we will 'mine' for underpinning studies | ( ✓ )[1] | ✓ |
| ii. Studies which have any form/quality of counterfactual. | ( ✓ )[1] | ✓[2] |
| iii. We will flag (but exclude) reviews and meta-analyses for reference mining and to inform the underpinning science or practice review strands | ✓ | ✓ |
| iv. We will flag (but exclude) pieces of relevance to the underpinning science or practice review strands | ✓ | ✓ |
| 4. **Language:** Include pieces written in English and peer reviewed (for journal articles). | ✓[3] | ✓[3] |
| 5. **Bodies of Literature:** | | |
| i. Include all peer reviewed journal articles, and reports based on research commissioned by policy makers, charitable or other non-commercial organisations | ✓[3] | ✓[3] |
| ii. Exclude conference proceedings, working papers and master's and doctoral dissertations/theses that were published before January 2017. | ✓[3] | ✓[3] |

[1] Assessing this item will be to some extent possible from title and abstract screening, with definite 'no's' being removed. We will assess after round 2 the false-negative rates of records marked for exclusion based only on titles and abstracts and screen on full papers to ensure accurate coding.

[2] As discussed above, a decision will be made about level of stringency for the study design and quality criteria following an initial literature mapping after round 2 (see below).

[3] These final criteria will mostly be applied during database searching, but remain as eligibility criteria during screening for any records for which initial information was missing or erroneous.

After initial calibration, training and quality assurance in the use of the eligibility criteria and screening approach within EPPI-Reviewer, the two rounds of screening will be implemented. This initial calibration will involve three researchers all screening around 30 records and comparing the results, repeating if necessary. After initial screening on the title and abstracts, records selected for further review will go through a second round based on the full text. At each stage, 20% of records will be double screened independently by a second researcher. The comparability of this screening will be reported as a measure of inter-rater reliability, with any discrepancies identified, described and resolved. In the case of disagreement between the two reviewers in the process of abstract screening, a third reviewer will be involved in the selection process. A flag/code will be used to mark studies which need to be reviewed by other team members to reach agreement and confidence in the coding. Studies which do not meet the core criteria yet have potential value for the sub-strands of the review (e.g. qualitative studies of cognitive science applications) will be retained in a separate folder, for potential use in contextualising the quantitative findings.

The screening and final selection process will be documented in a PRISMA chart produced within EPPI-Reviewer. As part of the searching process, the researchers may utilise tools that assist with the identification and extraction of records, but all records will be checked manually.

**Approach to Iteration**

The systematic review process is iterative in two respects:

- First, in the range of cognitive science concepts considered within the review. All included studies will meet the core eligibility criterion (2ii. approaches derived from cognitive science relating to the acquisition and retention of knowledge); however, how narrowly/broadly this is interpreted, and where and how to broaden it, will depend on emerging evidence.
- Second, in terms of how stringently the additional eligibility criteria (see below) are interpreted for study exclusion, particularly in relation to methodological quality appraisal criteria are in relation to specific cognitive science concepts. Where there is strong evidence in an area, the criterion will be tightened to include only the strongest and most ecologically valid, causal evidence. Where evidence is weak or limited, additional empirical pieces of lower quality will be retained to allow an account of the cognitive science concept and pave the way for future research.

For both of these, the central consideration is the potential for informing practice through gathering and presenting robust, and relevant evidence on cognitive science-informed interventions and practices.

***Bases for iterative inclusion* –** A decision on whether and which additional concepts will strengthen the systematic review and are feasible to include within its scope of the systematic review will be based on the following ***bases***:

a. The ***quantity and quality of evidence*** emerging from data gathering for the five core review concepts. This includes practical considerations, around the resource envelope available to the review, as well as quality considerations around, for example, coverage of related and complementary cognitive science concepts and their operationalisation emerging from the search database after the first round of searching and screening (see below).
b. ***Emerging findings from the underpinning science review*** examining specific cognitive science concepts related to the acquisition and retention of knowledge (see Section 3b for objectives), identifying specific cognitive science concepts, related concepts, terminology and their educational applications.

c. ***Emerging findings from the practice review*** (see Section 3c for objectives) identifying variants in classroom practices, interventions and the terminology surrounding them that derive for cognitive science.
d. ***Advice from the advisory group*** relating to priority areas for investigation.

***Criterion for iterative inclusion –*** A decision on whether and which additional concepts will strengthen the systematic review and are feasible to include within its scope of the systematic review will be based on the following ***criterion***:

1. **Quality** of study in relation to
    a. Internal validity for answering its own question
    b. **Relevance** to our questions
    c. Ecological validity

(See below for details of quality appraisal, including extraction and appraisal tools)

2. **Topic** boundaries
    a) Definition of cognitive science and its boundaries
    b) Application to learning and its boundaries – e.g. teaching and learning in formal learning contexts
    c) Meaningful connection with targeted interventions (and or other well-established interventions)
3. Amount of **detail** regarding factors that affect the potential of studies to inform guidance:
    a) Detail provided regarding teaching and learning processes
    b) Clarity of agency of intervention agent (teacher, machine, learning support assistant)
    c) Detail provided regarding relationship with particular subject or phase of the curriculum
    d) Detail provided regarding CPD provided to teachers
    e) Detail provided regarding links with relevant, broader school policies
4. Any known **conflict of interest**, relating to the independence of evaluation and its design and methods (including outcome measures).
5. Number of studies that can be accommodated within the **resource** envelope

These additional criteria will be applied following the initial general screening (above) on the remaining records. Several items (e.g. relating to internal and ecological validity) have coding items already built into the EEF education database extraction tool (see below). Where standard coding items are not included, additional items will be added. For purposes of screening, these criteria will be coded as closed-response ordinal/binary ratings (e.g. low, medium, high level of detail provided), increasing efficiency (for records ultimately excluded) and enabling more transparent reporting via a PRISMA diagram.

Following this additional coding, the weight of evidence will be assessed using these criteria in a mid-point review ***by cognitive science concept/area***. At this stage selected studies based on the application of the additional eligibility criteria will be progressed to a final round of data extraction, with all decisions fully recorded in EPPI-Reviewer. This will allow for additional more detailed (open-response) and broader set of codes used for extraction than used for screening, while still based on the same criteria.

**Duplicates:** We will remove all duplicate records. We will in the first instance include multiple publications from the same study or body of work but will subsequently remove any superseded by other related publications associated with the study. Where multiple studies are reported within a single publication, we will apply eligibility criteria to publication sections or chapters pertaining to individual studies and treat eligible sections as single records.

### *c. Data extraction and management*

Data will be extracted from the selected papers, using a coding framework based on several parts:

1) The EEF main data extraction tool
2) The EEF effect size data extraction tool
3) Our quality assessment tools used in iterative inclusion (above), comprising:
    - The Revised Cochrane Risk of Bias Tool (2)[7]
    - The Cochrane GRADE Tool
    - Items for ecological validity from the EEF extraction tools
    - Additional quality appraisal items for **relevance**, **topic**, **detail** and any **conflict of interest** (as above). NB. There are conflict of interest items in the EEF extraction codesets (e.g. for developer-led evaluations) which we will use.
4) Codes produced from our underpinning cognitive science review and practice review as well as input from the advisory group, to flag and categorise studies, and extract data relating to cognitive sciences concepts evident within the interventions (see Appendix 3g for a first draft of these, to be refined using the data).

Our data extraction tools (including EEF, risk of bias, and review-specific tools) is provided in Appendix 3. We provide further details of quality appraisal below.

As with the earlier screening, initial calibration will take place with three researchers coding around 30 records and comparing data. During the process of data extraction, queries will be flagged on the EPPI-reviewer system, and there will be close coordination of the team to ensure quality control. All members of the team working on data extraction tasks will keep detailed records (wherever possible in EPPI-reviewer) and confer with each other should any problems arise.

At the data extraction stage, 20% of records will be double-coded independently by a second researcher. This applies to all extraction items (such as those relating to quality and effects, described below). The comparability of this coding/extraction will be reported as a measure of inter-rater reliability, with any discrepancies identified, described and resolved. In the case of disagreement between the two reviewers, a third reviewer will be involved in the selection process.

Key details of the entire screening and subsequent data extraction process will be presented in tables and a PRISMA diagram produced in EPPI-Reviewer. These reports will contain, for examples, key information relating

[7] https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials

to the search (e.g. how many studies were included/excluded for each search, origin of studies (by continent and context), quality ratings per category, and so on).

**Appraisal of included studies**
We will quality appraise studies using the criteria set out in the additional criteria, above, and repeated for convenience below:

***Quality of study*** in relation to:

a) Internal validity for answering its own question
b) Relevance to our questions
c) Ecological validity

These three criteria are operationalised in our coding/extraction framework using code-sets in four areas (again, given above, but repeated here for convenience). Our quality assessment tool used in iterative inclusion (above), comprising:

***Quality appraisal tools:***

- The Revised Cochrane Risk of Bias Tool (2)[8]
- The Cochrane GRADE Tool[9]
- Items for ecological validity from the EEF extraction tools
- Additional quality appraisal items for **relevance**, **topic**, **detail** and **conflict of interest** (as above)

In overview, we will use 1) the Cochrane Risk of Bias (RoB2) tool to assess internal validity at individual study level; 2) the GRADE assessment criteria to quality assess at the level of cognitive science concept areas[10]; 3) selected items from the EEF tools along with intervention frequency and duration codes to assess ecological validity[11]; and additional items to code for relevance, topic boundaries and detail.

We provide brief further details of each of these below, and the full code sets in Appendix 3.

### *Revised Cochrane risk-of-bias tool for randomized trials (RoB 2) (Appendix 3c)*

Several quality assessment tools were considered for the purpose of assessing risk of bias. These were narrowed done to the Quality Appraisal Checklist for quantitative intervention studies (NICE), the Cochrane Risk of Bias tool (Rob 2), and the Quality Assessment Tool for Quantitative Studies (EPHPP). While all three of these tools provide the means of assessing study quality across several core domains (I.e., selection bias, allocation to groups, outcome measures, reporting bias), a close inspection of available tools indicated that the **RoB 2: A revised Cochrane risk-of-bias tool for randomized trials** was the most applicable and appropriate for the purposes of our systematic review.

[8] https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials
[9] https://training.cochrane.org/grade-approach
[10] To support the coding for the RoB2 and GRADE tools, we will draw on items from the EEF data extraction (and effect size extraction) frameworks relating to study quality (i.e. participant group assignment process and level, design strength for causal inference, sample size, attrition/drop-out, group comparability, outcome measure quality).
[11] We have been advised on suitable items by Prof. Steve Higgins

RoB 2 is structured into a fixed set of domains of bias, focussing on different aspects of trial design, conduct, and reporting. Within each domain, a series of questions ('signalling questions') aim to elicit information about features of the trial that are relevant to risk of bias. A proposed judgement about the risk of bias arising from each domain is generated by an algorithm, based on answers to the signalling questions. Judgement can be 'Low' or 'High' risk of bias, or can express 'Some concerns'. The RoB 2 will be applied to each of our included studies using EPPI-Reviewer.

***Cochrane GRADE Tool (Appendix 3d)***

To summarise the overall strength of evidence in each cognitive science area for all included studies, we will use the GRADE criteria. We will report the results by GRADE factor in a summary of results table by cognitive science area. We will use our coded data (as above) and follow the GRADE handbook guidance[12] to assess risk of bias, indirectness, inconsistency, imprecision, publication bias and dose effect. This assessment tool has been developed in a medical setting, and even in that setting is under development. However, we believe there are analogous considerations in the educational trial literature which allow its application. For example, we are interpreting indirectness in terms of translation and ecological validity and dose effects in terms of frequency and duration of interventions, and are coding for these within our tools (RoB, Ecological validity, and additional quality items). We will use a best-fit judgement and the 4-point GRADE scale when upgrading and downgrading.

***Ecological Validity Assessment (Appendix 3e)***

Given the discussion of ecological validity and translation in the opening sections, as well as the advice from the advisory group around maximising benefits through clear application to particular contexts and settings, we plan to use several EEF extraction and effect size items and create several additional codes which will be used to assess ecological validity (again, reported in an overview table by cognitive science concept area). When assessing external/ecological validity, we will assess whether research results represent what happens in typical classroom teaching and learning in the population of interest (in this case school pupils aged between 3 and 18). Key variables for the assessment relate to how realistic the study was (in relation to typical classroom teaching and learning), the number and type of schools involved, who was responsible for delivering the intervention, and the duration and frequency of intervention sessions (see Appendix 3c).

***Additional quality appraisal items for relevance, topic and detail (Appendix 3f)***

Additional data extraction codes will be used to assess to assess **relevance** to our questions, relevance in terms of **topic** boundaries, the quality of reporting in terms of **detail**, and any **conflict of interests** (see above). These are simple closed-response codes which a) allow us to make and transparently record quality appraisal decisions and b) flag records for more detailed data extraction in open-response items to support analysis, synthesis and reporting.

### *Effect size calculation*

We will record all effect sizes using the EPPI-Reviewer coding items. As discussed below, we do not expect to have a sufficient number of homogeneous studies to conduct a meta-analysis of the results. We will however

[12] https://gdt.gradepro.org/app/handbook/handbook.html#h.fzuoa9x107cu

extract and report effect sizes for all studies where this is applicable, reporting these using the author's original preferred calculation within the narrative review. Where an effect size is not reported within the original study, but it can be calculated using the data presented, we will use Hedge's g, as recommended by the EEF (see Borenstein et al., 2009, p.27 for the formula); with the numerator being the difference in means between the two respective groups, and the denominator being the pooled standard deviation. With a sufficient number of effect sizes (for sufficiently homogenous interventions), we will use the $I^2$ statistic and estimate $\tau^2$ to consider heterogeneity of the sample and sub-samples of studies (Borenstein et al., 2009, pp.114 and 117).

**Unit of analysis issues**

Again, while a meta-analysis is not planned, we will record the level of randomisation for randomised trials during the data extraction process using the standard EEF data extraction codes. We will also record group sizes and units of analysis when extracting data. We will calculate weighted mean effects using a fixed effects model and provide an overview using a forest plot. We will use the meta-analysis tools within EPPI-Reviewer (and its integration with R and the Metafor package) to calculate the effect sizes and statistics.

**Dealing with missing data**

Wherever possible, missing values will be calculated from the paper. This can be achieved in instances where effect sizes are not reported, but group scores, sizes and standard deviation statistics are. If that is not possible, the authors of the papers will be contacted by email and asked to supply the missing data if it is deemed of potential importance to the review findings.

### *d. Data synthesis*

At the highest level, we will report summary findings for each cognitive science concept area using the GRADE assessment tool, as described above, GRADE ratings will be provided in a summary table which also provides overview information about the concept and the studies with the area and our key findings for the overall concept/intervention area.

Within results sections for each cognitive science concept area, results will be reported by a combination of a study-level summary of results table followed by a structured narrative synthesis approach. The narrative synthesis approach will be structured into groups of studies based on our quality appraisal criteria, as above. Summary judgements from our quality appraisal (i.e. on ecological validity, risk of bias, question and topic relevance, and detail) will be provided in the overview table. The summary table will be based on selected, closed-response fields from the EPPI-Reviewer database in the focus area/report section.

We will follow the summary table in each section by summarising in narrative form, the overall (as per the GRADE rating) and study-level weight and quality of evidence in the area. This makes transparent and communicates to the reader the confidence vested in particular studies and groups of studies as a basis for inferences within the narrative synthesis. We will group studies based on their quality appraisal and proceed to report the narrative synthesis in relation to quality level/groups. For example, we might begin one section by reporting that that there were a group of studies with high internal validity (perhaps based on well-conducted, randomised controlled trials) but low ecological validity; we will indicate which these are in the summary table and then provide a narrative synthesis of the findings of these; then – in this example – perhaps we have a group of medium validity studies but with higher ecological validity and greater detail in intervention process and implementation reporting; we would then (again with transparency about which these studies are and the overview table for readers to consult) discuss how these studies support, refine or refute those of the first group.

Synthesis and reporting will therefore be centred on clear organisation and presentation of data to address the research questions in each strand (as described in this protocol), with clear indication of quality characteristics of the studies within the synthesis throughout. The exact approach to grouping and reporting the synthesis will be decided when all selected papers have been identified for review and in collaboration with the project advisory group. Possible options in terms of structure is to organise the findings around the type of intervention/practice, around different subjects or age groups. If, for example, our advisory group strongly advises reporting of studies in the primary age context separately from the secondary age ranges we would organise extracted data, sythesise and present results by age range. Crucially, we would retain our structured narrative synthesis approach, as described above, of first describing the weight and quality of evidence in the section for synthesis before analysis of studies grouped by quality appraisal results.

At this stage we are not deciding on the minimum numbers of studies required for the analysis, as this will depend on the amount of studies identified in total and the specific approach for organising the data. We are also not planning a meta-analysis of the quantitative data, as the studies are likely to be too heterogeneous and include too many contextual factors. However, we will record the key details required to enable researchers to potentially conduct a meta-analysis in the future, something likely to be of value once the body of evidence in this area has grown. Any effect sizes (as well as any other main study findings and results) will be reported in overview tables of study details and outcomes.

**Investigation of heterogeneity**

The present review will ensure that all potentially relevant study characteristics are captured during the data extraction process and considered and reported within the narrative synthesis. To this end, we have developed a carefully constructed data extraction tool, based on and to supplement items on the tool provided by the EEF with added elements derived from our scoping work and advisory group meeting.

Objective 3 of the review relates to differential effects on any groups of pupils, and the data extraction tool thus includes information about the educational stage, age, gender and ethnicity of the children, as well as the proportion eligible for Free School meals, and the proportion who have SEND. These factors others within the EEF extraction template will be coded during the data extraction process and used to evaluate whether there are key areas of difference. Based on the discussion in the first advisory group, we will also record information about the specific subject that the reported intervention was directed towards.

It is possible that additional moderators will be identified throughout the review process (e.g through the underpinning cognitive science review, the practice review or when reading the full texts of included papers). To ensure that the data extraction tool includes all relevant moderators, we will at an early stage of the study, review it and revise if necessary. The educators and cognitive neuroscientists present in the 1st Advisory Group meeting (Appendix 1) emphasised the importance of age, with concerns that some cognitive science-informed approaches may show greater effectiveness for older children than younger children. The age of participants in each study will be recorded and potentially used as a means of performing subgroup comparisons. We will record the basis for all sub-group/moderator analysis prior to conducting analysis and report all (including null) results of these.

The data extraction tool which specifies the factors that will be assessed is attached as an appendix.

*Sensitivity analysis*

A meta-analysis is not planned as part of this review. As noted above, relevant data will be captured so that this could be performed at a later date.

## 5. Reporting

Findings from the review will be presented in a final report, using the EEF review reporting template for evidence reviews. Within this framework, it will be possible to some extent to make decisions around organising findings around particular interventions, subjects or age groups. The most logical and useful way to present the data within the template will be decided when the publications have all been identified and we have a better overview of 1) how many trials or interventions have been carried out using particular cognitive science theories (and also the number of participants within those studies), 2) how many studies are conducted on each type of subject, 3) age spread of study, 4) the extent to which studies consider other contextual factors, e.g. FSM, ethnicity and special educational need. The overall judgement on the quality or certainty of the evidence base for a finding will be based on a combination of factors including: the volume of studies and their findings (do they come to the same conclusion or do they differ significantly in their outcomes), the amount of children who have been trialled and outcome-tested, and the extent to which studies include moderating, contextual factors.

We have described our approach to synthesis and reporting, in which overview tables and narrative synthesis presents results from study transparently differentiated by our quality appraisal criteria.

The research team will work with the EEF to produce a school-facing documents using data and findings from the review, organised to best inform practice.

## 6. Personnel

*Overview*

The review team brings together expertise in cognitive neuroscience, education policy and practice, education evaluation and measurement, systematic review, and interventions in the classroom. We draw together experts from the University of Birmingham, the Centre for the Use of Research and Evidence in Education (CUREE), and practitioner researchers from or associated with the Queen Anne's School and its neuroscience-focussed research centre. Together this team offers a unique combination of expertise – in undertaking systematic review of interventions in school as well as in the underpinning science; the translation of new approaches in cognitive neuroscience to education; in assessing the ecological validity of interventions; and in the application of interventions in schools.

*University of Birmingham*

- ***Professor Deborah Youdell (DY)*** is an expert in the links between education policy and practice and inequalities. She is renowned for her development of interdisciplinary approaches in education drawing on new biosciences and neuroscience. She is currently Working Group Co-Chair of the UNESCO-funded International Science and Evidence based Education Assessment (ISEEA), bringing global experts together to map the state-of-the art for science-informed education.
- ***Professor Kimron Shapiro (KS)*** is an expert in cognitive neuroscience; in particular inattention, short- and long- term memory and their enhancement, as well as using modern approaches including electrophysiology and functional imaging to understand the brain mechanisms that underpin behaviour. Professor Shapiro has published over 100 papers, many in upper tier journals, and has an H-index of 42.
- ***Dr Thomas Perry (TP)*** will lead the project team. He has conducted numerous reviews including rapid, scoping, systematic, and policy reviews (Perry et al., 2018; Cordingley et al., 2018; Morris and Perry, 2017) and is an expert in educational evaluation and improvement, research review and synthesis, and secondary data analysis. Tom leads the educational leadership Master's programmes at the University of Birmingham and is currently working on one systematic review and a systematic realist review
- ***Dr Rosanna Lea (RL)*** is a Research Fellow in the School of Education at the University of Birmingham. Rose is also an Associate Lecturer in School of Psychology at the University of Worcester, where she teaches across the British Psychological Society core curriculum. Rose's research interests include emotion and cognitive science, adolescent mental health and well-being, individual differences, and autism.
- ***Dr. Clara Jørgensen (CJ)*** is a Research Fellow in University of Birmingham School of Education. Clara is a social anthropologist with experience of international and interdisciplinary research in education, extensive experience of conducting reviews and expertise in involving stakeholders and communities in research and teaching.

*Centre for the Use of Research and Evidence in Education (CUREE)*

- ***Philippa Cordingley (PAC)*** is an expert in systematic review, education evaluation and education policy. She chaired the EPPI CPD Review group and has been Principal Investigator for three EPPI and numerous other full technical and less technical reviews.
- ***Paul Crisp (PEC)*** is an expert in the technical aspects of knowledge management, design, analysis and interpretation, quantitative analysis and reporting.

*Queen Anne's School and BrainCanDo*

- ***Julia Harrington (JH), Amy Fancourt (AF) and colleagues at Queen Anne's School*** are practitioner researchers with an expertise in and experience of using neuro-science research in practice in both state and independent schools. ***BrainCanDo*** is an education neuroscience and cognitive psychology research centre based at Queen Anne's. Julie and colleagues will support the design, analysis and reporting phases of the review, in particular in relation to review focus and usability. They will be part of the advisory group and contribute to the case studies. Julia is Headmistress of Queen Anne's School and CEO of the BrainCanDo centre (https://braincando.com/) and Amy is Head of Psychology at QAS and Director of Research at BrainCanDo.

## 7. Conflicts of interest

All members of the research team, as listed above, are committed to educational practice being informed, where appropriate, by cognitive science, and the development, configuration and application of cognitive science to that end. We are also committed to a disinterested, systematic and transparent review processes to ensure that our assessment of the evidence base is warranted on methodological grounds and open to scrutiny. We have no conflicts of interest in our ability or intention to carry out these commitments. Nonetheless, several members of the team are working on related projects, summarised below:

- Youdell is currently Co-Chair of UNESCO International Science and Evidence in Education Assessment
- Shapiro is currently PI on a research programme to test the application of transcranial stimulation (tCS) to improve working memory. He is currently seeking a patent and in discussion with various companies about commercialisation.
- Harrington is the CEO and Fancourt Director of Research for BrainCanDo, a charitable company whose aims include: the development of a strong neuroscientific evidence base to inform and underpin education. BrainCanDo aims to empower teaching professionals to use the latest findings from neuroscience research to transform and enrich their classrooms and to empower students to understand how learning happens. One of the programmes that they will be trialling in 2020 is a 'Neuroscience for Teachers' course that has been developed by Prof. Patricia Riddell. The programme is expected to launch in January 2021. BrainCanDo are also publishing the following book in July 2020: 'The BrainCanDo Handbook of Teaching and Learning: Practical strategies to bring psychology and neuroscience into the classroom'.

The team has been organised with a core team (TP, RL, CJ) who will be applying the study eligibility criteria and quality appraisal coding. The wider team will be involved in analysis and narrative synthesis, but (as we have set out) this will be firmly and transparently grounded in the quality appraisal results.

## 8. Registration

This systematic review will be registered on the Open Science Framework registry following this protocol being finalised after peer review.

We plan to publish one or more papers in peer-review journals based on the review and its sub-strands. A publication outlet to target has not been determined at this stage.

## 9. Timeline

| Activity | M | J | J | A | S | O | N | D | J | F | Lead |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Phase 1 - Set-up and project management* | | | | | | | | | | | |
| Set-up and Project Management (ongoing) | | | | | | | | | | | TP |
| Scoping work (inc. research and practice) | | | | | | | | | | | TP/all |
| Create and agree review protocols | | | | | | | | | | | TP |
| *Phase 2 - Searching* | | | | | | | | | | | |
| Database searching, duplicate removal | | | | | | | | | | | TP/RFs |
| Practioner survey and initial analysis | | | | | | | | | | | DY/AF |
| Searches for Underpinning Evidence Review | | | | | | | | | | | KS |
| Searches for Practice Review | | | | | | | | | | | DY/AF |
| Abstract/title screening using inc./excl. criteria | | | | | | | | | | | TP/RFs |
| 20% Double-blind Screening | | | | | | | | | | | TP/RFs |
| *Phase 3 - Extraction* | | | | | | | | | | | |
| Screening on full text | | | | | | | | | | | TP/RFs |
| Initial quality appraisal, finalise mapping and extraction strategy | | | | | | | | | | | TP/RFs |
| Coding and Data extraction in EPPI Reviewer | | | | | | | | | | | TP/RFs |
| Complete underpinning evidence review | | | | | | | | | | | KS |
| Complete practice review, with survey evidence | | | | | | | | | | | DY/AF |
| 20% Double Coding and mid-point analysis | | | | | | | | | | | TP/RFs |
| Data verification | | | | | | | | | | | TP/RFs |
| *Phase 4 - Analysis and Synthesis* | | | | | | | | | | | |
| Mapping and quantitative summary | | | | | | | | | | | TP/RFs |
| Additional data extraction for pertinent and high-quality studies | | | | | | | | | | | TP/RFs |
| Interrogation of evidence and consultation including translation review | | | | | | | | | | | TP/all |
| School case study visits inc. prep | | | | | | | | | | | DY/AF |
| *Phase 5 - Write up and project completion* | | | | | | | | | | | |
| Drafting and finalisation of final review report | | | | | | | | | | | TP/all |
| Drafting and finalisation of school-facing publication | | | | | | | | | | | TP/all |
| Archiving and project completion | | | | | | | | | | | TP |

## 10.References

Aronsson, L. and Lenz Taguchi, H. (2017) Mapping a collaborative cartography of the encounters between the neurosciences and early childhood education practices, Discourse Studies in the Cultural Politics of Education 39(2):1-16 DOI:10.1080/01596306.2017.1396732

Arwood, E. L., & Meridith, C. (2017). *Neuro-Education: a translation from theory to practice*. Tigard, Oregon: Arwood Neuro-Viconics.

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., . . . Dikker, S. (2018). Brain-to-Brain Synchrony and Learning Outcomes Vary by Student–Teacher Dynamics: Evidence from a Real-world Classroom Electroencephalography Study. *Journal of Cognitive Neuroscience, 31*(3), 401-411. doi:10.1162/jocn_a_01274

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis.* Wiley.

Clouter, A., Shapiro, K.L., Hanslmayr, S. (2017). Theta Phase Synchronization Is the Glue that Binds Human Associative Memory. *Current Biology*, 27, 1–6.

Cordingley, P., Greany, T., Crisp, B., Seleznyov, S., Bradbury, M., & Perry, T. (2018). Developing Great Subject Teaching: Rapid Evidence Review of Subject-Specific Continuing Professional Development in the UK. Wellcome Trust. Retrieved from http://www.curee.co.uk/node/5032

Davidesco, I., Laurent, E., Valk, H., West, T., Dikker, S., Milne, C., & Poeppel, D. (2019). Brain-to-brain synchrony predicts long-term memory retention more accurately than individual brain measures. *bioRxiv*, 644047. doi:10.1101/644047

Deans for Impact. (2015). *The Science of Learning.* Deans for Impact. Available: http://www.deansforimpact.org/wp-content/uploads/2016/12/The_Science_of_Learning.pdf

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science, 38*, 105-134. https://doi.org/10.1007/s11251-009-9110-0

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., . . . Poeppel, D. (2017). Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Current Biology, 27*(9), 1375-1380. doi:10.1016/j.cub.2017.04.002

Fischer, K. W., Goswami, U., & Geake, J. (2010). The Future of Educational Neuroscience. *Mind, Brain, and Education, 4*(2), 68-80.

Goswami, U. (2006). Neuroscience and Education: From Research to Practice? *National Review of Neuroscience, 7*(5), 406-411

Goswami, U. (2015). The neural basis of dyslexia may originate in primary auditory cortex. *Brain, 137*(12), 3100-3102. doi:10.1093/brain/awu296.

Gough, D., Oliver, S., & Thomas, J. (Eds.). (2017). An introduction to systematic reviews. Sage.

Howard-Jones, P. A. (2014). Neuroscience and education: myths and messages. *Nature Reviews Neuroscience, 15*, 817-824. doi:10.1038/nrn3817

Jensen, E., & McConchie, L. (2020). Brain-Based Learning: Teaching the Way Students Really Learn. Corwin.

Kirschner, P. A., & Hendrick, C. (2020). How Learning Happens: Seminal Works in Educational Psychology and what They Mean in Practice. Routledge.

Kroeger, L.A., Douglas Brown, R., & O'Brien, B.A. (2012).Connecting Neuroscience, Cognitive, and Educational Theories and Research to Practice: A Review of Mathematics Intervention Programs. Early Education and Development, 23(1), 37-58.

Kyle, F., Kujala, J. V., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly, 48*(1), 61-76.

MacCann, C., Jiang, Y., Brown, L., Double, K., Bucich, M., & Minbashian, A. (2020). Emotional intelligence predicts academic performance: A meta-analysis. *Psychological Bulletin, 146*(2), 150-186. https://doi.org/10.1037/bul0000219

Mason, A., Farrell, S., Howard-Jones, P., & Ludwig, C. J. H. (2017). The role of reward and reward uncertainty in episodic memory. *Journal of Memory and Language, 96*, 62-77. https://doi.org/10.1016/j.jml.2017.05.003

Morris, R. & Perry, T. (2017). Reframing the English Grammar Schools Debate. Educational Review. http://dx.doi.org/10.1080/00131911.2016.1184132.

Perry, T., Cordingley, P., Johns, P., & Bradbury, M. (2018). International Review of Teacher Evaluation Systems: Executive Summary, Main Report, Technical Report and System Case Studies. Prepared for the Inter-American Development Bank (IDB).

Rose, S., & Rose, H. (2013). *Genes, Cells and Brains: The Promethean Promise of the New Biology*. London: Verso.

Sawyer, R. K. (2008). Optimising learning implications of learning sciences research. Innovating to learn, learning to innovate, 45, 35-98.

Schlesinger, M.A., Hassinger-Das, B., Zosh, J.M., Sawyer, J., Evans, N., & Hirsh, P. (2020). Cognitive Behavioral Science behind the Value of Play: Leveraging Everyday Experiences to Promote Play, Learning, and Positive Interactions. *Journal of Infant, Child and Adolescent Psychotherapy, 19*(2). https://doi.org/10.1080/15289168.2020.1755084

Sweller, J. (1994). Cognitive Load Theory, Learning Difficulty, And Instructional Design. *Learning and Instruction, 4*, 295-312.

Sweller, J. (2016). *Cognitive load theory, evolutionary educational psychology, and instructional design*. In D. C. Geary & D. B. Berch (Eds.), Evolutionary psychology. Evolutionary perspectives on child development and education (p. 291–306). Springer International Publishing. https://doi.org/10.1007/978-3-319-29986-0_12

Tibke, J. (2019). Why the Brain Matters: A Teacher Explores Neuroscience. SAGE Publications Limited.

Utterly, L., & Montgomery, P. (2017). The influence of the team in conducting a systematic review. *Systematic Review, 6*(1). https://doi.org/10.1186/s13643-017-0548-x

Youdell, D., & Lindley, M. R. (2018). *Biosocial Education: the social and biological entanglements of learning*. London: Routledge.

Weinstein, Y., Madan, C. R., & Sumeracki, M. A. (2018). Teaching the science of learning. Cognitive Research: Principles and Implications, 3(1), 2.

Weinstein, Y., Sumeracki, M., & Caviglioli, O. (2018). Understanding how we learn: A visual guide. Routledge

Worth, J., Nelson, J., Harland, J., Bernardinelli, D., & Styles, B. (2018). GraphoGame Rime. Evaluation report and executive summary. Education Endowment Foundation.

## 11. Bibliography for Scoping

### *Books*

Agarwal, P. K., & Bain, P. M. (2019). Powerful teaching: Unleash the science of learning. John Wiley & Sons.

Brown, P. C., Roediger III, H. L., & McDaniel, M. A. (2014). Make it stick. Harvard University Press.

Busch, B., & Watson, E. (2019). The Science of Learning: 77 Studies that Every Teacher Needs to Know. Routledge.

Chartered College of Teaching (2020) Cognition and Learning: Applying key insights from cognitive science and psychology in the classroom. *Impact* (8, Spring 2020)

Dehaene, S. (2020). How We Learn: The New Science of Education and the Brain. Penguin UK.

Didau, D., & Rose, N. (2016). What Every Teacher Needs to Know about... Psychology. John Catt Educational Limited.

Harrington, J., Beale, J., Fancourt, A., & Lutz, C. (2020). The 'BrainCanDo' Handbook of Teaching and Learning: Practical Strategies to Bring Psychology and Neuroscience into the Classroom.

Jensen, E., & McConchie, L. (2020). Brain-Based Learning: Teaching the Way Students Really Learn. Corwin.

Kirschner, P. A., & Hendrick, C. (2020). How Learning Happens: Seminal Works in Educational Psychology and What They Mean in Practice. Routledge.

Kitchen, W. H. (2017). Philosophical Reflections on Neuroscience and Education. Bloomsbury Publishing.

Tibke, J. (2019). Why the Brain Matters: A Teacher Explores Neuroscience. SAGE Publications Limited.

Weinstein, Y., Sumeracki, M., & Caviglioli, O. (2018). Understanding how we learn: A visual guide. Routledge.

### *Reviews, Meta-analyses and Practice-focused Reports*

Centre for Education, Statistics and Evaluation (2018) Cognitive Load Theory in Practice: Examples for the classroom. Available at: https://www.cese.nsw.gov.au//images/stories/PDF/Cognitive_load_theory_practice_guide_AA.pdf

Churches, R., Dommett, E. J., Devonshire, I. M., Hall, R., Higgins, S., & Korin, A. (2020). Translating Laboratory Evidence into Classroom Practice with Teacher-Led Randomized Controlled Trials—A Perspective and Meta-Analysis. Mind, Brain, and Education. https://doi.org/10.1111/mbe.12243

Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980–2016. Educational Research, 60, 276–291. https://doi .org/10.1080/00131881.2018.1493353

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. Applied Developmental Science, 24(2), 97-140.

Deans for Impact. (2015). *The Science of Learning.* Deans for Impact. Available: http://www.deansforimpact.org/wp-content/uploads/2016/12/The_Science_of_Learning.pdf

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. Psychological Science in the Public Interest, 14(1), 4-58.

Howard-Jones, P. (2014). Neuroscience and education: A review of educational interventions and approaches informed by neuroscience. Education Endowment Foundation, Millbank, UK.

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004. National Center for Education Research.

Rosenshine, B. (2012). Principles of Instruction: Research-Based Strategies That All Teachers Should Know. American educator, 36(1), 12.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin.

Wittwer, J., & Renkl, A. (2010). How effective are instructional explanations in example-based learning? A meta-analytic review. Educational Psychology Review, 22(4), 393-409.

Weinstein, Y., Madan, C. R., & Sumeracki, M. A. (2018). Teaching the science of learning. Cognitive Research: Principles and Implications, 3(1), 2.

### *Other Papers and Policy Documents*

Aronsson, L. (2020). Reconsidering the concept of difference: A proposal to connect education and neuroscience in new ways. Policy Futures in Education, 18(2), 275-293.

Ofsted (2019) *Education inspection framework: overview of research.* Ref. 180045, Available: https://www.gov.uk/government/publications/education-inspection-framework-overview-of-research

Goswami, U. (2006). Neuroscience and Education: From Research to Practice? *National Review of Neuroscience, 7*(5), 406-411

GOV.UK (2019) *Early Career Framework*. DFE-00015-2019. Available: https://www.gov.uk/government/publications/supporting-early-career-teachers.

Howard-Jones, P. A. (2014). Neuroscience and education: myths and messages. *Nature Reviews Neuroscience, 15*, 817-824. doi:10.1038/nrn3817

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. Learning and instruction, 4(4), 295-312.

## 12. Appendix 1 – Summary of Advisory Group Meeting 1 – 10/07/2020

**Area 1 – Application/Translation**

The advisory group generally considered the application of cognitive science in the classroom as growing, but nevertheless still in its beginning. There was a sense that cognitive science inspired practices were gaining momentum and furthermore an anticipation that they will be increasingly incorporated in policy.

Approaches inspired by cognitive science take time to implement and embed, as they may differ from 'normal' ways of working. Key barriers to their application were considered to include teacher resistance to change, and questions of how to translate and implement cognitive science concepts and theories in classroom practice. One issue, which was mentioned in relation to this was the danger of separating small or selected aspects of the learning process from a broader conceptual framework and losing the broader dynamics when translating a defined result from cognitive scientific research into the curriculum. Furthermore, the question of how to understand learning was raised, and it was noted that evidence of memorised information is not necessarily evidence of learning in itself.

It was generally agreed that in order for cognitive science inspired approaches to be successful and work well, their particular relevance needs to be clear to teachers. Without a deeper understanding of why a particular approach is adopted, and what its specific benefits are for students, there is a danger of approaches being seen merely as "techniques." Although there is a lot of engagement in schools with issues about cognition, it is questionable whether there is a coherent and consistent understanding across educational practitioners and leaders about what the evidence says and how to make use of it. It is furthermore important to acknowledge the many contextual factors around education and how research evidence can be translated into pedagogy in a system.

**Area 2 – More/less promising Approaches**

It was noted that many of the techniques mentioned in the conceptual map of the project are not new (e.g. retrieval practice) and may furthermore have been practised previously by "hunch" (e.g. quizzes) rather than as officially inspired by cognitive science informed practices. Many of these practices resonate with established practice and good pedagogy, but cognitive science may help us move towards a shared understanding and common language. At the same time, insights from cognitive science may help establish redundant practices that are commonly used, and show why some work and some don't.

When discussing the project conceptual map more specifically, the advisory group commented on the following specific points:

- The first five strategies mentioned in the concept map (spaced practice, interleaving, retrieval practice, dual coding and strategies to manage cognitive load) were seen as being pushed a lot in some schools. It was mentioned that it would be good to look specifically at the classroom-based evidence behind them, as they may not translate from lab to classroom as well as we think. Furthermore, it was mentioned that we need to be clear about what these strategies mean, how we define them, and what they look like when they are done well.

- 'Brain training' may be too broad a concept which needs to be defined or narrowed, as it can potentially covers many cognitive interventions. One way of making this more specific could be to separate computerised training.
- 'Brain-to-brain synchrony' was mentioned as an area which should perhaps have low priority or which could be included in a more general category on the influence of social interaction and peers on learning.
- Mindfulness (included in stress/anxiety on table) is a promising area relating to effects on cognition, and could be included as an additional strategy.
- Analogies, future basing, active promotion of explicit links, mapping and visual schema as an aid to learning were also suggested as additional areas to potentially add to the map.
- Play was also suggested as a potential area to look at, particularly as younger year groups are included in the review.
- It was suggested that we might need to explore the role of the brain in affective domains and how it pertains to topics like values or character-based education and other more holistic questions.
- While the concept map needs to recognise limits and boundaries to what can sensibly be interrogated, at the same time these boundaries should not be treated as wholly exclusionary, as something being practically challenging to explore doesn't automatically equate to it not being relevant. Achieving a linkage between lesson objective, method and underpinning science about cognition would be a massive achievement, even if the linkages are somewhat loose.

In terms of contextual factors, age was mentioned as a very important element. Much existing work has been focused on secondary schools and universities, but practices tried and tested in older pupils may not be relevant/effective for younger pupils. SEND and cultural factors were also considered as important contextual factors, but it was noted that there might not be enough studies to make firm conclusions.

**Potential Project Benefits (whole group closing discussion)**

The group considered that a successful outcome of the project would be to produce a trusted source of documents that practitioners can refer and relate to. Being able to achieve something that even broadly fits that description would be a major and powerful return from the project. It was considered important that the interpretations of broader findings were linked to particular subjects or phases of schooling. Even in the case of findings being consistent across different phases or subject areas, teachers may still be inclined only to engage with the parts of findings that obviously and overtly relate to their particular subject discipline and/or the phase of education they engage with. The project would help teachers feel more confident about their own professionalism and capacity to interrogate research around cognitive knowledge and research.

Finally, from a policy perspective, success was seen to involve less rigidity and certainty in the expectations about how policies should be translated. This relates to a general and cross-cutting point arising from the meeting, which was to avoid a reductive view, not only of cognitive science, but also of the concept of learning and consequently, of links between the two.

## 13. Appendix 2 – Cognitive Science Concept Map

| Cognitive Science Concept Map | | | | |
|---|---|---|---|---|
| **1. Approaches Informing the Design of Classroom Teaching and Learning (Core Focus)** | | | | |
| *Technique/Theory from Cognitive Science* | *Classroom Teaching and Learning Theorie(s)* | *Synonyms and Conceptual Connections* | *Examples of Classroom Teaching and Learning Practice(s)* | *Concept-Specific Contextual and moderating factors* |
| **Spaced practice** | **Distributing learning and retrieval opportunities** over a longer period of time leads to better retention on delayed tests, compared to massed practice | *Synonyms/related terms:* Spacing; spaced repetition; distributed practice; often compared to 'cramming' (opposite concept)<br><br>*May overlap with:*<br>- Interleaving<br>- Retrieval practice | - Introducing rapid quizzes<br>- Curriculum planning/sequencing | - Delay to the final test matters<br>- Spacing matters (a few days after learning?)<br>- Taught, homework/practice, test (the space can be between different stages of learning, from intro to practice) |
| **Interleaving** | **Switching** between different types of problem or different ideas within the same study session leads to better retention on delayed tests, compared to block practice (e.g. abcbcacab instead of aaabbbccc) | *Synonyms/related terms:* interleaved practice; interweaving<br>*May overlap with:*<br>- Spaced practice | - Introducing rapid quizzes<br>- Lesson planning<br>- Varied practice | - Difficult to disentangle from spacing outside of the lab<br>- Could effects be because of spacing rather than interleaving (i.e. is interleaving merely a form of spacing?)<br>- How related should interleaved material be? |

| | | | | |
|---|---|---|---|---|
| **Retrieval practice** | **Recalling** information from memory can promote long-term learning | *Synonyms/related terms:* test-enhanced learning; testing effect; recall; quiz; mind-maps<br>*Conceptual links:*<br>- Includes elaboration: very broad (integrating and organising new information with what we already know). Involves asking and explaining how things work<br>- Interleaving<br>- Spacing | - Introducing rapid quizzes<br>- Practice tests<br>- Flashcards<br>- Mind-maps from memory<br>- Blank sheet of paper ('brain dumps'/open recall)<br>- Cloze/Retrieval guides/scaffolding<br>- Elaborative interrogation<br>- Multiple choice testsFilling in the blanks | - Better after a delay<br>- Short answer might be better than multiple choice, unless this involves retrieval (e.g. plausible distractors)<br>- Needs to be successful and address misunderstandings<br>- May be more beneficial where students have lower working memory<br>- Elaboration-based retrieval is most helpful when students have more background knowledge |
| **Dual coding** | Introducing concepts using both **verbal** (i.e. words) and **non-verbal** information (i.e. pictures) is thought to increase the chance of remembering that concept, compared to if the stimulus was only coded one way (i.e. words *or* images) | *Synonyms/related terms:* mental representation; multimedia design; verbal/non-verbal information; opposing theory = propositional theory, which claims that mental representations are stored as propositions rather than as images<br><br>*Links with:*<br>- Cognitive load | - Effective use of text and graphics on slides<br>- Use of diagrams or cartoons/comic strips<br>- Timelines with images<br>- Diagrams<br>- Infographics<br>- 5 minute lesson plan (i.e. visual guide to lesson) | - Controversial theory<br>- Which pictures/words?<br>- How conceptually relevant? |

| | | | | |
|---|---|---|---|---|
| **Strategies to manage cognitive load** | Attention and working memory, are essential to learning but are **'limited capacity' resources** which can be over-loaded. Cognitive load theory focuses teachers on the efficiency of their explanation and presentation of new knowledge | *Synonyms/related terms:* 3 types (intrinsic, extraneous, and germane); situational demands<br><br>Links with:<br>- Dual coding | - Worked examples<br>- Segmented presentation<br>- Pre-training<br>- Narration<br>- Strip back extraneous detail - Less 'redundant' information (i.e. not reading words off slides)<br>- Limit background distractions<br>- Optimum classroom seating arrangements | - Expertise reversal effect (i.e. if the instruction fails to provide guidance, low-knowledge learners often resort to inefficient problem-solving strategies, increasing cognitive load)<br>- Gradually fading out guidance best<br>- Situational interest is important<br>- Potential gender-specific effects for different types of cognitive load<br>- High cognitive load sometimes useful? |
| **Concrete examples** | Using **specific examples** to understand **abstract ideas** | *Synonyms/related terms:* analogies; models; modelling; real-life examples<br><br>Links with:<br>- Dual coding | - Analogies<br>- Models/modelling, especially in science (i.e. to help with visualising microscopic/subatomic structures)<br>- Check background knowledge<br>- Using real-life situations/scenarios when explaining new concepts<br>- Students generate their own concrete example | - Number of examples needed?<br>- Can backfire when a) distracts attention, b) surface features are too salient<br>- Important to use varied examples, so that students do not associate meaning of the concept with only one specific example |

| | | | | |
|---|---|---|---|---|
| **Brain training** | **Cognitive training** programmes that are designed to boost cognitive functioning, either as a whole, or a specific aspect (e.g. working memory) | *Synonyms/related terms:* cognitive training programmes; brain training games; cognitive training software; executive functioning training<br><br>Potentially, this category could overlap significantly with other concepts, including retrieval practice, game-based learning, spacing, exercise | - Many commercialised programmes (e.g. ACTIVATE by C8, BrainWare)<br>- Can include a number of different tasks/activities, including memory games, puzzles, multi-tasking games | - Can be classroom-based or online (e.g. apps) and/or at home<br>- Can be broad (e.g. for whole class) or specific (e.g. for students with ADHD or learning disabilities)<br>- How transferable/generalisable are the 'skills' developed in these programmes to everyday learning? (i.e. transfer effects)<br>- Age is important |
| **2. Approaches Involving Physical Factors (Wider Concepts)** | | | | |
| *Technique/Theory from Cognitive Science* | *Classroom Teaching and Learning Theorie(s)* | *Synonyms and Conceptual Connections* | *Examples of Classroom Teaching and Learning Practice(s)* | *Concept-Specific Contextual and Moderating Factors* |
| **Exercise** | It is thought that **exercise** increases efficiency of neural networks that are important for learning, whereby episodes and regimes of exercise can improve cognitive function and memory | *Synonyms/related terms:* physical activity (vs sedentary activity); embodied cognition<br><br>*Links with:*<br>- Interleaving (if used within lesson)<br>- Stress reduction | - Daily mile<br>- 'Educational Kinesiology' programmes, such as Brain Gym<br>-Physical activity can include curriculum content (e.g. recalling times-tables while performing activity) | - Variable format (e.g. additional slot on timetable, start of school day, embedded at start of lesson or within lesson)<br>- Effects can be measured acutely (e.g. working memory immediately after short burst of physical activity) or long-term (e.g. change in cognitive performance following |

| | | | | |
|---|---|---|---|---|
| | | | | implementation of physical activity regime) |
| **Nutrition/Hydration** | **Hunger and malnutrition** can affect cognitive performance. Hunger affects many aspects of cognition (e.g. working memory, attention), as well as emotional factors (e.g. motivation, engagement) Habitual ingestion of caffeine/dehydration can also reduce cognitive function | Interventions may also be implemented alongside other approaches in this category (e.g. sleep, exercise) as part of a broader health programme | - Breakfast clubs<br>- 'Breakfast after the Bell' programmes<br>- Education about nutrition in lessons<br>- Nutrition programmes | - Could be classroom-based, or school-wide approach |
| **Sleep** | Sleep is important for rest and for **consolidating** the day's learning in **long-term** memory | Interventions may also be implemented alongside other approaches in this category (e.g. nutrition, exercise) as part of a broader health programme | - Changes to start/finish times of school day<br>- Sleep education (i.e. promoting knowledge about sleep) | |
| **3. Approaches Involving the Motivational or Emotional State of the Learner (Wider Concepts)** | | | | |
| *Technique/Theory from Cognitive Science* | *Classroom Teaching and Learning Theorie(s)* | *Synonyms and Conceptual Connections* | *Examples of Classroom Teaching and Learning Practice(s)* | *Concept-Specific Contextual and Moderating Factors* |
| **Mindfulness** | Involves **attending** to and **focusing** non-judgementally on whatever is happening in any given moment, including | *Synonyms/related terms:* gratitude; relaxation; mindfulness practice; mindfulness meditation | - Mindfulness-based stress reduction programmes<br>- Journalling<br>- Affirmations | - Can be digital (e.g. app) or in the classroom<br>- Can be school-wide or classroom-based |

| | | | | |
|---|---|---|---|---|
| | thoughts, feelings, bodily sensations, and surrounding environment. This is thought aid cognitive processing (e.g. attention) | *Links:*<br>Could potentially overlap with stress reduction/SEL | - Mindful walks<br>- Mindfulness breathing exercises<br>- Often implemented as part of broader intervention (e.g. wellbeing/mental health/healthy minds) | |
| **Stress/anxiety reduction** | **Acute or chronic stress** and anxiety can have detrimental effects on higher order cognitive functions, such as working memory | *Synonyms/related terms:* relaxation; breathing exercises; meditation; yoga; wellbeing; anxiety<br><br>Could potentially overlap with mindfulness/SEL/exercise | - Breathing exercises<br>- Guided meditation<br>- Yoga or other physical activity<br>- Often implemented as part of broader intervention (e.g. wellbeing/mental health/healthy minds) | - Can be digital (e.g. app) or in the classroom<br>- Can be school-wide or classroom-based |
| **Social and emotional learning** | Strategies that help students to effectively apply the **knowledge, attitudes, and skills** necessary to **understand and manage emotions**, set and achieve positive goals, feel and show empathy for others, establish and maintain positive | *Synonyms/related terms:* Emotional intelligence (training); emotional competence; emotion recognition; emotion understanding; emotion regulation/management; resilience; SEAL | - Several commercial programmes (e.g. RULER,<br>- Mood meter (to help students learn emotion recognition)<br>- Meta-moment (helps challenge impulses and negative behaviours) | - Can be school-wide or classroom-based<br>- Can have a compensatory effect (e.g. emotional intelligence can provide extra resources for a student to draw upon if cognitive ability is low)<br>- Emotional intelligence can be conceptualised using a trait (i.e. emotional self-efficacy) or ability (i.e. |

| | | | | |
|---|---|---|---|---|
| | relationships, and make responsible decisions. Being emotionally competent can have benefits for academic performance | *Potential overlap with:*<br>- Stress reduction<br>- Mindfulness<br>- Game-based learning | - School charters (to establish supportive and productive learning environment)<br>- Games to help students learn emotion words (and thus increase emotion knowledge) | emotion-related cognitive skills, such as emotion recognition) |
| **Reward/game-based learning** | Games provide rapid schedules of uncertain **reward/reinforcement** that stimulate the brain's reward system, which can positively influence the rate at which we learn | *Synonyms/related terms:* educational games; gameplay; positive reinforcement; operant conditioning; reward<br><br>*Potential links:*<br>- Concrete examples<br>- Retrieval practice | - Minecraft/Lego can be used to help with visuospatial learning<br>- Competitions<br>- Reward/behaviour reinforcement charts | - Enjoyment of game is an important factor<br>- Can involve creativity (e.g. Lego/Minecraft building) |
| **4. Approaches Involving Direct Measurement or Manipulation of Neural Activity (Wider Concepts)** | | | | |
| *Technique/Theory from Cognitive Science* | *Classroom Teaching and Learning Theorie(s)* | *Synonyms and Conceptual Connections* | *Examples of Classroom Teaching and Learning Practice(s)* | *Concept-Specific Contextual and Moderating Factors* |
| **Transcranial electrical stimulation** | Applying small **currents** to the scalp can benefit some cognitive functions and learning processes, | *Synonyms/related terms:* TDCS; TACS; TRNS; evoked potentials | Not applicable - not yet established as classroom practice | - Transcranial electrical stimulation may improve learning difficulties in atypical brain development? |

| | | | | |
|---|---|---|---|---|
| | potentially by increasing neuroplasticity | | | |
| **Brain-to-brain synchrony** | Brain-to-brain synchrony - the coupling of behavioural and biological signals during social contact - is a possible **neural marker for dynamic social interactions**, likely driven by shared attention mechanisms. Greater synchrony may lead to more positive learning outcomes | *Synonyms/related terms:* brain coordination; EEG; fMRI; BOLD; neural synchrony; interbrain synchrony; social connectedness; dyadic interactions; group dynamics; student engagement | Not applicable – not yet established as classroom practice | - Social context is important<br>- Potential for investigating teacher-student dynamics and student engagement<br>- Can depend on how much people like each other (close relationships = greater synchronisation) |

# 14. Appendix 3 – Data Extraction and Coding Tools

*Appendix 3a – EEF main data extraction tool*

| | | |
|---|---|---|
| Publication information | Publication Type | Journal article<br>Report<br>Dissertation or thesis<br>Technical report<br>Book or book chapter<br>Conference paper<br>Other |
| Research type and method | Name of intervention | |
| | Description of the interventions | |
| | Objectives of intervention | |
| | Is there more than one treatment group | Yes<br>No<br>Not specified or N/A |
| | Assignment of participants | Random,<br>Non-random/Matched<br>Non-random/non-matched prior to treatment<br>Natural sample<br>Retrospective Quasi Experimental Design<br>Regression discontinuity<br>Unclear |
| | Level of assignment | Individual<br>Class |

| | | |
|---|---|---|
| | | School-cluster<br>School whole site<br>Region<br>Not provided |
| | How realistic was the study? | High/Low/Unclear Ecological Validity |
| Location | Study country | |
| | Additional information | Name of city, region or district<br>Rural/urban/sub-urban<br>No further information |
| Educational setting | Preschool/Nursery<br>Primary school<br>Middle school<br>Secondary/High school<br>Residential/Boarding School<br>Private/Independent School<br>Home<br>Further education/Junior or Community College<br>Other educational setting<br>Outdoor adventure setting | |
| Study Sample | Overall number of participants (both intervention and control) | |
| | Gender | Male<br>Female<br>Mixed<br>No Information |
| | Age | 3 – 18 |
| | Proportion of FSM/low SES children in the sample | Add specific indicators of FSM/Low SES |

| | | |
|---|---|---|
| The intervention | Who was responsible for the intervention? | School<br>Charity/NGO<br>University researchers<br>Local Authority<br>Private or commercial company<br>Other |
| | Was training provided for the delivery team? | Yes<br>No<br>Unclear |
| | Who was the focus of the intervention? | Students<br>Teachers<br>Teaching assistants<br>Other education Practitioners<br>Non-teaching staff<br>Senior Management<br>Parents<br>Others |
| | Teaching/intervention approach | Large groups<br>Small groups<br>Pairs<br>One-to-one<br>Students alone |
| | Was digital technology involved? | Y/N |
| | Were parent/community volunteers involved? | Y/N |
| | When was the intervention delivered? | During regular school hours<br>Before/After school |

| | | |
|---|---|---|
| | | Evenings/Weekends<br>Summer holiday period<br>Other<br>Not specified |
| | Who was responsible for the teaching at the point of delivery? | Research staff<br>Class Teachers<br>Teaching assistants<br>Other school staff<br>External teachers<br>Parents/Carers<br>Lay persons<br>Peers<br>Digital technology<br>Unclear |
| | Duration of the intervention | |
| | Frequency of the intervention | |
| | Length of intervention sessions | |
| | Are implementation details and/or fidelity details provided? | Quantitative<br>Qualitative<br>No details |
| | Costs of the intervention | Amount<br>Not-specified |
| Evaluation of the interventions | Who undertook the evaluation? | The developer<br>A different organization paid by developer<br>An organization commissioned independently to evaluate<br>EEF evaluation |

| | | |
|---|---|---|
| | | Unclear/not stated |
| | Reported primary outcomes | Standardised test<br>Researcher developed test School-developed test National test or examination International tests |
| | Curriculum subjects tested | Literacy (first language)<br>• Reading comprehension<br>• Decoding/Phonics<br>• Spelling<br>• Reading other<br>• Speaking/listening<br>• Writing<br>Mathematics<br>Science<br>Social studies<br>Arts<br>Languages<br>Other curriculum test |
| | Other reported outcomes | Yes<br>No |
| | If yes, which outcomes | Cognitive outcomes measured<br>Other types of student outcomes<br>Other participant outcomes |

*Appendix 3b – EEF effect size data extraction tool*

| | | |
|---|---|---|
| Study design | What type of study design is used for the evaluation of impact? | Individual RCT<br>Cluster RCT<br>Multisite RCT<br>Prospective QED<br>Retrospective QED<br>Interrupted time series QED<br>Regression Discontinuity with randomisation<br>Regression Discontinuity - not randomised<br>Regression Continuity - naturally occurring |
| | Are details of randomisation provided? | Yes<br>Not applicable<br>No/Unclear |
| Number of schools | What is the number of schools involved in the intervention group(s)? | |
| | What is the number of schools involved in the control or comparison group? | |
| | What is the total number of schools involved? | |
| | Not provided/ unclear / not applicable | |
| Number of classes involved | What is the total number of classes involved in the intervention group? | |
| | What is the total number of classes involved in the control or comparison group? | |
| | What is the total number of classes involved? | |
| | Not provided/ unclear / not applicable | |
| | What is the sample size for the intervention group? | |

| | | |
|---|---|---|
| Sample description | What is the sample size for the control group? | |
| | What is the sample size for the second intervention group? | |
| | What is the sample size for the third intervention group? | |
| | Does the study report any group differences at baseline? | Yes<br>No/Unclear |
| | Is comparability taken into account in the analysis? | Yes<br>No<br>Unclear or details not provided |
| | Is attrition or drop out reported? | Yes<br>No<br>Unclear (please add notes) |
| | What is the attrition in the treatment group? | |
| | Are the variables used for comparability reported? | Yes<br>No<br>N/A |
| | If yes, which variables are used for comparability? | Educational attainment<br>• Gender<br>• Socio-economic status<br>• Special educational needs<br>• Other (please specify) |
| | What is the total or overall percentage attrition? | |
| | Is clustering accounted for in the analysis? | Yes<br>No<br>Unclear |

| | | |
|---|---|---|
| Outcome details | Are descriptive statistics reported for the primary outcome? | Yes/No |
| | If yes, please add for the intervention* group | Number (n)<br>Pre-test mean<br>Pre-test standard deviation<br>Post-test mean<br>Post test standard deviation<br>Gain score mean (if reported)<br>Gain score standard deviation (if reported)<br>Any other information? |
| | If yes please add for the control group | (as previous) |
| | If yes, please add for a second intervention* group (if needed) | (as previous) |
| | If needed, please add for the second control group | (as previous) |
| | If yes, please add for a third intervention* group (if needed) | (as previous) |
| | If needed please add for a third control group | (as previous) |
| | Is there follow up data? | Yes<br>No |
| | Primary outcome | |
| | Secondary outcome(s) | |
| | SES/FSM outcome | |
| Outcome classification | Sample (select one from this group) | Sample: All<br>Sample: Exceptional<br>Sample: High achievers<br>Sample: Average<br>Sample: Low achievers |

| | | |
|---|---|---|
| | Test type (select one from this group) | Test type: Standardised test<br>Test type: Researcher developed test<br>Test type: National test<br>Test type: School-developed test<br>Test type: International tests |
| Effect size calculation (select one from this group) | What kind of effect size is being reported for this outcome? | Post-test unadjusted<br>Post-test adjusted for baseline attainment<br>Post-test adjusted for baseline attainment AND clustering<br>Pre-post gain |
| Toolkit strand(s) | Arts participation / Aspiration interventions / Behaviour interventions / Block scheduling / Built environment / Collaborative learning / Digital technology / Early years intervention / Extending school time / Feedback / Homework / Individualised instruction / Learning styles / Mastery learning / Metacognition and self-regulation / Mentoring / One to one tuition / Oral language interventions / Outdoor adventure learning / Parental engagement / Peer Tutoring / Performance pay / Phonics / Reading comprehension strategies / Reducing class size / Repeating a year / School uniform / Setting or streaming / Small Group Tuition / Social and emotional learning / Sports participation / Summer schools / Teaching assistants | |

*Appendix 3c – Revised Cochrane risk-of-bias tool for randomized trials (RoB 2)*

The following tool will be applied to all randomised trials. We will follow the guidance provided in the handbook when implementing the tool:

https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/current-version-of-rob-2

| | | |
|---|---|---|
| Preliminary | Study design | Individually-randomized parallel-group trial<br>Cluster-randomized parallel-group trial<br>Individually randomized cross-over (or other matched) trial |
| | | |
| | Intervention definition, Experimental: | |
| | Intervention definition, Comparator: | |
| | Specify which outcome is being assessed for risk of bias | |
| | Specify the numerical result being assessed. In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed. | |
| | Is the review team's aim for this result…? | • to assess the effect of assignment to intervention (the 'intention-to-treat' effect)<br>to assess the effect of adhering to intervention (the 'per-protocol' effect) |
| | If the aim is to assess the effect of adhering to intervention, select the deviations from intended intervention that should be addressed (at least one must be checked): | • occurrence of non-protocol interventions<br>• failures in implementing the intervention that could have affected the outcome<br>non-adherence to their assigned intervention by trial participants |
| | Which of the following sources were obtained to help inform the risk-of-bias assessment? (tick as many as apply) | Journal article(s)<br>Trial protocol<br>Statistical analysis plan (SAP)<br>Non-commercial trial registry record (e.g. ClinicalTrials.gov record)<br>Company-owned trial registry record (e.g. GSK Clinical Study Register record)<br>"Grey literature" (e.g. unpublished thesis)<br>Conference abstract(s) about the trial |

| | | |
|---|---|---|
| | | Regulatory document (e.g. Clinical Study Report, Drug Approval Package)<br>Research ethics application<br>Grant database summary (e.g. NIH RePORTER or Research Councils UK Gateway to Research)<br>Personal communication with trialist<br>Personal communication with the sponsor |
| Domain 1: Risk of bias arising from the randomization process | 1.1 Was the allocation sequence random? | Y/PY/PN/N/NI |
| | 1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions? | Y/PY/PN/N/NI |
| | 1.3 Did baseline differences between intervention groups suggest a problem with the randomization process? | Y/PY/PN/N/NI |
| | Risk-of-bias judgement | Low / High / Some concerns (Calculated using algorithm based on previous items) |
| | Optional: What is the predicted direction of bias arising from the randomization process? | NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |
| Domain 2: Risk of bias due to deviations from the intended interventions (effect of assignment to intervention) | 2.1. Were participants aware of their assigned intervention during the trial? | Y/PY/PN/N/NI |
| | 2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial? | Y/PY/PN/N/NI |
| | 2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context? | Y/PY/PN/N/NI |
| | 2.4 If Y/PY to 2.3: Were these deviations likely to have affected the outcome? | Y/PY/PN/N/NI |
| | 2.5. If Y/PY/NI to 2.4: Were these deviations from intended intervention balanced between groups? | Y/PY/PN/N/NI |
| | 2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention? | Y/PY/PN/N/NI |
| | 2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomized? | Y/PY/PN/N/NI |
| | Risk-of-bias judgement | Low / High / Some concerns (using algorithm) |

| | | |
|---|---|---|
| | Optional: What is the predicted direction of bias arising from the randomization process? | NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |
| Domain 2: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention) | 2.1. Were participants aware of their assigned intervention during the trial? | Y/PY/PN/N/NI |
| | 2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial? | Y/PY/PN/N/NI |
| | 2.3. [If applicable:] If Y/PY/NI to 2.1 or 2.2: Were important nonprotocol interventions balanced across intervention groups? | Y/PY/PN/N/NI |
| | 2.4. [If applicable:] Were there failures in implementing the intervention that could have affected the outcome? | Y/PY/PN/N/NI |
| | 2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes? | Y/PY/PN/N/NI |
| | 2.6. If N/PN/NI to 2.3, or Y/PY/NI to 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of adhering to the intervention? | Y/PY/PN/N/NI |
| | Risk-of-bias judgement | Low / High / Some concerns (using algorithm) |
| | Optional: What is the predicted direction of bias due to deviations from intended interventions? | NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |
| Domain 3: Risk of bias due to missing outcome data | 3.1 Were data for this outcome available for all, or nearly all, participants randomized? | Y/PY/PN/N/NI |
| | 3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data? | Y/PY/PN/N/NI |
| | 3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value? | Y/PY/PN/N/NI |
| | 3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value? | Y/PY/PN/N/NI |
| | Risk-of-bias judgement | Y/PY/PN/N/NI |
| | Optional: What is the predicted direction of bias due to missing outcome data? | Y/PY/PN/N/NI |
| Domain 4: Risk of bias in measurement | 4.1 Was the method of measuring the outcome inappropriate? | Y/PY/PN/N/NI |
| | 4.2 Could measurement or ascertainment of the outcome have differed between intervention groups? | Y/PY/PN/N/NI |

| | | |
|---|---|---|
| of the outcome | 4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants? | Y/PY/PN/N/NI |
| | 4.4 If Y/PY/NI to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received? | Y/PY/PN/N/NI |
| | 4.5 If Y/PY/NI to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received? | Y/PY/PN/N/NI |
| | Risk-of-bias judgement | Low / High / Some concerns (using algorithm) |
| | Optional: What is the predicted direction of bias in measurement of the outcome? | NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |
| Domain 5: Risk of bias in selection of the reported result | 5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis? | Y/PY/PN/N/NI |
| | Is the numerical result being assessed likely to have been selected, on the basis of the results, from... | Y/PY/PN/N/NI |
| | 5.2. ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain? | Y/PY/PN/N/NI |
| | 5.3 ... multiple eligible analyses of the data? | Y/PY/PN/N/NI |
| | Risk-of-bias judgement | Low / High / Some concerns (using algorithm) |
| | Optional: What is the predicted direction of bias due to selection of the reported result? | NA / Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable |

| | | |
|---|---|---|
| Overall risk of bias | | Low / High / Some concerns |
| | | Favours experimental / Favours comparator / Towards null /Away from null / Unpredictable / NA |
| ***Where:*** | Low risk of bias | The study is judged to be at low risk of bias for all domains for this result |
| | Some concerns | The study is judged to raise some concerns in at least one domain for this result, but not to be at high risk of bias for any domain. |
| | High risk of bias | The study is judged to be at high risk of bias in at least one domain for this result. Or The study is judged to have some concerns for multiple domains in a way that substantially lowers confidence in the result |

*Appendix 3d – Cochrane GRADE Tool*

The GRADE tool will be used to make an overall assessment of evidence in a cognitive science concept area. We will adhere to guidelines for the GRADE handbook: https://gdt.gradepro.org/app/handbook/handbook.html

**Quality of Evidence Levels Definitions**

- High: Very confident that the true effect lies close to the estimate of effect
- Moderate: Moderate confidence that the true effect lies close to the estimate of effect
- Low: Limited confidence that the true effect lies close to the estimate of effect
- Very Low: Very little confidence in the estimate of effect

**GRADE Quality of Evidence Assessment Process and Rating**

<table>
<tr><th>Study Design <u>At Entry</u> Into GRADE System</th><th>Quality of Evidence on Entry</th><th>Lower Category If…</th><th>Higher Category If…</th><th colspan="2"><u>Final</u> Quality of Evidence Rating (Select One)</th></tr>
<tr><td>RCT</td><td>HIGH</td><td rowspan="5">Risk of Bias<br>-1 Serious<br>-2 Very Serious<br>Inconsistency<br>-1 Serious<br>-2 Very Serious<br>Indirectness<br>-1 Serious<br>-2 Very Serious<br>Imprecision<br>-1 Serious<br>-2 Very Serious<br>Publication Bias<br>-1 Serious<br>-2 Very Serious</td><td rowspan="5">Effect Size<br>+1 Large<br>+2 Very Large<br>Dose Response<br>+1<br>All plausible confounders would reduce a demonstrated effect<br>+1<br>All plausible confounders would suggest a spurious effect when the results show no effect<br>+1</td><td>HIGH<br>++++</td><td></td></tr>
<tr><td rowspan="2">Observational Study</td><td rowspan="2">LOW</td><td>MODERATE<br>+++0</td><td></td></tr>
<tr><td>LOW<br>++00</td><td></td></tr>
<tr><td colspan="2" rowspan="2"></td><td>VERY LOW<br>+000</td><td></td></tr>
<tr><td colspan="2"></td></tr>
</table>

*Appendix 3e – Ecological Validity Assessment*

| | | |
|---|---|---|
| From EEF Extraction Tool | Duration of the intervention | |
| | Frequency of the intervention | |
| | Length of intervention sessions | |
| | How realistic was the study? | Higher ecological validity<br>Lower ecological validity<br>Unclear |
| Additional | Give details of previous | (Open response) |
| From EEF ES Extraction Tool | What is the number of schools involved in the intervention group(s)? | |
| | Who was responsible for the teaching at the point of delivery? | Research staff<br>Class teachers<br>Teaching assistants<br>Other school staff<br>External teachers<br>Parents/carers<br>Lay persons/volunteers<br>Peers<br>Digital technology<br>Unclear/not specified |

*Appendix 3f – Additional quality appraisal items for relevance, topic and detail*

| | | |
|---|---|---|
| Relevance to our questions | Relevance to our core systematic review objectives/questions | High/Medium/Low/None |
| Internal validity for answering its own question | Study research question(s) | (Open response) |
| | Based on RoB, EEF extraction items and other threats to validity, what is the level of internal validity of the study to answer its own question? | High/Medium/Low/None |
| | Give details of previous | (Open response) |
| Topic boundaries | Adherence to definition of cognitive science and its boundaries | High/Medium/Low/None |
| | Application to learning and its boundaries – e.g. teaching and learning in formal learning contexts | High/Medium/Low/None |
| | Meaningful connection with targeted interventions (and or other well-established interventions) | High/Medium/Low/None |
| Amount of detail regarding factors that affect the potential of studies to inform guidance | Detail provided regarding teaching and learning processes | High/Medium/Low/None |
| | Clarity of agency of intervention agent (teacher, machine, learning support assistant) | High/Medium/Low/None |
| | Detail provided regarding relationship with particular subject or phase of the curriculum | High/Medium/Low/None |
| | Detail provided regarding CPD provided to teachers | High/Medium/Low/None |
| | Detail provided regarding links with relevant, broader school policies | High/Medium/Low/None |
| Conflicts of interest | Researcher developed test | Y/N |
| | Researcher developed intervention | Y/N |
| | Other | (State what) |

*Appendix 3g – Additional items for topic and detail*

(Codes produced from our underpinning cognitive science review and practice review as well as input from the advisory group, to flag and extract data relating to cognitive sciences concepts evident within the interventions – we expect to make small revisions and additions to this after evidence mapping and prior to the main synthesis/analysis, using the codes mainly to support and record our categorisation of studies against our conceptual framework and organisation within the analysis and reporting).

| | | |
|---|---|---|
| Core cognitive science concepts [I.e. does intervention include elements of...] | Spaced practice / Interleaving / Dual coding / Retrieval practice / Cognitive load | Y/N (for each item) |
| Wider cognitive science concepts | Concrete examples / Brain training / Exercise / Nutrition/hydration / Sleep / Mindfulness / Stress/anxiety reduction / Social and emotional learning / Reward/game-based learning / Transcranial electrical stimulation / Brain-to-brain synchrony / Other | Y/N (for each item) |
| State and trait individual factors | Did the study measure or account for any *state/situational* factors? | Stress / Anxiety / Affect / Fatigue / Hunger/Thirst / Motivation / Other (specify) |
| | Did the study measure or account for any *trait* factors? | Personality traits / Cognitive ability / Autistic traits / Motivational traits / Mental health/well-being / Other (specify) |
| Sample (additional to EEF) | Is there provided any information about ethnicity? | Yes (specify)<br>No |
| | Is there provided any information about SEND? | Yes (specify)<br>No |