Ambition Institute

When and how to use visuals effectively: a meta-analysis of dual coding in the classroom

Briony Banks, Jennifer Curran, Rebecca Hall, Sam Sims, Stefanie Meliss

2026

Contents

Contributions

Conceptualisation

Briony Banks
Sam Sims
Jennifer Curran

Methodology

Briony Banks
Sam Sims
Jennifer Curran

Software

EPPI Reviewer
RStudio

Abstract screening

Briony Banks
Rebecca Hall
Jennifer Curran
Sam Sims
Stefanie Meliss

Full article screening

Briony Banks
Rebecca Hall
Jennifer Curran

Data extraction

Briony Banks
Rebecca Hall
Jennifer Curran

Analysis

Briony Banks
Jennifer Curran

Data curation

Briony Banks

Writing – original draft

Briony Banks
Jennifer Curran

Writing – review and editing

Sam Sims
Stefanie Meliss

Visualisation

Briony Banks

Supervision

Rebecca Curtis

Acknowledgements

The project was funded by Ambition Institute with funding acquired by Hilary Spencer and Marie Hamer, MBE. Special thanks to Dr Tom Perry for reviewing this paper.

Abstract

Visuals such as pictures and diagrams are an essential part of a teacher's toolkit. When used alongside verbal or written text (here termed 'dual coding'), they offer a way to explain concepts that words alone cannot, and an additional route to retaining information (Clark & Paivio, 1991). Much research has explored the effectiveness of dual coding for learning, but it has often been conducted in lab settings rather than schools, and results vary. This gives teachers mixed signals about when and how to use visuals effectively, and where to invest their limited teaching and preparation time. We carried out a meta-analysis of 56 studies conducted in school settings to address these questions and provide teachers with evidence-based recommendations for applying dual coding effectively. We found that dual coding benefits learning at all ages but more so for older pupils. Particularly, it may have no effect on primary pupils' transfer. Using relevant visuals can help pupils to retain and understand information in the short term, but we found limited evidence of long-term benefits. Effects were largest when interventions applied one of three principles from the Cognitive Theory of Multimedia Learning (Mayer, 2008, 2014): signalling (linking visuals to verbal explanations or written text); segmenting content meaningfully; and avoiding irrelevant details. We recommend teachers use dual coding in whole-class teaching as part of their core offer to all pupils, particularly for teacher-paced or timed activities, and focus on the three most effective learning principles to ensure pupils gain maximum benefits.

Keywords: meta-analysis, dual coding, multimedia, cognitive theory, visual aids, teaching

Introduction

Teachers frequently use visuals such as pictures, diagrams or animations to enhance pupil learning. They are often used alongside verbal or written text to help teachers communicate complex or abstract ideas, and to help pupils understand and retain what they are learning. The use of visuals alongside verbal or written content is often called 'dual coding', based on Dual Coding Theory (DCT; Paivio, 1990). This predicts that learning will improve when visuals are included alongside written or verbal instruction. A large amount of research has been conducted into the benefits of dual coding, as well as the use of visuals more broadly. The Cognitive Theory of Multimedia Learning (CTML; Mayer, 2008, 2014) has led to much research into the way in which visuals can be used optimally, with the use of visuals in teaching often being termed 'multimedia learning' in the literature. However, much of this research has taken place with adult learners and particularly undergraduates, in well-controlled lab experiments rather than authentic school settings. Moreover, results are varied: a recent review of cognitive science in the classroom by the Education Endowment Foundation (EEF; Perry et al., 2021) concluded that 'dual coding' with pictures and diagrams is overall beneficial, but there may also be many moderating factors. Because of this, teachers may be unclear about how to use visuals effectively. Findings from the review suggest that more guidance is needed for teachers to understand exactly when and how visuals can be used to enhance pupils' learning.

As teachers have limited time to plan resources, it is essential for them to understand how they can use visuals optimally and efficiently in their teaching. The present systematic review and meta-analysis aims to address these questions. We focused on research conducted in schools to ensure the validity and applicability of our findings to school settings and only included randomised experimental studies to ensure they provided strong causal evidence. We examined multiple aspects of using visuals that would be beneficial to educators that are likely to moderate the effects of dual coding. Our review was based on key predictions from Dual Coding Theory but also took into account elements of the Cognitive Theory of Multimedia Learning.

Visuals in teaching

Teachers use a range of visuals to explain concepts, skills and processes, and to support pupil understanding. These might include pictures in a story book, diagrams in a science text, a map in a geography lesson, multimedia software, or concept maps and graphic organisers. They might be displayed on an interactive whiteboard, provided for pupils in their exercise books or shown on a working wall. They are most commonly used by teachers to improve pupils' learning (to support memory and understanding), but can also be used for other purposes such as increasing motivation, focusing attention, or clarifying text (Peeck, 1993). Previous research has suggested five core functions of visuals that teachers regularly use to accompany text (Carney & Levin, 2002; Levin, 1981; Levin et al., 1987): decoration (purely to decorate the page); representation (a literal illustration of the text); organisation (providing a structural framework for the text such as a flow chart); interpretation (helping to interpret a more abstract or complex concept); and mnemonic (memory enhancing).

Several publications aimed at teachers (for example, Caviglioli, 2019; Payne, 2022; Sumeracki, 2019) have advocated for them to apply 'dual coding' to support pupil learning – that is, using visuals in addition to verbal explanations or written text. However, there is concern that the theory can be misapplied, leading to tokenistic adding of symbols or icons which may not be appropriate or effective. Indeed, purely decorative pictures have been shown not to benefit learning, albeit from a small number of studies (Levin et al., 1987). In the UK, the use of dual coding to support learning is supported by the Initial Teacher Training and Early Career Framework, which recommends "combining a verbal explanation with a relevant graphical representation of the same concept or process, where appropriate" (Department for Education, 2024, p. 18). However, teachers need a clear understanding of what constitutes 'appropriate' use, otherwise they risk spending valuable time on resources that are not effective. Overall, it is crucial that teachers are supported to understand the benefits and limitations of using visuals in teaching, to ensure they enhance pupils' learning rather than hinder it.

Theoretical context

There are many theoretical reasons why visuals might benefit pupils' learning, but the most prominent stems from DCT (Paivio, 1979, 1990). DCT is a prominent psychological theory of how concepts are formed and represented in long-term memory. The theory recognises the contribution of both non-linguistic sensory input (visual information such as pictures and images) and language (words and phrases) to long-term conceptual memory. For example, our concept of a cat is likely represented by the image of a cat (a furry animal with four legs and a tail) and the word 'cat'. These two representations or 'codes' are connected in long-term memory to form a coherent concept (mental model) of a cat. Although our understanding of conceptual memory is now far more advanced, DCT offered a simple and useful prediction for teaching: that presenting pupils with visuals (such as pictures or diagrams) alongside verbal or written text would improve their learning and particularly long-term retention of information (Clark & Paivio, 1987, 1991). Encoding visual and linguistic 'codes' for a concept or mental model should lead to more robust mental representations in long-term memory. Indeed, 'dual-coding' in an education context is often synonymous with presenting a picture alongside written or verbal text.

However, teaching is highly complex and requires teachers to adapt their instruction to a range of situations, pupils and learning needs (Vogt & Rogalla, 2009; Wang, 1980). Teachers must also take into account the limited working memory and attentional capacities of pupils. Dual coding on its own, therefore, doesn't provide teachers with much guidance about how and when to use visuals effectively. The CTML (Mayer, 2008, 2014) attempted to do just this.

Building on DCT, it aimed to provide a framework for using visuals in teaching based on knowledge of working memory and attention, incorporating Cognitive Load Theory (Chandler & Sweller, 1991; Sweller, 2010) as well as Baddeley's model of working memory (Baddeley, 2000; Baddeley & Hitch, 1974). CTML emphasises the limitations of the visual and verbal channels (the limited capacity assumption). For example, the theory acknowledges that it may be difficult for pupils to process a picture and written text at the same time, as they both involve visual attention and visual working memory, which have a limited capacity. This limitation can be addressed by providing a verbal explanation with a visual instead of written text (the modality effect). CTML also highlights that visual and linguistic information need to be actively integrated for pupils to form a coherent mental model (the active processing assumption). Mayer (2014) proposed that multimedia designed to encourage active learning (selecting, organising and integrating information) would optimise the benefits that pupils can gain from accompanying visuals. He proposed five key design principles to encourage active learning when using visuals:

1) **Signalling** (also known as cueing) directs learners' attention to the salient elements of visual and linguistic teaching content and can help pupils to integrate them by highlighting correspondences between them. For example, a teacher might present a diagram and refer to it as they explain key concepts, or a picture might have annotations taken from the accompanying text.
2) **Contiguity** refers to visual content that is spatially integrated with written text or presented simultaneously with a verbal narrative. Contiguity supports the integration and organisation of learning content by avoiding split attention and helping learners to form an integrated mental model. For example, a teacher might present a picture at the same time as giving a verbal explanation, or pictures are presented directly next to each corresponding paragraph of text.
3) **Coherence** refers to the exclusion of irrelevant, extraneous or 'seductive' details in learning content. Coherent materials do not include details that are irrelevant to the lesson content and learning objectives, for example, irrelevant pictures, examples, animations or sound effects.
4) **Verbal redundancy** refers to linguistic content that is presented concurrently in two different modalities – that is, speaking over a written text. For example, a teacher might provide pupils with some written text but also read it to them, requiring pupils to listen and read at the same time. This splits pupils' attention and strains their verbal working memory, making it more difficult to process the content. To avoid verbal redundancy, pupils could read text themselves or teachers could present content verbally with an accompanying visual instead of written text. Note that using written key terms alongside a spoken narration does not count as verbal redundancy, and is associated with greater learning (Adesope & Nesbit, 2011). This is in contrast to fully redundant written presentations where the teacher reads aloud the same text the pupils are given in written form.
5) **Segmentation** refers to content presented in meaningful and coherent segments rather than as a continuous unit. For example, written materials are clearly presented as 'chunks' of content using separate headers and visuals, or a teacher uses spacing/pauses between steps when verbally explaining a process. Segmenting helps pupils to chunk new information, thus managing their cognitive load and creating discrete and structured concepts in memory (Rey et al., 2019). Segmented content can be system-paced (segments are presented and paced by the instructor) or self-paced (segments are paced by the learner).

These five principles aim to support pupils to integrate and organise learning content presented in different modalities, without splitting their attention or straining their working memory. The CTML therefore predicts that if teachers follow these learning principles in their teaching and resources, visuals will be more effective and provide greater benefits to pupils' learning, for example improving retention and understanding.

Current evidence for Dual Coding Theory and the Cognitive Theory of Multimedia Learning

There is a large evidence base around the use of visuals in learning and much empirical support for DCT and the CTML. Evidence has shown that visuals can improve retention and transfer (Carney & Levin, 2002; Noetel et al., 2022), reading comprehension (Guo et al., 2020), and problem solving (Hu et al., 2021). Several meta-analyses have examined evidence for the CTML principles such as the modality effect (Ginns, 2005; Reinwein, 2012), contiguity (Ginns, 2006; Schroeder & Cenkci, 2018), signalling (Alpizar et al., 2020; Richter et al., 2016; Schneider et al., 2018; Xie et al., 2017), and coherence (Rey, 2012; Sundararajan & Adesope, 2020). These analyses largely support the active learning principles, but several moderators have been identified. For example, the modality effect and coherence have a greater effect for system-paced content (Ginns, 2005; Rey, 2012), contiguity is more important for older (especially post-secondary) learners (Schroeder & Cenkci, 2018), and signalling has a greater effect when learners' prior knowledge is low (Richter et al., 2016). Recently, a meta-review of this literature examined the effectiveness of several design principles, including those from CTML (Noetel et al., 2022). The meta-review broadly found support for them, indicating that adhering to them should ensure that visuals are effective.

An important limitation of this evidence base is that a large proportion of it has been conducted in lab settings with university students. This is problematic for two key reasons. Firstly, a young adult population is cognitively and developmentally more advanced than school-age pupils, meaning that university students may benefit more from visuals than younger pupils. Indeed, a meta-analysis by Guo et al. (2020) observed that adults obtained a greater benefit from visuals during reading comprehension than secondary and primary pupils. Similarly, Schroeder and Cenkci (2018) found that the contiguity principle had a greater effect in older (post-secondary) learners compared to school-age learners.

Secondly, lab studies involve participants following a very controlled procedure in a controlled environment, especially when compared to how pupils learn in a typical school classroom. This difference in setting may mean that effects are larger from lab studies, especially as large-scale field studies tend to have much smaller effects (for example, Lortie-Forgues & Inglis, 2019). One meta-analysis (Sundararajan & Adesope, 2020) indeed found that extraneous details (such as decorative images) impaired learning significantly more in lab studies compared to field studies (the effect in lab studies was moderate, but low in field studies). Conversely, Noetel et al. (2022) found that effects of visuals were not moderated by the location of the research (higher education colleges versus schools), although they did not differentiate between secondary (high) schools and primary (elementary) schools. Most meta-analyses examining the use of visuals have not differentiated between lab and school-based contexts, and it is therefore unclear just how applicable the majority of findings are to everyday school settings.

A recent review from the EEF (part of 'Cognitive Science in the Classroom', Perry et al., 2021) addressed this issue by reviewing research conducted specifically with school-age pupils in school settings. Part of the review focused on experimental evidence of dual coding in schools. The authors identified 48 articles which met these criteria, covering an age range of 7-18 years. They concluded that pictures and diagrams potentially have a small positive effect on learning (they did not calculate an overall estimate), but that this varies greatly and can even have a negative effect. They highlighted the heterogeneity of results and suggested several potential moderators to explain this, including the function of visuals (decorative or informational), pupil characteristics (such as age and prior attainment) and the format of learning. They also highlighted the low ecological validity of many studies in terms of their delivery, for example interventions comprised one-off sessions delivered by researchers with minimal input from a teacher. Methodological limitations were also noted, such as varied and often small sample sizes ($N < 100$). A limitation of the review was that search terms were limited to 'dual coding' and 'cognitive science', meaning that some relevant articles may have been missed. Further, the authors did not conduct a meta-analysis to aggregate effects, or test whether potential moderators did indeed account for the heterogeneity of effects. Nevertheless, the review rightly indicates that in practice, visuals may have a more mixed effect on pupil learning, and that more evidence is needed to understand exactly when visuals are beneficial.

The present study

Despite the large evidence base for 'dual coding' in teaching, there is a clear need for more guidance for teachers on when and how to use visuals effectively alongside verbal or written content. Moreover, a careful moderator analysis of evidence from classroom settings is needed to derive reliable recommendations, and to understand which factors influence the benefit pupils can gain from dual coding. To address these needs, we expanded on the review by the EEF (Perry et al., 2021) by conducting a wider literature search of research on visuals in teaching, and a meta-analysis of effects. Meta-analysis is well-suited to analysing a large and variable evidence base as it provides a quantitative estimate of effects while taking into account variable sample sizes (more weight is given to studies with larger sample sizes). Importantly, we used more granular coding to classify aspects of studies (such as the function and design of visuals) that were not captured in the EEF review. This allowed heterogeneity to be analysed based on a range of moderators, helping to identify when and how visuals are most effective.

Aims and hypotheses

Our overall hypothesis was that lessons and learning resources which include visuals (non-linguistic content), presented in addition to verbal or written text (linguistic content), would benefit learning outcomes in real-life teaching environments when compared to verbal or written content alone. The main aim of our meta-analysis was to examine heterogeneity of this 'dual coding' effect – why visuals can be more or less beneficial to learning outcomes, or sometimes not effective at all. We did this by analysing moderators relating to the instructional design of interventions, pupil characteristics, and the research design. We predicted that the heterogeneity of effects would be explained by these moderators, and our pre-registered hypotheses are detailed in the following section.

Instructional design

Moderator 1: CTML active learning principles

We hypothesised that dual coding would have a greater effect when interventions applied one of the five active learning principles from the CTML (signalling, contiguity, coherence, verbal redundancy or segmentation). We predicted that contiguity and signalling would have the greatest moderating effect as they encourage integration of linguistic and visual content (for example, Schroeder & Cenkci, 2020) which, according to the CTML, is important for learning with multimedia (Mayer, 2014). We also hypothesised that the more principles included in the design of the intervention, the more effective the intervention would be. Lastly, we hypothesised that the five design principles would interact with learning outcomes (see Moderator 3), with a greater effect on understanding and transfer than retention, as the principles are designed to enhance deeper conceptual understanding rather than remembering.

Moderator 2: the function of visuals

We predicted that the function of visuals would explain some of the heterogeneity of dual coding effects. Perry et al. (2021) suggested that the function of a visual (such as decoration or to provide information) likely moderates the effect of dual coding on pupils' learning. Previous research indicates that decorative pictures have no effect on learning, while mnemonic, interpretation, organisation and representation do, with mnemonic visuals potentially having the largest effect (Levin et al., 1987). We analysed five core functions of visuals commonly used in education (Carney & Levin, 2002; Levin, 1981):

1) Representation: a literal representation of linguistic content that accurately mirrors part or all of it. Makes the text more 'concrete' and helps the reader visualise a particular place, person, thing or event. Example: an accurate story illustration.
2) Decoration: a visual that simply decorates the page, with little or no relationship to the learning content. Example: a picture of a dinosaur alongside a maths problem.
3) Interpretation: a visual analogy that makes abstract or complex concepts more concrete and understandable. Example: a diagram of an electrical circuit or the circulatory system.
4) Organisation: helps organise information by depicting relationships between concepts visuo-spatially. Provides a structural framework for linguistic content or depicts the organisational plan of a text. Example: a story map; a flow chart of the recycling process; an illustrated map.
5) Mnemonic: helps to remember the linguistic content by using mnemonic (memory enhancing) components that are designed to improve a reader's recall. Example: a picture of a sandwich to demonstrate use of speech marks.

We also predicted that the function of visuals would interact with learning outcomes: representation and mnemonic visuals would have the greatest effect on pupils' retention as they help pupils to form stronger representations in long-term memory. Meanwhile, organisation and interpretation visuals would have the greatest effect on understanding and transfer as they help pupils to form more complex mental models.

Moderator 3: learning outcome

We predicted that visuals would benefit three key learning outcomes: retention, understanding and transfer (see Methods for definitions). DCT predicts that dual coding will benefit retention of information by providing two 'codes' to form a stronger representation in long-term memory. The CTML predicts that dual coding also benefits pupils' understanding as it provides complementary information (visual and linguistic) that can be integrated to form a coherent mental model. Neither theory predicts whether dual coding is more beneficial to retention or understanding, so we did not predict a significant difference in effects for these outcomes. However, previous research indicates that dual coding may have a smaller effect on transfer; for example, visuals that include signalling have a greater effect on retention than transfer (Schneider et al., 2018; Xie et al., 2017). We also predicted that learning outcomes would interact with the CTML design principles, the function of visuals, and pupils' age (see Moderators 1, 2 and 8 for details).

Moderator 4: lesson pacing

Lessons can be self-paced (where pupils can process materials in their own time, including backtracking) or system-paced (pupils cannot pause or backtrack and have to move at the pace of the instructor or software). We hypothesised that dual coding would be more beneficial for system-paced lessons as it will support pupils' cognitive load and help them to process lesson content effectively; for example, the modality effect is greater for system-paced lessons (Ginns, 2005).

Moderator 5: subject

We hypothesised that dual coding would benefit all subjects (such as maths or science) being taught and tested in the intervention, but that effects would vary. Empirical evidence as to how much subject moderates the effect of visuals is unclear; a review of the modality effect found that subject was a significant moderator (Ginns, 2005), while a review of the contiguity principle found no difference between subjects (Ginns, 2006). Mixed results may in part be due to a predominance of science and maths interventions in the literature (for instance, Perry et al., 2021). We predicted that dual coding would be most beneficial in maths where visuo-spatial processing is related to maths performance (Allen et al., 2019), and least beneficial for learning second languages which may rely on the connections between speakers' first and second language (Hall & Cook, 2012).

Moderator 6: level of knowledge

Based on the knowledge dimension of the 2001 revision of Bloom's taxonomy (Anderson & Krathwohl, 2001), there are three main levels of knowledge that pupils might be required to learn: factual knowledge (such as dates or vocabulary), conceptual knowledge (more complex, organised knowledge such as theories, principles or schema), or procedural knowledge (how to do something)[1]. We hypothesised that visuals would benefit learning of all levels of knowledge, but particularly conceptual and procedural knowledge. Use of visuals is theorised to help learners form more complex mental models from the two modes of input (for example, Mayer, 2014), while procedural knowledge may benefit from the additional visuo-spatial representation (that is, making abstract procedures more concrete; for instance, Paivio, 1990).

Moderator 7: manual interaction with the visual

Some dual coding interventions involve pupils manually interacting with the visual by performing an action such as tracing, copying, colouring or labelling (Caviglioli, 2019), while in others pupils only view and mentally process the visual without performing an action. We hypothesised that manually interacting with visual content would lead to greater effects on learning outcomes compared to only viewing it, as pupils could form more robust referential connections between the visual content and the concept/referent through multimodal encoding (for example, Engelkamp, 1998).

1 The taxonomy also includes metacognitive knowledge, which refers to pupils having knowledge of cognitive strategies or their own strengths and weaknesses. We did not include this form of knowledge as we considered it less likely to be used in the studies we reviewed. Indeed, only two studies (Alvermann, 1981; Sherman, 1976) did not fit into the three categories of knowledge, but this was due to unclear reporting rather than pupils learning metacognitive knowledge.

Pupil characteristics

Moderator 8: pupil age/education phase

We hypothesised that visuals would benefit pupils of all ages and phases of education, but that effects would be larger for older (secondary/high school) compared to younger (primary/elementary) pupils. Older pupils may have better active learning skills and a better ability to integrate visual and linguistic content (Guo et al., 2020; Schroeder & Cenkci, 2020). We also hypothesised that pupil age/education phase would interact with learning outcomes, in that older pupils are better able to form complex mental models from non-linguistic visual and linguistic content, leading to larger effects for understanding and transfer.

Moderator 9: prior knowledge

We hypothesised that pupils with low prior knowledge of the topic being taught would benefit more from dual coding than pupils with high prior knowledge. Additional visual content will likely help pupils with lower prior knowledge to form robust mental models. For example, the learning principle signalling (which helps pupils to actively integrate visual and verbal information) is more effective for pupils with low prior knowledge (Richter et al., 2016). Note that we do not consider prior knowledge effects to be related to pupil age, as pupils will always be taught age-appropriate topics based on their prior learning. For instance, a secondary pupil learning algebra for the first time may not have more prior knowledge of the topic than a primary pupil would when learning multiplication for the first time.

Moderator 10: pupil population

We hypothesised that effect sizes would be larger for studies conducted on a targeted subgroup of pupils (such as pupils with low prior knowledge or non-native speakers) compared to a broader/more universal sample (for example, all pupils from a year group) due to less heterogeneity between pupils (Kraft, 2020).

Research design

Moderator 11: instructor

We hypothesised that dual coding would be more effective when the person conducting the intervention was a qualified teacher due to their greater skill and experience (compared to a researcher, for example).

Moderator 12: teaching group size

We hypothesised that the effects of dual coding would be greater when pupils learnt individually or in small groups compared to whole-class teaching, as small-group teaching is known to be more effective (EEF, 2021). We categorised whether an intervention was carried out with a whole class, small group or pair, or with individual pupils.

Moderator 13: type of content

Content that is part of a carefully sequenced curriculum will be better connected to pupils' prior knowledge, which may lead to better integration of visual and verbal content and therefore larger dual coding effects. However, better prior knowledge may also mitigate the need for additional visual materials, leading to smaller effects. We therefore hypothesised that the type of content would moderate effects, but we did not predict this in a particular direction.

Moderator 14: outcome measure/intervention alignment

Alignment between the outcome measure and intervention refers to whether an outcome only tests knowledge and understanding of the materials and content provided in the intervention (narrow alignment), or whether it also tests knowledge and understanding beyond what was covered in the intervention (broad alignment). An example of a narrowly aligned outcome might be a researcher-developed test for an intervention, while a broadly aligned measure might be a standard measure of attainment such as Key Stage reading. We hypothesised that effects would be greater for narrow compared to broadly aligned measures (Kraft, 2020).

Moderator 15: timing of outcome measure

We hypothesised that the effect of dual coding would be smaller for longer periods between the intervention and outcome measurement, as pupils are required to retain the learnt knowledge for longer (Kraft, 2020).

Moderator 16: number of lessons

We hypothesised that the effects of dual coding would be larger for longer interventions due to greater exposure to the learning materials, leading to improved understanding and retention.

Moderator 17: learning method

Many studies in the literature base for dual coding have relied on using computer software for the intervention. As most teaching takes place in person, and computer interventions might be considered less ecologically valid, we felt this was an important moderator to examine. We categorised interventions in terms of whether pupils were taught in-person or using a computer, but we did not have specific predictions about when dual coding would be most effective.

Methods

The study was pre-registered as a meta-analysis on the Open Science Framework (OSF): https://osf.io/cn6qf. Full data, analyses and materials are also available on the OSF: https://osf.io/fx7e9/. Ethical approval was provided by the UCL IOE Research Ethics Committee.

PICO statement

We developed the scope of the project around the PICO (Population, Intervention, Comparison, Outcome) framework for meta-analysis.

The **population** of interest was pupils in early years, primary, secondary and further mainstream education (or equivalent if outside the UK), aged 3-18 years at the time of intervention.

The **intervention** of interest was the use of non-linguistic visual content in learning materials, in addition to linguistic content (spoken or written text). We defined non-linguistic visual content as pictures, photos, icons, diagrams, animations, videos or any visuo-spatial graphics (for example, charts, graphic organisers or concept maps).

The **comparator** of interest was verbal (spoken) or written/printed learning content that did not include any visuals, or 'business as usual' (as long as this did not include the visual content included in the intervention).

The primary **outcome** of interest was pupil learning measured using a test that captured retention, understanding or transfer. Retention refers to how well students remember what was explicitly taught in the intervention, measured in a test of recognition or recall of knowledge (for example, verification, matching, forced choice). Understanding refers to how well students understand what they were explicitly taught in the intervention. Testing understanding may involve interpreting, exemplifying, classifying, summarising, inferring, comparing, explaining, analysing, organising, attributing, evaluating, checking or critiquing. Transfer means transfer of knowledge and skills from one domain or context to another. Testing this might look like applying a learned skill, knowledge or strategy to a new or unfamiliar but relevant problem or task that was not explicitly taught in the intervention.

Inclusion criteria

Inclusion criteria (listed in Table 1) were piloted on a sample of 34 articles included in the review by Perry et al. (2021). The 34 articles were double-screened (by authors Briony Banks (BB) and Jennifer Curran (JC)) to test and refine the criteria, until we reached full agreement on a final set of inclusion criteria. Note that our final criteria are different to those used by Perry et al. (2021).

Table 1. Full inclusion criteria

Criteria	Definition
Research type	Primary research studies
Publication type	Journal articles, working papers, institutional reports, doctoral theses
Language	Reported in English
Intervention	Studies test use of non-linguistic visual content (see 'type of visual content') in teaching materials, during instruction. Experimental design must contain at least one condition with linguistic and non-linguistic visual content AND at least one control condition comprising the same without the visuals
Research setting	Intervention and assessment occur in a mainstream school, nursery or further education college in any country
Research design	Randomised experimental studies using a between-group design
Type of visual content	Illustrations, pictures, photos, diagrams, charts, animations, videos, concept/knowledge maps, graphic organisers, icons, story maps
Population	Early years/pre-kindergarten, primary/elementary, or secondary/middle/high school/sixth form pupils
Outcomes	Studies that measured pupil achievement using tests, in any school curriculum subject, capturing at least one of retention, understanding or transfer
Statistics	Sufficient statistics reported to calculate Cohen's *d* effect sizes (means and SDs/SEs/CIs, or *t*-/*F*-statistics, or *p*-values, plus group sample sizes)

Literature searches

We developed our literature searches from our PICO statement and inclusion criteria. Search terms for four key criteria (population, intervention, outcome and design) were developed using relevant articles from Perry et al. (2021). They were tested and refined using an iterative process, using PsycInfo (EBSCO) database to ensure that we retrieved seven key articles from Perry et al. (2021). We also removed any synonyms (for example, 'performance' for outcome; 'pupil' for population) which retrieved many irrelevant articles (for instance, 'pupil' retrieved many articles relating to vision research). This process ensured that the search was sufficiently broad to include all key search terms, and precise enough so the number of articles to screen would be practical given our resources. Appendix A1 lists our final search terms, and the databases and review articles used in our literature searches.

We searched six databases using these criteria (see Figure 1). We also retrieved references from 18 relevant systematic reviews and meta-analyses that had searched at least two databases (see Appendix A1). These review articles were identified during an initial literature search of reviews and meta-analyses relating to DCT and the CTML on Google Scholar, and from citations and reference lists of key articles (Carney & Levin, 2002; Clark & Paivio, 1991; Guo et al., 2020; Mayer, 2014; Noetel et al., 2022).

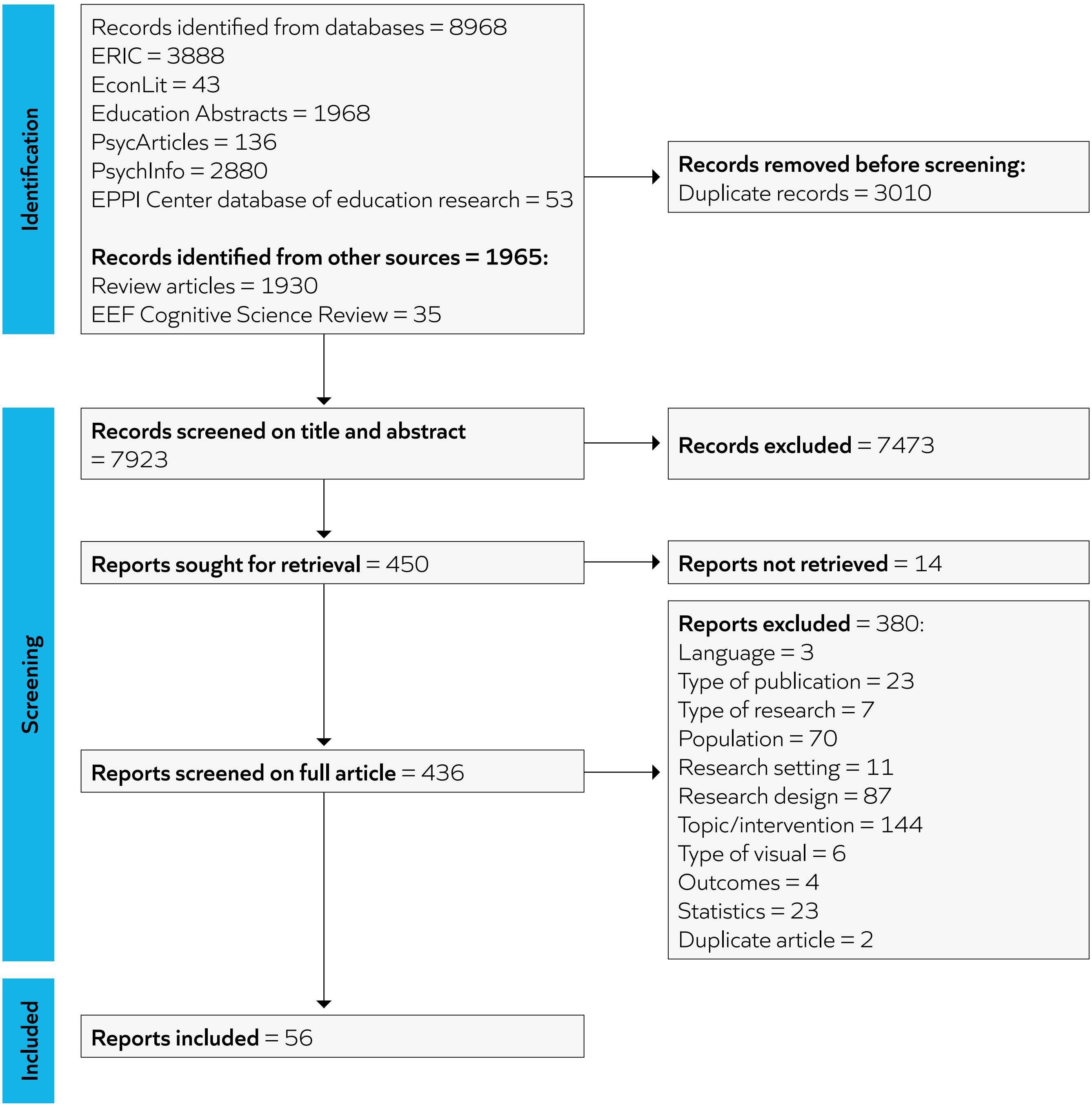
Article screening

Figure 1 shows the number of articles included and excluded at each screening stage. We used EPPI Reviewer software (Thomas et al., 2022) to screen all articles retrieved from our literature searches. The software automatically identified duplicate articles which were checked and excluded by the first author. We then single-screened remaining articles based on their title and abstract using all inclusion criteria except statistics. All authors contributed to this stage. We used inbuilt priority screening – a machine learning approach that learns over time as screening progresses, resulting in a more efficient process (O'Mara-Eves et al., 2015; Tsou et al., 2020). However, we did not identify a clear 'drop-off' in inclusions to indicate that we could stop screening, so we continued to screen all articles. Queries were resolved through discussion.

Next, we screened all included articles using the full text and inclusion criteria. Authors BB, JC and Rebecca Hall (RH) carried out this stage. To ensure consistency, we triple-screened the first 10% of articles (45 articles) – that is, the three reviewers independently screened the same sample. Agreement was 80%[2] between one pair of reviewers, but lower between the other two pairs (66% and 71%). We discussed disagreements and reached a consensus on each article and then triple-coded another 10% of articles (41 articles). Following this, agreement ranged from 80-83% between each pair of reviewers, so we proceeded to single-coding for all remaining inclusions (each remaining article was screened by one reviewer only). The three reviewers discussed a further 82 articles to reach a consensus on how to classify them. These discussions were often due to (lack of) clarity in reporting. For example, where papers were not explicit regarding randomisation methods, the exact content and procedure of the intervention, or how populations were referred to (American college versus UK university students). We discussed types of visual that should/should not be included, for example agreeing on a clearer definition of graphic organisers. We also discussed the statistics that were reported and whether we could use them to calculate relevant effect sizes that would answer our primary research question of whether dual coding is effective.

2 Our pre-registration stated we would use Cohen's kappa to measure reliability (with a threshold of ≥.80). However, due to the large number of articles, we required a third reviewer, so instead used % agreement based on inclusion/exclusion of articles.

Figure 1: PRISMA diagram of screened articles



Data extraction

We extracted data from all included articles using a detailed coding framework for each moderator (available on the OSF). Authors BB and JC piloted the framework prior to pre-registration by independently double-coding the 14 articles included from the screening pilot. Discrepancies were discussed and the coding tool was revised accordingly until agreement reached Cohen's $k \geq .80$ on all moderators.

Data extraction was carried out by authors BB, JC and RH. We triple-coded the first 12 articles (20%) and resolved disagreements through discussion. We then proceeded to single coding but continued to resolve queries through discussion.

Effect size extraction and calculation

To calculate effect sizes, we extracted group means and the pupil-level pooled standard deviation (*SD*) for each intervention and comparison group that met our criteria. If these statistics were not available, we extracted *t* or *F* statistics. We calculated standardised mean difference effect sizes (Cohen's *d* and Hedges *g*) so that positive effects indicated an improvement in learning from the visual intervention. We used the esc package in R (Lüdecke, 2019) to calculate all effect sizes plus the standard error (SE), variance, and 95% confidence intervals (CIs) for each outcome. We did not specify in our pre-registration which effect size we would use for analysis. A sensitivity analysis comparing Cohen's *d* and Hedges *g* showed little difference between the two (available on the OSF). However, we decided to use Hedges *g* for all analyses due to the small sample size in some studies (Cohen's *d* can have a slight positive bias for smaller sample sizes; Cooper et al., 2019). We extracted multiple effect sizes from the same study/experiment as separate data points if they related to different learning outcomes (for instance, retention and transfer, immediate and delayed measures), different age groups, or differences in the intervention that met our inclusion criteria (for instance, different types of visual, or different instructional design).

Data preparation

The first author checked the data from EPPI reviewer for errors, missing data and consistency (for instance, the format and spelling of moderator levels). For the following moderators, we extracted data exactly as it was reported and defined suitable categories post hoc based on the studies we included: education phase, subject, timing of outcome measure, and duration of intervention. We defined these categories to be meaningful to teachers, and to maximise the number of studies for each category level (for instance, we combined science and technology due to similarity in teaching content, and because only one study reported the subject as technology). We cleaned all data for analysis using the tidyverse package (Wickham et al., 2019) in RStudio 2024.12.0 running R version 4.4.2.

Analysis

We conducted a random effects meta-analysis for all included studies and for each moderator analysis. We used the robust variance estimation (RVE) method to address the statistical dependency between different outcomes within the same study (Hedges et al., 2010; Pustejovsky & Tipton, 2022). Correlated effects may arise from multiple types of dependency, for example if a study analyses different outcome measures or different time points for the same sample, or if it uses the same control group to compare different interventions. However, such correlations are rarely reported. RVE is appropriate as it does not require information on the covariance structure of the effect size estimates and is simple to implement.
We used the robumeta package (Fisher et al., 2023) in RStudio 2024.12.0 (running R version 4.4.2) with small-sample adjustments to residuals and degrees of freedom recommended by Tipton (2015). These adjustments result in more reliable estimates when there are fewer than 40 studies, as is the case with our subgroup analyses. RVE requires the within-study effect size correlation (Spearman's ρ) to be specified. A sensitivity analysis (available on the OSF) using the robumeta package assessed the impact of different ρ values on τ^2 and the pooled estimate, but indicated no differences. We therefore used the default value of Spearman's $\rho = 0.80$ for all meta analyses.

Moderator analyses

The primary moderators 'function of visual' and 'learning outcome' had insufficient studies per category to conduct meta-regression as planned. We therefore conducted subgroup analyses for each moderator as stated in our pre-registration. To analyse whether each moderator contributed to heterogeneity, we compared effect sizes for each subgroup level and conducted *t*-tests using the robumeta package (see Fisher & Tipton, 2015). We again used small-sample corrections (Tipton, 2015); however, the adjustments can inflate the Type 1 error rate if degrees of freedom < 4. This can occur with small sample sizes, particularly for study-level and highly imbalanced covariates. The recommended solution is to either use multivariate meta-analysis or to remove unusual data points (Tipton, 2015). As neither of these was suitable for our data, we decided not to report estimates for any subgroup categories with < 4 degrees of freedom to avoid misleading or incorrect conclusions (full analytical output with all subgroup categories is available on the OSF). This applied to the function of visual (mnemonic and decorative), one CTML principle (verbal redundancy), subject (maths), knowledge level (procedural), and outcome alignment (broad).

Publication bias

To assess the likelihood of publication bias in our analysis, we first plotted a contour-enhanced funnel plot (Peters et al., 2008) to identify small-study effects. Second, we conducted a trim-and-fill analysis (Duval & Tweedie, 2000) to identify theoretically 'missing' studies due to potential small-study effects. Third, we conducted a p-curve analysis (Simonsohn et al., 2014b, 2014a) to assess if there was a true effect in our data or not (that is, whether we could detect motivated *p*-hacking). Fourth, we used a selection model (Vevea & Hedges, 1995) to estimate the probability that studies would be published if results had different levels of significance. We modelled a weight function with two cut-points ($p = .1$ and $p = .05$), reflecting commonly used thresholds above which studies are less likely to be published (Sutton, 2009). We used one-tailed tests which imply publication in favour of increasingly significant positive outcomes. As these four methods assume independent effect sizes (that is, they assess publication bias of studies rather than outcomes), we first aggregated effects at the study level. To do this we used the 'aggregate' function from the metafor package in R which uses inverse variance weighting to combine effects from the same study. We specified the variance-covariance matrix of sampling errors for each study using a within-study correlation of Spearman's $\rho = 0.80$ (that is, we assumed the same within-study correlation as in our other meta-analyses). We conducted all publication bias analyses using the metafor package (Viechtbauer, 2010) in RStudio 2024.12.0 (running R version 4.4.2), except the *p*-curve analysis for which we used the dmetar package (Harrer et al., 2019).

Results

Descriptives

Our final meta-analysis included 56 studies, 189 outcomes and a total of 6675 pupils. The majority of interventions used pictures or diagrams, focused on science or literacy, and lasted for one lesson (see Table 2). Primary and secondary school age groups were equally represented in our analysis.

Most interventions were not carried out by the class teacher (instead using a researcher) and used an experimental control group rather than 'business as usual'. A summary of all included studies and outcomes are listed in the Appendix A2.

Table 2. Characteristics of the 56 studies included in the meta-analysis

Characteristic type	Characteristic	Number of studies (outcomes)	Proportion of studies*
Type of visual	Animation Concept map Diagram Graphic organiser Photo Picture	6 (17) 2 (6) 15 (38) 4 (15) 1 (1) 34 (112)	11% 4% 27% 7% 2% 61%
Education phase	Early Years and Primary (Preschool/ Kindergarten/Reception; Year 6/Grades 1-6)	29 (96)	52%
	Secondary and Further (Years 7-13/Grades 7-12; High School/ Middle school)	28 (93)	50%
Subject	English and literacy Graphic organiser Picture	17 (76) 1 (4) 16 (67)	30%
	Humanities Concept map Diagram Graphic organiser	6 (32) 1 (4) 4 (10) 1 (6)	11%
	Maths Animation Diagram	5 (17) 2 (6) 3 (11)	9%

Characteristic type	Characteristic	Number of studies (outcomes)	Proportion of studies*
Subject	Science and technology Animation Concept map Diagram Graphic organiser Photo Picture Second languages Picture	20 (52) 4 (11) 1 (2) 8 (17) 2 (5) 1 (1) 10 (28) 8 (17) 8 (17)	36% 14%
Duration of intervention	1 lesson 2-3 lessons 4-5+ lessons	38 (141) 7 (15) 11 (33)	68% 13% 20%
Instructor	Teacher Non-teacher Not reported	13 (42) 28 (85) 15 (62)	23% 50% 27%
Comparison group	Business as usual Experimental control	3 (10) 53 (179)	5% 95%
Country	Australia Belgium Canada Finland Germany Hong Kong Iran Israel Korea New Zealand South Korea Spain Taiwan The Netherlands Turkey UK USA	4 (10) 1(3) 2 (3) 1 (2) 5 (15) 1 (4) 1 (1) 1 (2) 1 (2) 1 (2) 1 (4) 1 (2) 1 (8) 1 (2) 1 (1) 2 (5) 32 (123)	7% 2% 4% 2% 9% 2% 2% 2% 2% 2% 2% 2% 2% 2% 2% 2% 57%

Note. *Proportions may not sum to 100% due to rounding, or because studies included more than one intervention, population or outcome.

Main meta-analysis: is dual coding overall effective in school settings?

The estimated impact of visuals on learning outcomes was positive, significant and relatively large ($g = 0.47$, $p < .001$; see Table 3). As expected, we observed substantial heterogeneity between studies[3] (forest plots available on the OSF). τ^2 indicated that on average, individual effects differed from the average pooled effect by 0.57 standard deviations, while I^2 indicated that 89% of this between-study variance was due to true heterogeneity of effects rather than sampling error. This means that the benefit pupils gained from visuals varied depending on different factors. The subgroup analyses in the following section aimed to explore potential sources of this heterogeneity.

Table 3. Meta-analysis indicating the average effect of dual coding on pupil learning

Average effect (*SE*)	95% *CI*	*p*	*k* (*n*)	Estimates of heterogeneity	
0.47 (0.10)	0.26, 0.68	<.001	56 (189)	$I^2 = 89.11$	$\tau^2 = 0.57$

Note. Dependence of within-study effect sizes specified as $\rho = 0.80$. k = studies, n = outcomes.

Subgroup analyses: instructional design

Moderator 1: CTML active learning principles

Meta-regression of the summary variable indicated that interventions which followed more active learning principles resulted in a small increase in pupil learning, but this was statistically non-significant: $b = 0.15$, SE = 0.09, $p = .108$, 95% CI = [-0.04, 0.33], $k = 53$. The general impact of the principles on pupil learning is therefore uncertain. Indeed, subgroup analyses of each principle indicated that some had a greater impact than others (see Figure 2). Visual content that included signalling, that was segmented into meaningful chunks, and that was coherent (with no irrelevant details), resulted in much greater learning on average, compared to visual content where these principles were not followed (signalling: $g = 0.66$, $p = .001$; no signalling: $g = 0.29$, $p = .001$; segmented content: $g = 0.60$, $p < .001$; non-segmented content: $g = 0.31$, $p = .011$; coherent materials: $g = 0.49$, $p = .001$; non-coherent materials: $g = 0.23$, $p = .283$). Although effect sizes were substantially larger when each of these principles was followed, the differences were not statistically significant, and there was still substantial heterogeneity in the data (see Table 4).

The contiguity principle (spatial or temporal integration of visuals) only slightly impacted learning; pupils significantly improved whether visuals followed this principle or not, but they improved slightly more when it was followed ($g = 0.50$, $p = .001$) compared to when it was not ($g = 0.43$, $p = .002$). Most studies did not include any verbal redundancy ($k = 46$) so we cannot conclude whether this principle had an impact on pupil learning.

Moderator 2: function of visual

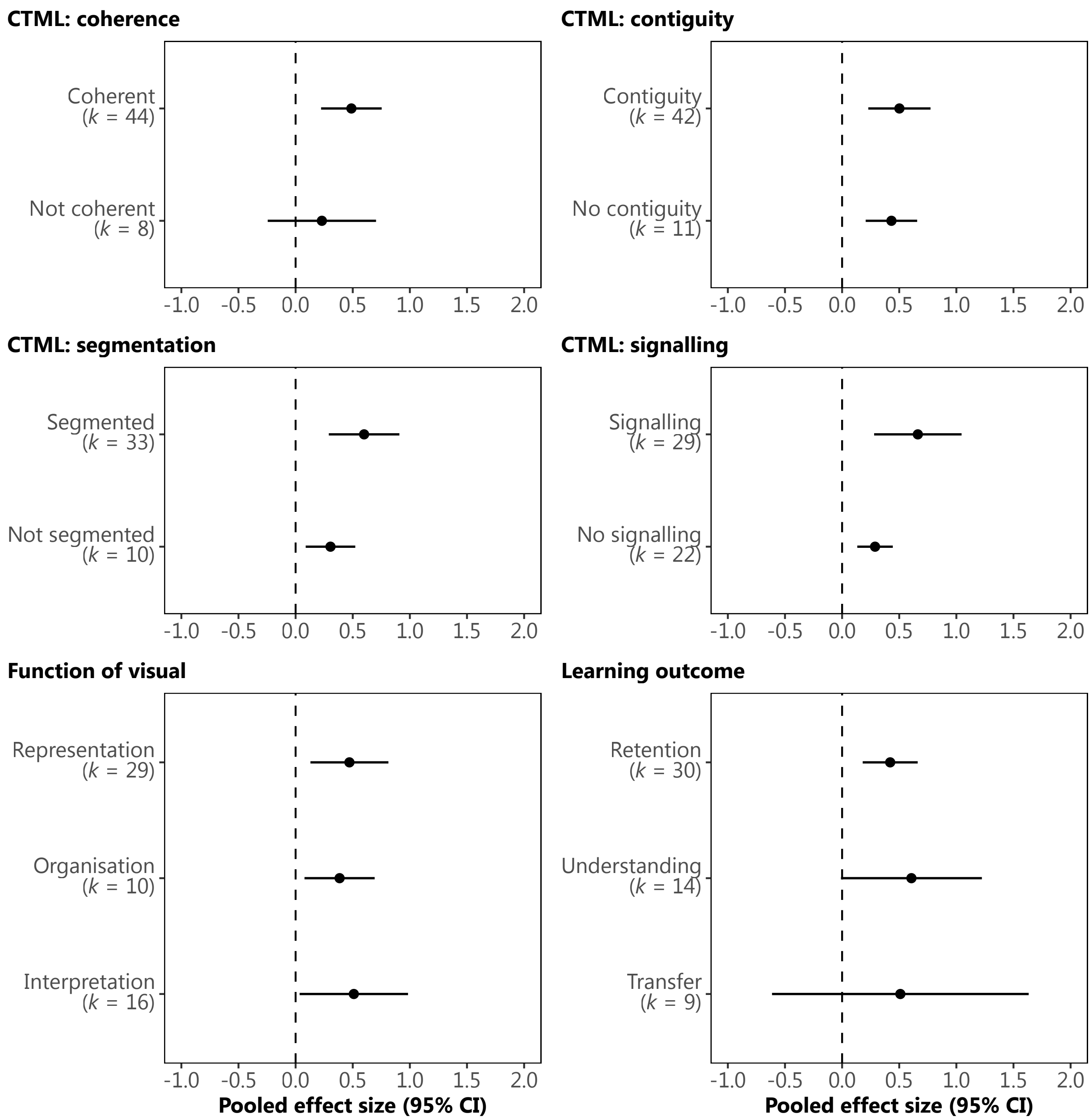
Visuals that represented linguistic content literally (representation), visuals that organised content (organisation), or that interpreted a concept (interpretation), all significantly improved pupil learning (see Figure 2). Effect sizes for these three functions were similar in magnitude and precision, although the average effect for interpretation was the largest (interpretation: $g = 0.52$, $p = .039$; representation: $g = 0.44$, $p = .011$; organisation: $g = 0.39$, $p = .020$). The difference in effects between representation and organisation/interpretation were small and not statistically significant (see Table 4). Estimates for mnemonic and decorative visuals were unreliable due to the low number of studies included.

3 Note that the RVE method (as used in the robumeta package) does not produce Cochrane's Q statistic as a measure of heterogeneity.

Moderator 3: learning outcome

Dual coding had a statistically significant and positive effect on pupils' retention ($g = 0.42$, $p = .001$; see Figure 2). It also had a large positive impact on understanding ($g = 0.61$, $p = .056$) and on transfer ($g = 0.51$, $p = .322$), although these estimates were not statistically significant and therefore less certain. In particular, 95% CIs for transfer were very wide and contained negative and positive values [-0.61, 1.63]. In contrast, 95% CIs for understanding contained mostly positive values [-0.01, 1.22], so we can be more confident that dual coding would have a positive impact on pupils' understanding than transfer. However, differences between the effect for transfer compared to those for retention and understanding were small and not statistically significant (see Table 4).

Figure 2. Subgroup analyses of moderators 1-3 relating to instructional design



Note. k = number of studies. The total number of studies for each analysis may not equal 56 as we were unable to code some outcomes (for example, if not reported clearly). Effects are not statistically significant if 95% CIs cross 0 (dashed line).

Table 4. Subgroup analyses for moderators 1-3 relating to instructional design.

Moderator	Mean difference in effect (*SE*)	95% *CI* (lower, upper)	*t*	*p*	*k* (*n*)	I^2	τ^2
1. CTML principles						-	-
Signalling / no signalling	0.37 (0.20)	-0.03, 0.78	1.86	.069	22 (79) 29 (87)	90.06	0.62
Segmented / not segmented	0.29 (0.18)	-0.09, 0.68	1.64	.123	33 (95) 10 (50)	87.83	0.54
Coherent / not coherent	0.26 (0.23)	-0.28, 0.80	1.10	.301	44 (156) 8 (17)	90.10	0.67
Contiguity / no contiguity	0.07 (0.17)	-0.29, 0.43	0.42	.682	42 (127) 11 (48)	89.70	0.61
2. Function of visual						89.82	0.62
Interpretation / representation	0.04 (0.28)	-0.53, 0.60	0.14	.890	16 (57) 29 (90)	-	-
Organisation / representation	-0.09 (0.21)	-0.54, 0.37	-0.40	.694	10 (30) 29 (90)	-	-
3. Outcome type						90.83	0.72
Retention / transfer	-0.09 (0.49)	-1.17, 0.99	-0.18	.860	30 (99) 9 (34)	-	-
Understanding / transfer	0.10 (0.56)	-1.09, 1.28	0.17	.864	14 (33) 9 (34)	-	-

Note. the reference for each comparison is stated as the second group level. For example, the mean difference (0.37) for signalling / no signalling indicates that dual coding effects were on average larger with signalling compared to no signalling. *k* = studies, *n* = outcomes.

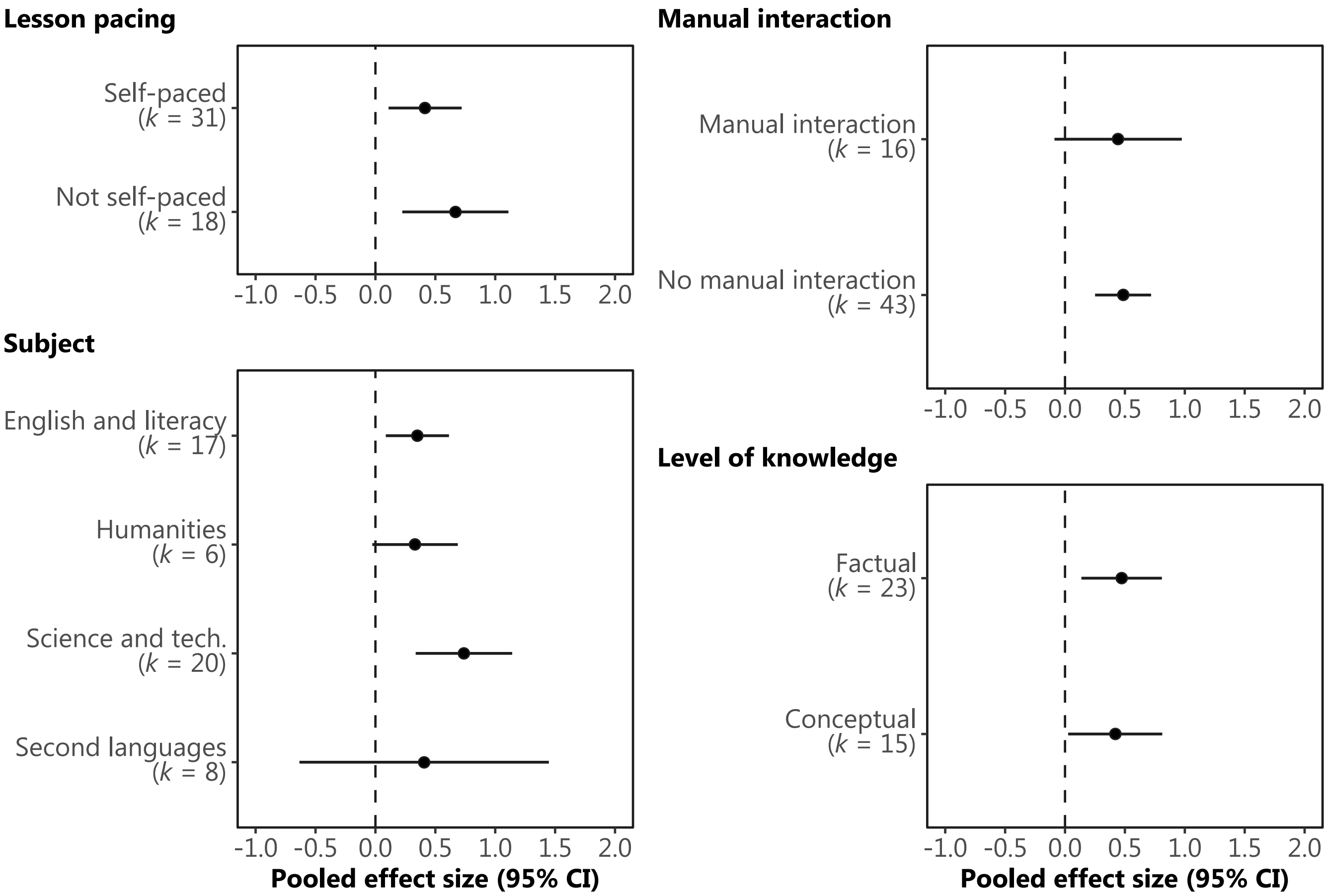
Moderators 4-7: lesson pacing, subject, level of knowledge, and manual interaction with visual

Three further moderators relating to instructional design potentially contributed to the heterogeneity of effects (see Figure 3). Firstly, dual coding had a greater effect when content was not self-paced ($g = 0.67$, $p = .006$), compared to when it was self-paced ($g = 0.42$, $p = .009$). Although this was a substantial difference on average, it was not statistically significant (see Table 5). Effects also varied for different subjects. Dual coding was the most effective for science and technology, and the estimate was statistically significant ($g = 0.74$, $p = .001$). Dual coding was also significantly effective for English and literacy, but the effect was smaller ($g = 0.35$, $p = .012$). Dual coding had a similar positive effect for humanities subjects (history, geography, social studies), but the estimate was not significant and so the impact for these subjects is less certain ($g = 0.33$, $p = .063$). The estimate for second language studies was positive ($g = 0.41$, $p = .385$) but had very large confidence intervals including negative and positive estimates, indicating that effects are very uncertain. We return to this in the discussion. The difference in effects between science and all other subjects was not statistically significant (see Table 5).

Dual coding had a positive effect on learning when pupils interacted with the visual by performing a manual action such as copying ($g = 0.44$, $p = .095$), and when they did not interact with the visual ($g = 0.49$, $p < .001$). However, the estimate for manual interaction was not statistically significant and was less precise (CIs were wider), indicating that effects are less certain. There was no statistical difference in effect sizes between subgroup levels.

Dual coding was effective whether pupils were learning factual knowledge ($g = 0.48$, $p = .008$) or conceptual knowledge ($g = 0.42$, $p = .038$), and effects were significant for both types of knowledge. This moderator is therefore unlikely to contribute to heterogeneity of effects. However, there were too few studies to draw reliable conclusions about procedural knowledge ($k = 2$).

Figure 3. Subgroup analyses of moderators 4-7 relating to instructional design



Note. *k* = number of studies. The total number of studies for each analysis may not equal 56 as we were unable to code some outcomes (for example, if not reported clearly). Effects are not statistically significant if 95% CIs cross 0 (dashed line).

Table 5. Subgroup analyses for moderators 4-7 relating to instructional design, and moderators 8 and 10 relating to pupil characteristics.

Moderator	Mean difference in effect *(SE)*	95% CI	*t* (df)	*p*	*k* (*n*)	I^2	τ^2
4. Lesson pacing						90.04	0.69
Not self-paced / self-paced	0.25 (0.26)	-0.27, 0.78	0.98	.333	18 (43) 31 (113)	-	-
5. Subject						88.83	0.58
English and literacy / Science and tech.	-0.39 (0.23)	-0.85, 0.08	-1.70	.099	17 (71) 20 (64)	-	-
Humanities / Science and tech.	-0.41 (0.24)	-0.95, 0.14	-1.72	.122	6 (20) 20 (64)	-	-
Second languages / Science and tech.	-0.33 (0.48)	-1.37, 0.71	-0.69	.501	8 (17) 20 (64)	-	-
6. Level of knowledge						88.10	0.51
Conceptual / factual	-0.05 (0.24)	-0.55, 0.45	-0.22	.827	15 (42) 23 (74)	-	-
7. Manual interaction with visual						89.43	0.61
Interaction / No interaction	-0.04 (0.27)	-0.61, 0.53	-0.15	.882	16 (49) 43 (138)	-	-
8. Education phase						89.02	0.57
Secondary / primary	0.52 (0.20)	0.12, 0.92	2.62	.011*	28 (93) 29 (96)	-	-
10. Pupil population						88.80	0.55
Targeted / broad	-0.19 (0.27)	-0.75, 0.36	-0.72	.479	15 (38) 40 (147)	-	-

Note. the reference level for each comparison is stated second. For example, the mean difference (0.25) for lesson pacing indicates that on average dual coding effects were larger for self-paced interventions compared to not self-paced.

We were not able to analyse moderator 9 (prior knowledge). *k* = studies, *n* = outcomes. * Significant at $p < .05$.

Subgroup analyses: pupil characteristics

Moderator 8: education phase

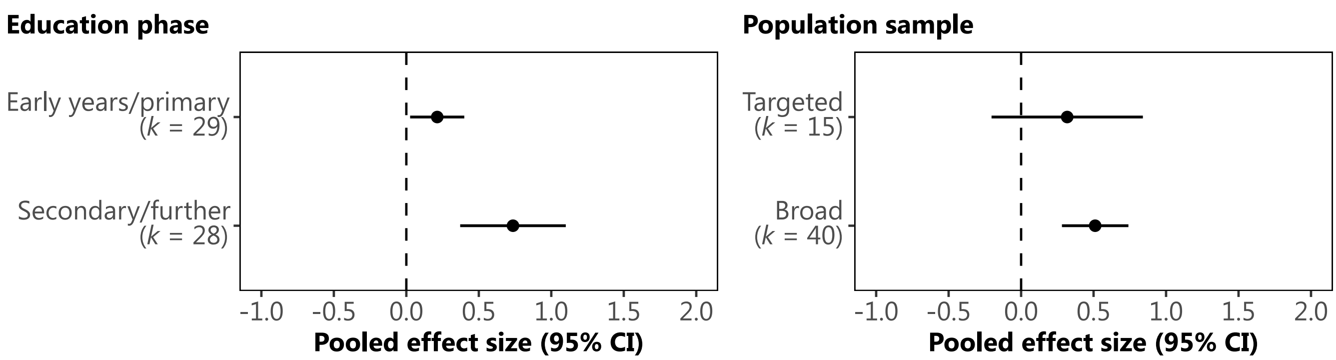
Visuals had a significant, positive effect on pupil learning regardless of their learning phase (see Figure 4). However, visuals were more effective for older pupils in secondary and further education (approximately 12-18 years old; $g = 0.74$, $p = .001$) compared to younger pupils in early years and primary (approximately 4-11 years old; $g = 0.21$, $p = .027$). The large difference in effects was statistically significant (see Table 5), indicating that this moderator significantly contributed to heterogeneity of effects.

Moderators 9 and 10: prior knowledge and pupil population

Apart from education phase, we were unable to identify any pupil characteristics that contributed to heterogeneity. First, we were unable to analyse pupils' prior knowledge as a moderator for several reasons. Although prior knowledge was measured in 28 studies, the measures that were used varied (examples include: vocabulary tests, standardised tests for reading or maths, or researcher-developed measures). This meant that we did not have a suitable continuous outcome variable for a sufficient number of studies, but also we could not easily classify pupils' prior knowledge as 'low' or 'high'. Only two studies classified student's prior knowledge, and both excluded pupils with high prior knowledge rather than comparing it to lower prior knowledge (Samuels, 1967; Stebner et al., 2017).

Second, there were too few studies to analyse any specific population group such as pupils with and without special educational needs and disabilities (SEND). Visuals had a positive and statistically significant effect on learning for studies that included a broad pupil population ($g = 0.51$, $p < .001$; see Figure 4); for example, pupils from multiple classes in a year group. However, effects were smaller and non-significant for more targeted pupil populations, for example pupils with learning difficulties or low attainment in a specific subject ($g = 0.32$, $p = .212$). This difference was not statistically significant (see Table 5). We return to this moderator in the Discussion.

Figure 4. Subgroup analyses of moderators relating to pupil characteristics



Note. k = number of studies. The total number of studies for each analysis may not equal 56 as we were unable to code some outcomes (for example, if not reported clearly). Effects are not statistically significant if 95% CIs cross 0 (dashed line).

Subgroup analyses: research design

Moderators 11-17: instructor, teaching group size, type of content, outcome measure/intervention alignment, timing of outcome measure, number of lessons, learning method

We analysed seven moderators relating to research design (see Figure 5) and found that most of them potentially contributed to heterogeneity of effects, but to different degrees and with different levels of certainty.

Firstly, when dual coding was part of whole-class teaching it was substantially more effective ($g = 0.81$, $p = .004$) than in interventions for individual pupils ($g = 0.50$, $p = .009$) or for small-group teaching ($g = 0.22$, $p = .116$). This difference was statistically significant (see Table 6), and indicates that teaching group size significantly contributed to heterogeneity of effects.

The timing of the outcome measure also significantly contributed to heterogeneity. When pupils were assessed immediately or up to one week after the intervention, dual coding had a positive and significant effect on learning (immediate: $g = 0.53$, $p < .001$; 1-7 days: $g = 0.45$, $p = .002$), but it had no effect when pupils were assessed two weeks or more after the intervention ($g = 0$, $p = .997$)[4]. The difference between these long term and short term outcomes was statistically significant (see Table 6). However, note that fewer studies assessed long-term effects (seven studies compared to 49 that measured learning immediately).

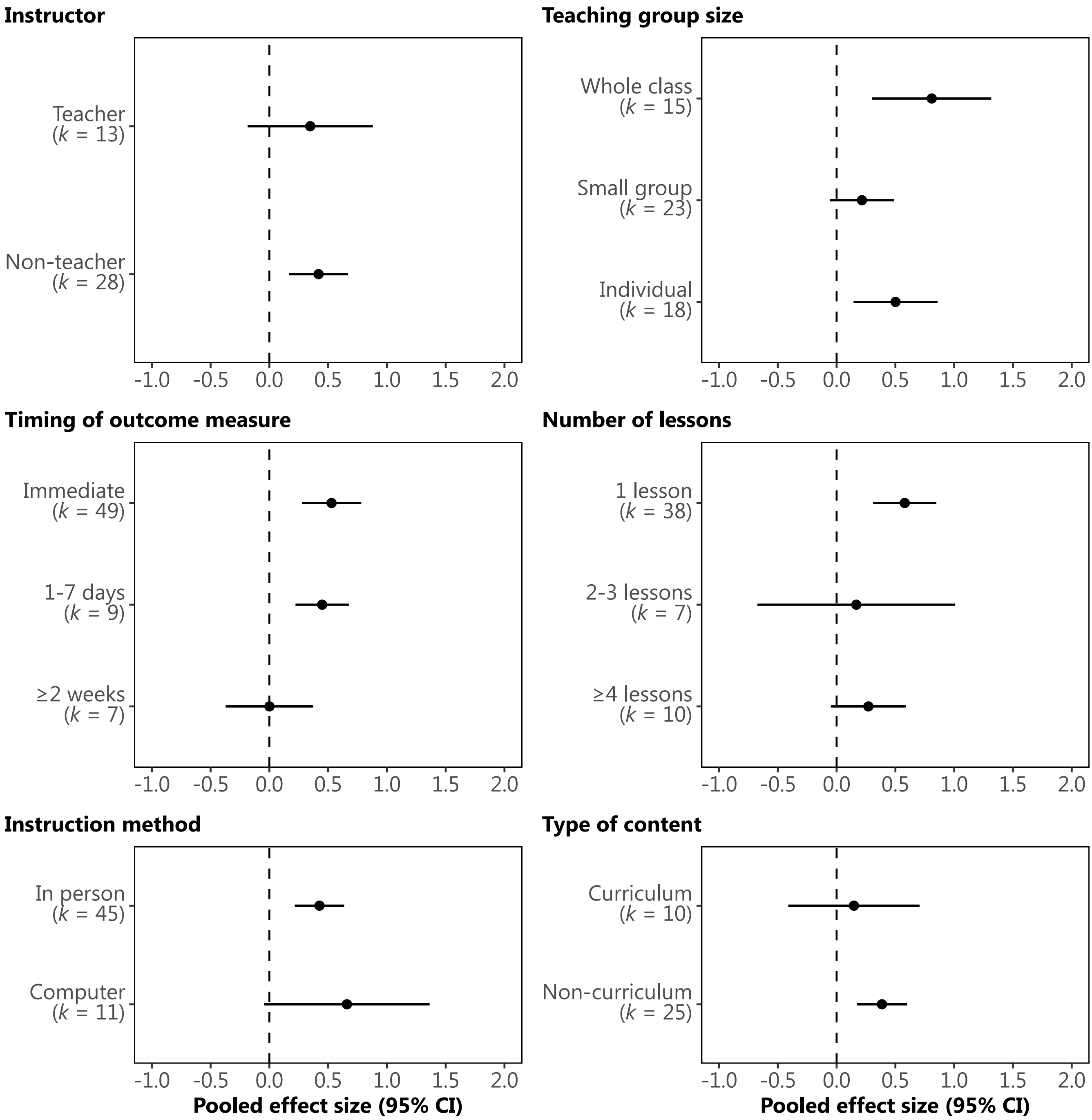
Visuals had a greater impact on interventions that only lasted one lesson ($g = 0.58$, $p < .001$) compared to interventions that lasted for four or more lessons ($g = 0.27$, $p = .088$). Interventions that lasted two to three lessons also had a positive impact, but the estimate was smaller and not statistically significant ($g = 0.17$, $p = .644$). Differences in these effects (that is, between longer and shorter interventions) were not statistically significant (see Table 6), so the impact of this moderator on heterogeneity is less certain.
Dual coding was, on average, more effective for computer-based interventions ($g = 0.66$, $p = .063$) compared to interventions carried out in person ($g = 0.43$, $p < .001$). However, the estimate for computer-based interventions was not statistically significant, meaning the impact of computer-based interventions was less certain. The difference between effect sizes for this moderator was not statistically significant (see Table 6).

Dual coding had a greater effect for content that was not part of the curriculum ($g = 0.39$, $p = .001$), compared to content that was ($g = 0.15$, $p = .565$). The latter estimate was very imprecise and CIs included negative values, indicating uncertain impact and potentially negative effects on pupil learning. The difference in effect sizes was not significant, although this moderator did slightly reduce between-study variation ($\tau^2 = 0.41$; see Table 6) so may still have contributed to heterogeneity.

Average effects were similar whether the intervention was conducted by the class teacher ($g = 0.35$, $p = .180$) or a researcher ($g = 0.42$, $p = .002$). This difference was not statistically significant. However, the estimate for interventions conducted by teachers was also not significant, meaning that the impact of dual coding when used by teachers was less certain. Lastly, almost all studies used outcome measures that were closely aligned with the intervention ($k = 53$), so no conclusions could be drawn from a subgroup analysis comparing broadly and narrowly aligned outcomes.

4 To examine this finding in more detail, we conducted an exploratory meta-regression using a continuous measure (the number of days the outcome was measured after the intervention; see supplementary analyses on the OSF). However, as the covariate was very imbalanced, with most outcomes being measured 0 days after the intervention, degrees of freedom were too low for results to be reliable.

Figure 5. Subgroup analyses of moderators 11-17 relating to research design



Note. *k* = number of studies. The total number of studies for each analysis may not equal 56 as we were unable to code some outcomes (for example, if not reported clearly). Effects are not statistically significant if 95% CIs cross 0 (dashed line).

Table 6. Subgroup analyses of moderators 11-17 relating to research design

Moderator	Mean difference in effect (*SE*)	95% *CI*	*t*	*p*	*k* (*n*)	I^2	τ^2
11. Instructor						88.33	0.50
Non-teacher / teacher	0.07 (0.27)	-0.49, 0.63	0.26	.797	28 (85) 13 (42)	-	-
12. Teaching group size						88.77	0.56
Small group / whole class	-0.59 (0.27)	-1.15, -0.04	-2.20	.036*	23 (88) 15 (58)	-	-
Individual / whole class	-0.31 (0.29)	-0.90, 0.29	-1.06	.298	18 (43) 15 (58)	-	-
13. Type of content						86.36	0.41
Non-curriculum / curriculum	0.24 (0.27)	-0.33, 0.80	0.89	.385	25 (95) 10 (31)	-	-
15. Timing of outcome measure						89.14	0.59
Immediate / 2 or more weeks	0.53 (0.19)	0.07, 0.98	2.75	.029*	49 (138) 7 (27)	-	-
1-7 days / 2 or more weeks	0.45 (0.19)	0.04, 0.85	2.40	.033*	9 (22) 7 (27)	-	-
16. Number of lessons						88.69	0.59
2-3 lessons / 1 lesson	-0.41 (0.37)	-1.25, 0.43	-1.12	.294	7 (15) 38 (141)	-	-
≥ 4 lessons / 1 lesson	-0.31 (0.19)	-0.73, 0.11	-1.60	.133	10 (31) 38 (141)	-	-
17. Learning method						89.21	0.58
Computer / in person teaching	0.23 (0.33)	-0.47, 0.94	0.70	.493	11 (41) 45 (148)	-	-

Note. The reference level for each comparison is stated second. For example, the mean difference (0.24) for type of content indicates that on average, dual coding effects were larger for non-curriculum content compared to curriculum content.

We did not analyse moderator 14 as there were too few studies (see Subgroup analyses: research design).
k = studies, *n* = outcomes. * Significant at $p < .05$.

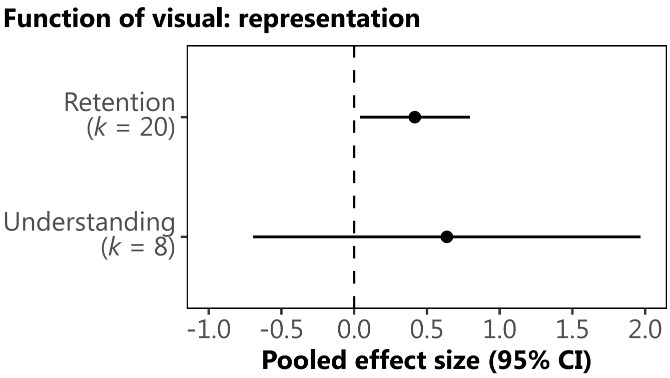
Interactions between moderators

We hypothesised that the function of visuals, the CTML principles and pupils' education phase would all interact with learning outcomes (see Aims and hypotheses). We planned to conduct a full meta-regression of our predicted interactions, but this was not possible due to the small number of studies in each subcategory ($k < 10$). We therefore carried out subgroup analyses to investigate differences in learning outcomes for each of the three moderators. The exception to this was the summary variable of CTML principles which, as a continuous variable, was suitable for meta-regression. Some subgroup analyses resulted in too few degrees of freedom for the estimates to be reliable (see Analysis). As before, we only report results that had reliable estimates and confidence intervals, but full analytical output is available on the OSF.

Function of visual x outcome measure

For the function of visuals, the only subcategory with a sufficient number of studies was representation (k = 28), and only for the outcome measures of retention and understanding. We broadly observed the same pattern of results for this analysis (Figure 6) as when we analysed outcomes across all functions (see Figure 2): representative visuals had a greater effect on understanding ($g = 0.64$, $p = .288$) than retention ($g = 0.42$, $p = .032$). However, the average effect for understanding was very imprecise, not statistically significant and included negative values. This might indicate that representative visuals could sometimes be harmful to pupils' understanding, but the smaller number of studies for this subgroup means that this is not certain. Furthermore, the difference in effects between retention and understanding was not statistically significant (see Table 7).

Figure 6. Subgroup analysis of outcome measures for representative visuals



Note. k = number studies. Analyses of the functions interpretation and organisation are not plotted (results were unreliable), but are available on the OSF.

Table 7. Interaction analyses between outcome measures and the function of visuals, CTML principles and education phase

Moderator	Mean difference in effect (*SE*)	95% *CI*	*t*	*p*	*k* (*n*)	I^2	τ^2
Function of visual						91.42	0.79
Understanding / retention	0.22 (0.57)	-1.05, 1.48	0.39	.708	8 (14) 20 (66)	-	-
CTML: sum variable						90.63	0.71
Understanding / retention	-0.04 (0.18)	-0.45, 0.38	-0.21	.838	14 (33), 30 (99)	-	-
Transfer / retention	0.38 (0.38)	-0.53, 1.29	1.01	.351	9 (34) 30 (99)	-	-
CTML: coherence						91.96	0.87
Understanding / retention	0.21 (0.34)	-0.51, 0.93	0.62	.545	12 (29) 25 (83)	-	-
Transfer / retention	0.28 (0.65)	-1.21, 1.76	0.42	.683	7 (29) 25 (83)	-	-
CTML: contiguity						92.54	0.87
Understanding / retention	0.03 (0.41)	-0.84, 0.90	0.07	.947	10 (17) 22 (68)	-	-
Transfer / retention	0.13 (0.64)	-1.34, 1.60	0.20	.846	6 (27) 22 (68)	-	-
CTML: segmentation						90.58	0.75
Understanding / retention	0.56 (0.38)	-0.29, 1.41	1.48	.173	7 (14) 20 (53)	-	-
CTML: signalling						92.54	0.87
Understanding / retention	0.03 (0.41)	-0.84, 0.90	0.07	.947	10 (24) 12 (27)	-	-
Transfer / retention	0.13 (0.64)	-1.34, 1.60	0.20	.846	7 (24) 12 (27)	-	-
Education Phase: early years and primary						73.57	0.22
Retention / transfer	0.36 (0.15)	0.02, 0.71	2.45	.042*	18 (46) 6 (25)	-	-
Education Phase: secondary and further						93.90	1.07
Understanding / retention	-0.08 (0.42)	-0.97, 0.80	-0.20	.843	10 (19) 13 (53)		

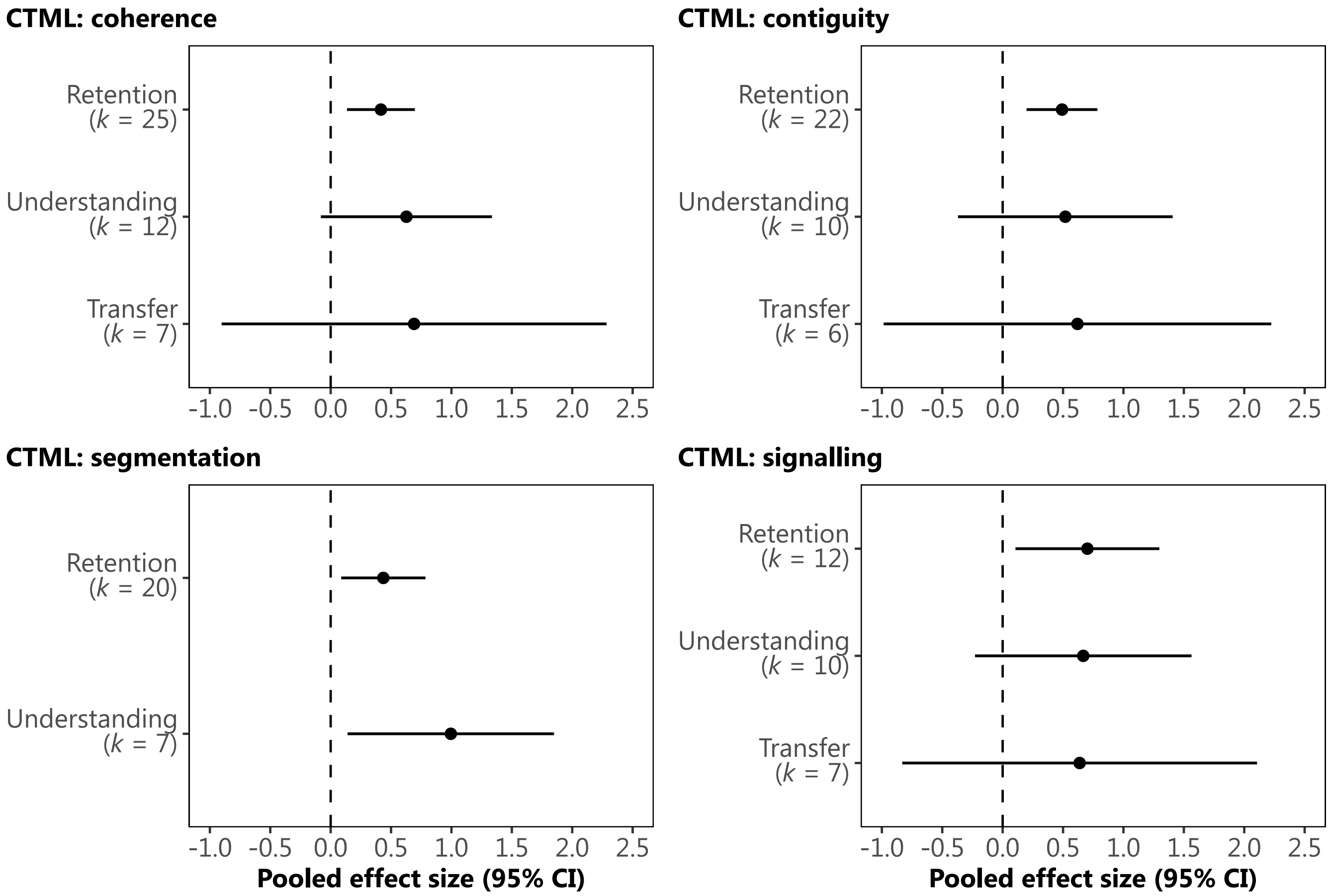
Note. The reference level for each comparison is stated second. For example, for early years and primary pupils, the mean difference (0.36) indicates that on average, dual coding effects were larger when the outcome was retention compared to transfer.

k = studies, n = outcomes. * Significant at $p < .05$.

CTML principles x outcome measure

Meta-regression of the summary variable (see Table 7) indicated that dual coding had a greater impact on transfer compared to retention when more principles of the CTML were followed (g_{diff} = 0.38, p = .351). However, this difference was not statistically significant and so the interaction is uncertain. Following more principles did not result in greater understanding compared to retention (g_{diff} = -0.04, p = .838). Subgroup analysis of each individual principle (see Figure 7) indicated that coherence, contiguity and signalling all had a similar positive effect on retention, understanding and transfer, but the impact on understanding and transfer was much less precise (and not statistically significant). Note that this is the same pattern when all outcome measures are grouped – see Figure 2. In contrast, when learning content was segmented, dual coding had a much greater impact on pupil understanding, although this difference was not statistically significant (g_{diff} = 0.56, p = .173). This suggests a possible interaction between this principle and learning outcomes; that is, segmenting content may increase the impact of dual coding on pupils' understanding, but it is uncertain.

Figure 7. Subgroup analysis of outcome measures for each CTML principle

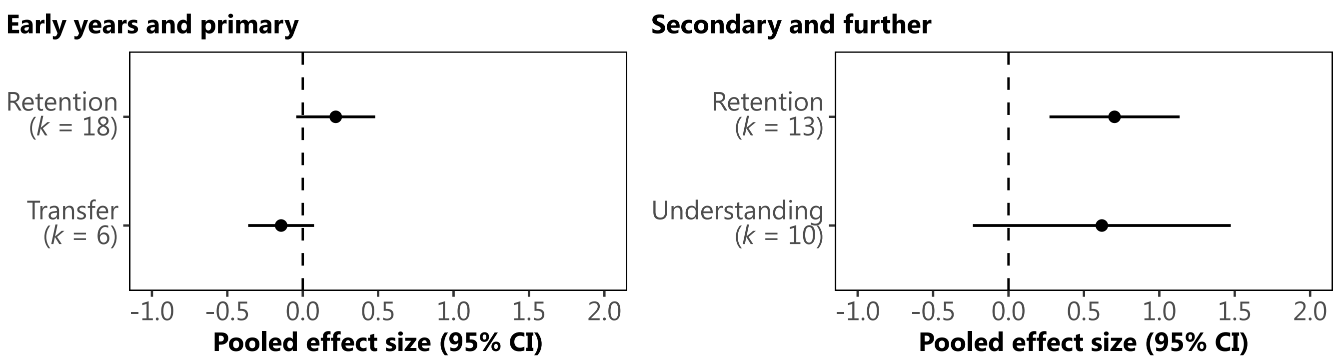


Note. k = number studies. The estimate for transfer is not reported for CTML: segmentation as there were too few studies to be reliable.

Education phase x outcome measure

For early years and primary phases, there were too few studies that measured understanding. For secondary and further education, there were too few studies that measured transfer. This meant that we could not reliably compare effects for these outcome measures in younger and older pupils. However, we found a statistically significant difference between retention and transfer for early years and primary studies (g_{diff} = 0.36, p = .042), and heterogeneity was reduced when both moderators were included in the regression model ($I^2 = 74$, $\tau^2 = 0.22$; see Table 7). These results indicate a significant interaction between education phase and learning outcomes. Specifically, the positive estimate for retention suggests that dual coding positively impacts younger pupils' retention, while the small negative estimate for transfer suggests that that at best, dual coding would likely have no impact on primary pupils' transfer (see Figure 8). We also observed that dual coding had a much larger and statistically significant impact on retention for pupils in secondary and further education (g = 0.70, p = .004), but a smaller and less certain (statistically non-significant) impact for early years and primary pupils (g = 0.22, p = .096).

Figure 8. Subgroup analysis of outcome measures for different education phases



Note. k = number of studies. As there were too few studies to be reliable, the estimates for understanding (for early years and primary) and transfer (for secondary and further) are not plotted.

Exploratory analyses

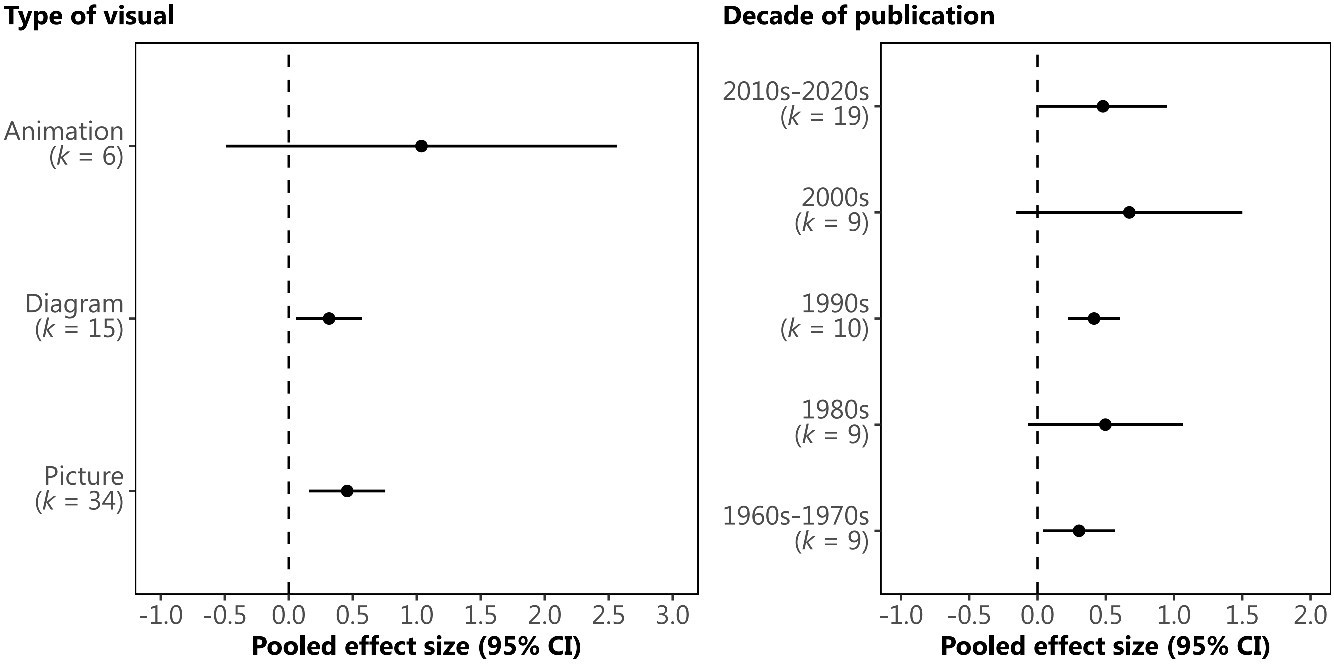
Type of visual

We did not have a priori hypotheses regarding the type of visual (such as picture or animation) and its contribution to heterogeneity. Nevertheless, we recorded the type of visual used in each intervention and explored differences between them through subgroup analysis. Pictures ($g = 0.46$, $p = .004$) and diagrams ($g = 0.32$, $p = .021$) had a positive and significant effect on learning (see Figure 9). Animations had a large positive effect on learning ($g = 1.04$, $p = .014$) but this was statistically not significant, and the wide confidence intervals suggest that their effect is uncertain. There were too few studies to make reliable conclusions about graphic organisers, concept maps or photos.

Decade

The studies included in our analysis spanned research that took place over 70 years – from the 1960s to the 2020s. We examined differences between decades as changes in research methods and publication standards could have impacted the size of effects. We grouped studies according to the decade of publication and if there were not enough studies for a particular decade, we grouped them with the subsequent or preceding decade. Effect sizes were positive for all decades, but not all of them were significant (see Figure 9). Particularly, the impact of visuals was non-significant (and therefore less certain) from research published in the 1980s ($g = 0.50$, $p = .078$) and in the 2000s (g = 0.67, $p = .098$). Effects from research conducted in the 1960s-1970s were noticeably smaller ($g = 0.30$, $p = .028$), but this estimate was statistically significant and effects are therefore more certain. The decade of publication (and quality of research) could potentially have contributed to heterogeneous effects, but we did not observe a clear pattern of effect sizes increasing or decreasing over time.

Figure 9. Exploratory subgroup analyses of type of visual and decade of publication

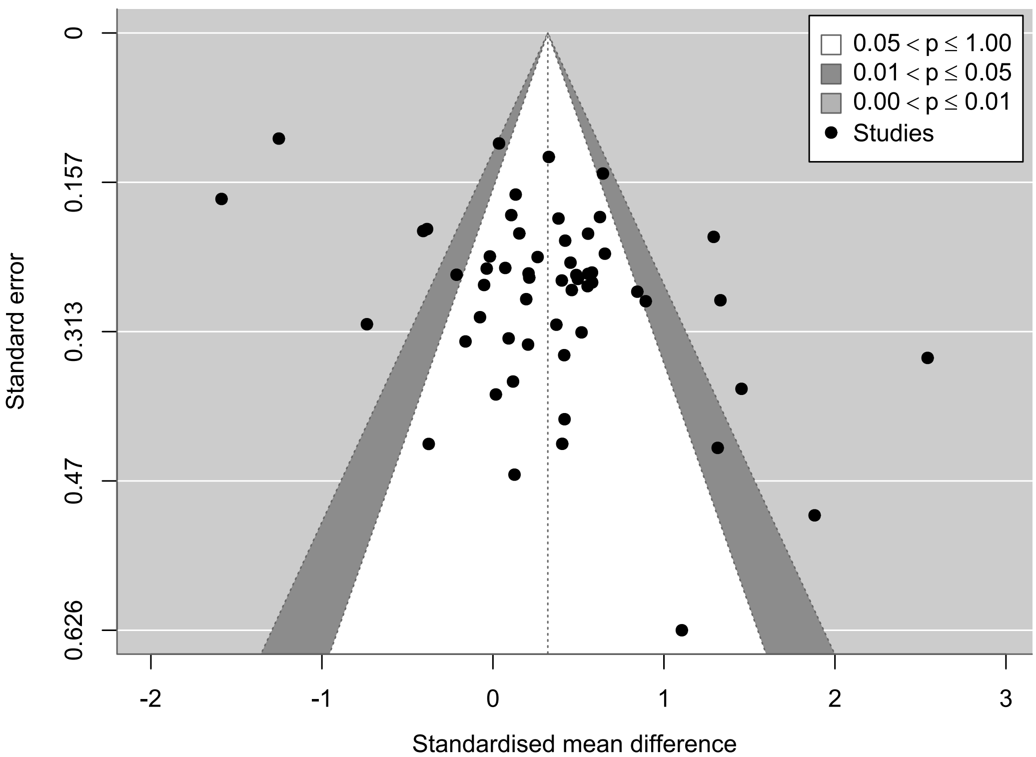


Note. *k* = number of studies per subgroup level. The total number of studies for each analysis may not equal 56 as we were unable to code some outcomes (for example, if not reported clearly). Effects are not statistically significant if 95% CIs cross 0 (dashed line).

Publication bias

A contour-enhanced funnel plot (see Figure 10) showed a fairly symmetrical pattern with a large number of non-significant effects, both positive and negative. The plot suggests a minimal presence of publication bias due to small studies (that is, smaller studies with high standard errors were not less likely to be published or identified in our literature search). This was supported by the trim-and-fill analysis which identified no 'missing' studies due to small-study effects (that is, non-significant results from smaller studies were not less likely to be published). The analysis did not impute additional effect sizes, and the pooled estimate and 95% CIs were identical to the uncorrected estimate for aggregated effects ($g = 0.32$, $SE = 0.09$, 95% CIs = [0.15, 0.49]. *P*-curve analysis of $k = 24$ studies with $p < .05$ did not indicate evidence of motivated *p*-hacking – that is, our analysis had evidential value, and there was likely a true non-zero effect of visuals on pupil learning. This was indicated by significant tests for right skewness (all $ps < .001$) and non-significant tests of flatness (all $ps > .94$; see plot and full results on the OSF). A selection model with 'steps' (significance thresholds) at $p < .05$ and $p < .10$ indicated a small amount of publication bias. The model showed that publication was slightly less likely for effects with *p*-values >.10 (likelihood = 0.95, $k = 30$) compared to effects with *p*-values < .05 (likelihood = 1.0, $k = 25$). Effects on the threshold of significance ($.10 > p > .05$) appeared much less likely to be published, but this only applied to one study (likelihood = 0.27, $k = 1$). The estimated average effect from the selection model ($g = 0.31$, $SE = 0.14$, 95% CIs [0.03, 0.59] was slightly smaller than the uncorrected estimate for aggregated effects, and 95% CIs were wider with the lower bound very close to zero. This finding indicates that taking publication bias into account would result in a less precise estimate but would not substantially change our conclusions. Overall, the four tests do not indicate major concerns regarding publication bias in our data, although they cannot fully rule it out.

Figure 10. Contour enhanced funnel plot of effects aggregated at the study level



Note. The vertical dotted line indicates the estimated average of aggregated effects. The white area of the plot indicates non-significance at the standard α-level of .05.

Discussion

In our meta-analysis of 56 studies and 189 outcomes, we found that on average, dual coding had a moderate, positive and significant effect on learning outcomes. This means that previous findings from lab experiments do, in general, translate to classroom settings. However, our analyses revealed high levels of heterogeneity. We identified several moderators relating to instructional design, pupil age and research methodology which potentially contributed to this heterogeneity. In particular, the education phase of pupils, teaching group size, length of time between the intervention and outcome measure, and the type of learning outcome, all significantly explained some of the heterogeneity. Nevertheless, it remained high even when accounting for differences in subgroups. This means that the impact of dual coding in a classroom setting can vary considerably, and could in some cases have no impact on learning, or even a negative impact (see also Perry et al., 2021). The moderators we identified relating to research methodology suggest that the estimated benefit from dual coding may, in more ecologically valid conditions, be smaller than our overall estimate. Teachers should carefully consider several factors when deciding when and how to use visuals in their teaching, to maximise benefits and prevent them from having a negative effect on pupil learning.

When to use dual coding

We found that the impact of dual coding will likely differ depending on the learning outcome of a lesson or intervention. Firstly, we found that dual coding on average has a positive impact on pupils' retention. This is in line with DCT which predicts that when visuals accompany written or verbal content, they support the encoding of concepts in memory by providing an additional 'code' to strengthen mental representations. However, it is less effective for early years and primary pupils compared to secondary pupils, and we found limited evidence that dual coding benefits long-term retention (we return to this later).

Secondly, we found that dual coding benefits pupils' understanding, although effects are less certain and may require more careful design of materials. For example, segmenting content into meaningful units may increase the impact of visuals on pupils' understanding. Visuals that literally represent written or verbal content (representation function) also have a less certain impact on understanding (see Figure 6), perhaps because they are somewhat redundant or can act as a distraction.

Thirdly, we found that dual coding has a less reliable impact on pupils' transfer and could potentially have a negative impact on this outcome, particularly for younger (early years and primary) pupils. This may be due to the more complex cognitive processes required to integrate visual information alongside written or verbal text but could also be due to limited long-term benefits. A limitation to these findings is that studies measuring understanding and transfer used a greater variety of tests and measures than retention (which is a more clearly defined outcome and simpler to measure). The variety of outcome measures used by researchers could potentially explain the non-significant results for understanding and transfer (as well as the confound between transfer and long-term effects). Indeed, outcomes for just over 20% of studies could not easily fit into one of the three learning outcome categories.

Effects for all outcome measures were only present up to one week following the lesson or intervention, and were negligible for longer time periods. This suggests that dual coding, when used during a lesson or intervention, is possibly only beneficial in the short term. However, only seven studies measured long-term impacts, and most of these measured transfer; it is therefore unclear whether this finding is due to the type or timing of the outcome measure. The finding may also be due to the way in which dual coding was used in the studies we reviewed. As most of the interventions were researcher-designed and not part of the normal class curriculum, dual coding may have been used intentionally to benefit immediate learning or understanding rather than having longer term learning as a goal. The long-term impact of dual coding therefore requires much more evidence from valid classroom settings.

We predicted that dual coding would improve learning in all subjects, regardless of education level, but that it would be most beneficial in maths (which involves visuo-spatial processing), and least beneficial for learning second languages (which can be achieved by forming connections with a speaker's first language). We did not retrieve enough maths studies to analyse this subject reliably, but we found that visuals are reliably effective when used in science and technology, and English and literacy, with a greater impact for science and technology. The impact on humanities subjects was also positive but less certain, potentially due to the small number of studies in this category. Examining the topics for each subject in more detail, we did not observe any clear patterns in terms of within-subject variation. Topics for science and technology were very broad and covered biology, physics and chemistry, and a variety of concepts, processes, theory and facts. Humanities (often termed social studies in the USA) mostly involved history as well as two studies of economics, while the English and literacy category mostly comprised reading comprehension. A wider variety of studies on humanities/social science and different aspects of English may indicate more precisely when dual coding can be beneficial.

Most notable was the very imprecise estimate for teaching second languages. The eight studies we reviewed for this subject took place in a variety of countries, involved foreign and native language learning (for example, English as a second language in the USA), and a variety of language skills (vocabulary, reading and listening comprehension). There was no identifiable pattern to clearly explain the imprecise estimate, except for one very small study (N = 11) with very large effects (Barnitz & Speaker Jr, 1991), which would have been down-weighted in the meta-analysis. This finding suggests our hypothesis may be correct, and that dual coding has a less reliable impact on second language learning. However, more evidence is required to reach a firm conclusion and to understand why it may be less effective for this subject.

Regardless of the subject, we did not find substantial differences for the type of knowledge pupils were learning, whether conceptual (such as schemas, mental models, principles or theories) or factual (such as dates or vocabulary). This suggests that teachers can use dual coding to effectively teach both types of knowledge.

Dual coding was more effective when used as part of whole-class teaching rather than with small groups or individual pupils. We had predicted the opposite, as small-group teaching is known to be more effective (EEF, 2021). The reason for this result is unclear, but suggests (along with the finding that visuals were effective for broad pupil populations) that dual coding may be effective as part of the core offer for all pupils during class teaching, rather than as an additional resource provided only for some (for example, to provide additional scaffolding for individual pupils).

Dual coding had a greater impact on pupils' learning when lessons were not self-paced (for example, when the pacing of a lesson was set by the teacher or had a time limit). This is in line with previous meta-analyses which observed that certain CTML principles have a greater impact on learning for system-paced lessons (Ginns, 2005; Rey, 2012). Our result suggests that visuals can particularly help pupils learn when they have limited time to process the lesson content, and may give pupils a useful shortcut to help them process and learn information more efficiently.

Summary: when to use dual coding

To get the greatest benefit from visuals alongside verbal or written content, teachers could consider using them more often in these contexts:

> To improve pupils' retention and understanding in the short term.
> For science and technology, English and literacy (including reading comprehension).
> To teach factual or conceptual knowledge.
> In whole-class teaching for all pupils.
> For teacher-paced or timed activities.

Which pupils benefit most from dual coding?

As we predicted, dual coding improved outcomes for pupils of all ages regardless of their education phase, but we also observed differences between phases, and a significant interaction between education phase and learning outcome. Secondary-age pupils benefitted far more from visuals than primary-age pupils overall, on average gaining more than double the impact, but particularly when retention was the learning outcome. Dual coding also had no impact on primary pupils' transfer (although there were too few studies to compare this to secondary pupils). This has implications for primary teachers, who may wish to more carefully consider how and when dual coding can benefit their pupils' learning.

Previous meta-analyses have also observed smaller or non-significant effects for younger pupils (Guo et al., 2020; Schroeder & Cenkci, 2018). This difference may be because older pupils have better active learning skills such as the ability to integrate visual and linguistic content successfully. It may also reflect differences in the way visuals are used for different age groups. In our data, over half of studies of primary pupils used visuals that were literal representations of content (18 out of 30 studies). For secondary pupils, functions were more varied with an equal number of representation (11 studies out of 31) and interpretation (12 studies). This does suggest that visuals are used for different purposes in primary and secondary education. Nevertheless, effects were very similar for these different functions in our analysis, so this can't fully explain the large difference in age groups. The implication of this finding is that teachers of pupils aged 12-18 could consider using visuals more often to support learning, while primary teachers should be aware that it could potentially have a negative impact on pupils' longer term learning (transfer).

We predicted that effect sizes would be larger for studies that included a targeted subgroup of pupils compared to studies with broader inclusion criteria, due to less heterogeneity between pupils (Kraft, 2020). However, we found the opposite. This is likely because the 'targeted' studies were too varied. They included pupils with and without SEND, low and high attaining pupils, pupils with low prior knowledge, and pupils with English as a second language. It is unsurprising that the impact of visuals varied in these different groups, and unfortunately there were insufficient studies to analyse effects in any one subgroup. Particularly, we predicted that pupils with low prior knowledge would benefit more from visuals, while pupils with greater prior knowledge would benefit less (Perry et al., 2021), but we were unable to analyse this moderator despite its potential importance for adaptive teaching. More research is needed to explore which pupil populations will and will not benefit from dual coding, particularly for pupils with SEND when it may form part of their specialist support.

Summary: which pupils benefit most from dual coding?

We recommend that teachers use dual coding for all education phases and age groups, but it can particularly benefit older pupils aged 12-18. Primary teachers should carefully consider when to use dual coding as it could negatively impact younger pupils' transfer. There is currently a lack of evidence regarding the benefit of dual coding for pupils with particular characteristics, particularly pupils with SEND.

How to apply dual coding effectively

Our analyses indicate several ways that teachers can maximise the benefits pupils gain from dual coding. Firstly, by applying the most effective active learning principles when creating resources. The CTML outlines several key principles for how to create and use visuals effectively, but prior meta-analyses have always focused on evidence for a single principle (Alpizar et al., 2020; Ginns, 2005, 2006; Reinwein, 2012; Rey, 2012; Richter et al., 2016; Schneider et al., 2018; Schroeder & Cenkci, 2018; Sundararajan & Adesope, 2020). We considered several of the learning principles at once, allowing us to compare which principles have the greatest benefit to pupil learning. Our results largely support the use of the CTML principles, but we found that their impact varies, and including more principles in general may only lead to small benefits for pupils. Instead, we found that three principles in particular are effective when applied to dual coding, and the most important is signalling (also known as cueing). Signalling directs learners' attention to relevant information to help them integrate visuals with verbal or written content. For example, a teacher might verbally refer to and point to a picture while explaining a concept, or a diagram might have annotations linking it to a written explanation. Compared to the other principles, signalling had a particularly large impact on pupils' retention, and overall, learning content that included signalling had more than double the impact of content without it.

We also found that segmenting content is important for pupils' understanding. This means presenting content in meaningful segments rather than one continuous unit; for example, presenting written text and pictures as chunks of content with separate paragraphs or headers, or spacing verbal content with pauses between segments. Segmentation likely helps manage pupils' cognitive load when processing additional visual content, helping them to connect visuals with the relevant text or verbal content without overloading their working memory.

Avoiding extraneous details that are irrelevant to the lesson content or learning objectives (the coherence principle; for example, decorative icons) also increased the impact of dual coding. Importantly, if lesson content is not coherent, this may result in visuals having a negative impact on learning. Coherent content likely helps pupils to focus their attention on only the most important information, allowing them to form and learn coherent mental models without being distracted, or straining their cognitive load.

Our findings suggest that teachers' time may not be best spent focusing on the contiguity principle, as it likely only has a small benefit to learning. This principle is intended to avoid splitting pupils' attention by ensuring that visual content is spatially integrated with written content (for example, aligning visuals with specific sections of text), or temporally integrated with verbal content. Lastly, we did not find enough studies to evaluate the verbal redundancy principle. Our results overall support the use of the active learning principles, but suggest that teachers prioritise signalling, segmentation and coherence.

Contrary to our predictions, the benefit of asking pupils to perform a manual interaction with visuals (for example, copying or labelling) is uncertain. Studies in this category involved a range of manual interactions such as copying icons and graphs, filling in diagrams and graphic organisers, and interactive software where pupils clicked on images and animations, or dragged shapes on screen. Out of 16 studies that included manual interactions, eight observed negative effects. This could partly be explained by a confound with outcome measure, as four out of the eight studies measured pupils' long-term transfer (and these studies contributed over half of all analysed outcomes in this category). Another explanation relates to the pedagogical skill involved in choosing when and how to use manual interaction alongside dual coding; that is, designing effective manual interactions may rely on teacher expertise. Future research could examine the types of manual interactions that are effective, especially as it is often recommended as part of dual coding in teaching practice (Carney and Levin, 2002; Caviglioli, 2019); based on our findings, we found limited evidence to support this recommendation.

Contrary to our predictions, the impact of computer-based interventions was less certain than in-person interventions, although on average they both had positive effects. It is important to note that these studies often involved pupils working independently on a computer, with a programme designed to teach a specific topic (for example, probability or vocabulary). The uncertain effect for computer-based interventions may therefore be due to less involvement and guidance from an instructor, but we cannot be certain.

Dual coding was most effective when used in just one lesson, indicating that visuals are perhaps most beneficial when focused on one learning activity or topic. They were less effective for longer interventions over several lessons, although this could be explained by the fewer number of studies in this category; most of the studies we reviewed investigated the application of dual coding in only one lesson.

Summary: how to apply dual coding effectively

To maximise the benefit from dual coding and to optimise time spent on preparing effective resources, we recommend teachers focus on the following:

> Include signalling to help pupils integrate visuals with verbal or written content.

> Segment content into meaningful, coherent sections to improve pupils' understanding.

> Ensure resources do not include irrelevant details.

> Use dual-coding during in-person teaching rather than solely relying on a computer-based intervention.

> Use visuals tailored to individual lessons or activities.

What types of visual are most effective?

Previous research suggested that the function of visuals may be an important moderator of dual coding (Carney & Levin, 2002; Perry et al., 2021). We found that visuals representing written or verbal content literally (for instance, a story illustration), visuals that organise content (for instance, a flow chart or timeline) or that interpret an idea or complex concept (for instance, a diagram of the circulatory system) can all improve pupils' learning. Visuals that interpreted an idea or concept had the greatest impact, and teachers may wish to consider using this type of visual more often. However, visuals that literally represented the verbal or written content did not always benefit pupils' understanding, suggesting that this function of visual is more effective for helping pupils retain information. Lastly, we did not find enough studies to evaluate the impact of mnemonic or decorative visuals. These two functions may explain more heterogeneity than the three we analysed; for example, mnemonic visuals have been shown to have the greatest impact and decorative visuals no impact (Levin et al., 1987).

In line with the review by Perry et al. (2021), most studies in our analysis used pictures or diagrams, both of which benefitted learning. Pictures had a greater impact than diagrams, while the impact of animations was positive but much less certain. It is not clear why, theoretically, there would be any difference in the effects of these three types of visual, except that diagrams and animations may be more complex and may need to be particularly well-designed for them to be effective. It should be noted that the number of studies in each category was very uneven (60% used pictures, about 27% used diagrams, and only 11% used animations), and that this analysis was exploratory, so results should be interpreted with caution. Previous research supports our finding that pictures and diagrams are reliably effective (see also Perry et al., 2021), but overall there is less evidence to make firm conclusions about the other media included in our search criteria (animations, videos, photos, concept maps and graphic organisers).

Summary: which visuals are the most effective for dual coding?

Pictures and diagrams are likely to be effective, but evidence is weaker or lacking for other types of media. To help pupils learn, teachers should consider applying dual coding using pictures or diagrams that:

> Interpret complex ideas or concepts (for example, diagrams).

> Accurately represent verbal or written content (for example, an accurate illustration), to help pupils retain information rather than understand it.

> Organise learning content (for example, flow charts).

Aspects of research design that contributed to heterogeneity

The way research studies were designed varied considerably, and several factors potentially contributed to the high heterogeneity of effects. These factors should be considered in any future research on dual coding in teaching, and when interpreting the present results, especially regarding the validity of the findings and the likely benefit pupils can gain from dual coding.

Firstly, the time interval between the intervention and outcome measure significantly contributed to heterogeneity of effects. We predicted that effects would be smaller for longer intervals compared to when outcomes were measured immediately. However, the positive benefits of dual coding disappeared with periods of two or more weeks between intervention and assessment. This calls into question the long-term benefits of dual coding, particularly in relation to transfer of learning. It indicates that visuals provide an immediate benefit to pupils' learning, but that this learning will need to be consolidated into long-term memory for the benefit to be maintained. It may also reflect that dual coding in these studies was used for an immediate purpose, and a very specific learning objective closely aligned with the intervention, especially as most of the materials were designed specifically for the research study. This area requires further research to understand the implications for teachers, and how they can enhance pupils' learning from dual coding in the long term.

Contrary to our predictions, the impact of dual coding was less certain when the intervention was taught by a class teacher compared to a researcher, and when content was taken directly from the normal class curriculum compared to when it was designed specifically for the research. Interventions taught by a researcher, or that used resources developed specifically for the research, may have been better controlled (for instance, implemented more consistently and accurately). This calls into question the validity of the results and suggests that the overall impact of visuals may be smaller and less reliable when applied in real-life teaching contexts rather than as part of an experimental research study.

Lastly, an exploratory analysis of the decade in which research was published indicated that this may have contributed to the heterogeneity of effects, although we did not make predictions about this moderator. We found that research published in the 1980s and 2000s was less reliable than other decades. It is not clear why these particular decades were different, but trends in research methods and publication standards could account for the differences.

Limitations

Despite analysing multiple moderators, we were not able to fully explain the observed heterogeneity in our meta-analysis, and heterogeneity remained high in all moderator analyses. This means that despite our recommendations, the impact of dual coding is still likely to vary when applied in the classroom. As with any systematic review, our analyses and the conclusions we can draw from them are limited to the literature we found in our searches, and the quality, amount and type of data that has already been published. This meant that some of the moderator analyses we had planned were not possible, for example examining all possible interactions between certain moderators. We could not draw conclusions about some moderator categories because not enough studies met our inclusion criteria (or not enough studies exist). For example, pupils' prior knowledge, decorative and mnemonic visuals, concept maps and graphic organisers, use of visuals in maths lessons, use of visuals to learn procedural knowledge, and verbal redundancy (one of the key principles from the CTML). Some subgroup analyses also had an uneven number of studies in each category, meaning that differences in effects (and particularly less precise or non-significant effects) could be related to a lack of studies in one category rather than the absence of a true effect. When interpreting the present results, it is important to note that a lack of evidence does not mean visuals are not effective, just that we cannot give reliable recommendations for these factors.

Another limitation of meta-analysis is that it does not indicate causation (that is, we cannot conclude that larger effects were definitively caused by the factors we analysed), even though we only reviewed randomised experimental designs which could determine causation. Lastly, conducting subgroup analyses requires studies to be categorised according to specific criteria, which risks some judgements being subjective (for example, the function of visuals). We mitigated this risk by double-screening and triple-coding studies to ensure consistency, but we cannot rule out that some of our decisions might be categorised differently if repeated by different researchers.

Finally, we agree with findings from the EEF (Perry et al., 2021) that much of the research we reviewed was less ecologically valid than we hoped. Most interventions were conducted by a researcher over just one lesson, with materials and topics that were not part of the normal curriculum, and with measures that would not necessarily be used in the lesson. This means that effects may be smaller in reality and highlights the need for better, more authentic research into dual coding in teaching.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Acha, J. (2009). The effectiveness of multimedia programmes in children's vocabulary learning. *British Journal of Educational Technology, 40*(1), 23–31.

Adesope, O., & Nesbit, J. (2011). Verbal Redundancy in Multimedia Learning Environments: A Meta-Analysis. *Journal of Educational Psychology, 104*, 250–263.

Allen, K., Higgins, S., & Adams, J. (2019). The Relationship between Visuospatial Working Memory and Mathematical Performance in School-Aged Children: A Systematic Review. *Educational Psychology Review, 31*(3), 509–531.

Alpizar, D., Adesope, O. O., & Wong, R. M. (2020). A meta-analysis of signaling principle in multimedia learning environments. *Educational Technology Research and Development, 68*(5), 2095–2119.

*Alvermann, D. E. (1981). The compensatory effect of graphic organizers on descriptive text. *Journal of Educational Research, 75*(1), 44–48.

*Alvermann, D. E., & Boothby, P. R. (1986). Children's Transfer of Graphic Organizer Instruction. *Reading Psychology, 7*(2), 87–100.

Anderson, L. W., & Krathwohl, D. R. (2001). *A Revision of Bloom's Taxonomy of Educational Objectives*. Addison Wesley Longman, Inc.

*Ardasheva, Y., Wang, Z., Roo, A. K., Adesope, O. O., & Morrison, J. A. (2018). Representation visuals' impacts on science interest and reading comprehension of adolescent English learners. *The Journal of Educational Research, 111*(5), 631–643.

*Babaei, S., & Izadpanah, S. (2019). Comparing the effects of different advance organizers on EFL learners' listening comprehension: *Key vocabularies, previewing comprehension questions, and multimedia annotations*. *Cogent Education, 6*(1), 1705666.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.

Baddeley, A., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press.

*Barnitz, J. G., & Speaker Jr, R. B. (1991). Second language readers' comprehension of a poem: *Exploring contextual and linguistic aspects. World Englishes, 10*(2), 197–209.

Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education, 101*, 150–167.

*Bobis, J., Sweller, J., & Cooper, M. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction, 3*(1), 1–21.

*Brandt, L., Elen, J., Hellemans, J., Heerman, L., Couwenberg, I., Volckaert, L., & Morisse, H. (2001). The impact of concept mapping and visualization on the learning of secondary school chemistry students. *International Journal of Science Education, 23*(12), 1303–1313.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review, 14*(No. 1), 5–26.
Caviglioli, O. (2019). *Dual Coding with Teachers.* Hachette UK.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

*ChanLin, L. (2001). Formats and prior knowledge on learning in a computer-based lesson. *Journal of Computer Assisted Learning, 17*(4), 409–419.

Clark, J. M., & Paivio, A. (1987). A Dual Coding Perspective on Encoding Processes. In M. A. McDaniel & M. Pressley (Eds), *Imagery and Related Mnemonic Processes: Theories, Individual Differences, and Applications (pp. 5–33)*. Springer.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149–210.

*Coleman, J. M., McTigue, E. M., & Dantzler, J. A. (2018). What Makes a Diagram Easy or Hard?: The Impact of Diagram Design on Fourth-Grade Students' Comprehension of Science Texts. *The Elementary School Journal, 119*(1), 122–151.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds). (2019). *The Handbook of Research Synthesis and Meta-Analysis*. Russell Sage Foundation.

Department for Education. (2024). *Initial Teacher Training and Early Career Framework*.

*DeRose, T. (1976). *The effects of verbally and pictorially induced and imposed strategies on children's reading comprehension* (Technical Report No. 394). Wisconsin Research and Development Center for Cognitive Learning, Wisconsin University.

*Désiron, J. C., de Vries, E., Bartel, A. N., & Varahamurti, N. (2018). The influence of text cohesion and picture detail on young readers' knowledge of science topics. *The British Journal of Educational Psychology, 88*(3), 465–479.

*Dexter, D. D. (2010). *Using graphic organizers to teach content area material to students with learning disabilities* [Pennsylvania State University].

*Diana, E. M., & Webb, J. M. (1997). Using geographic maps in classrooms: The conjoint influence of individual differences and dual coding on learning facts. *Learning and Individual Differences, 9*(3), 195–214.

*Dolan, T. (1994). *The Impact of Illustrations and Cultural Schemata on Hong Kong Pupils' Reading Comprehension and Recall of Text*. Annual International Language in Education Conference.

*Dunham, T. C., & Levin, J. R. (1979). Imagery instructions and young children's prose learning: No evidence of "support". *Contemporary Educational Psychology, 4*(2), 107–113.

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics, 56*(2), 455–463.

EEF. (2021, July). *Small group tuition*. Education Endowment Foundation.

*Ehlers-Zavala, F. P. (1999). *Reading an illustrated and non-illustrated story: Dual coding in the foreign language classroom* [Illinois State University]. Engelkamp, J. (1998). Memory for Actions. Routledge.

*Feldman, M. J. (1985). Evaluating pre-primer basal readers using story grammar. *American Educational Research Journal, 22*(4), 527–547.

Fisher, Z., & Tipton, E. (2015). *robumeta: An R-package for robust variance estimation in meta-analysis* (No. arXiv:1503.02220). arXiv.

Fisher, Z., Tipton, E., & Zhipeng, H. (2023). *robumeta: Robust Variance Meta-Regression.* (Version R package version 2.1) [Computer software].

*Gambrell, L. B., & Jawitz, P. B. (1993). Mental Imagery, Text Illustrations, and Children's Story Comprehension and Recall. *Reading Research Quarterly, 28*(3), 264.

*Gerjets, P., Scheiter, K., Opfermann, M., Hesse, F. W., & Eysink, T. H. S. (2009). Learning with hypermedia: The influence of representational formats and different levels of learner control on performance and learning behavior. *Computers in Human Behavior, 25*(2), 360–370.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*(4), 313–331.

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*(6), 511–525.

*Goldberg, F. (1974). Effects of imagery on learning incidental material in the classroom. *Journal of Educational Psychology, 66*(2), 233–237.

*Gonzalez, C. (2017). *Effects of causal diagrams and outlines on fifth-grade students' understanding of scientific texts* [Doctoral Thesis, The University of Florida].

*Griffin, C. C., Simmons, D. C., & Kameenui, E. J. (1991). Investigating the effectiveness of graphic organizer instruction on the comprehension and recall of science content by students with learning disabilities. *Journal of Reading, Writing, & Learning Disabilities International, 7*(4), 355–376.

*Gruhn, S., Segers, E., & Verhoeven, L. (2020). Moderating Role of Reading Comprehension in Children's Word Learning with Context versus Pictures. *Journal of Computer Assisted Learning, 36*(1), 29–45.

Guo, D., Zhang, S., Wright, K. L., & McTigue, E. M. (2020). Do You Get the Picture? A Meta-Analysis of the Effect of Graphics on Reading Comprehension. *AERA Open, 6*(1), 233285842090169.

Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. *Language Teaching, 45*(3), 271–308.

*Haring, M. J., & Fry, M. A. (1979). Effect of Pictures on Children's Comprehension of Written Text. *Educational Communication and Technology: A Journal of Theory, Research, and Development, 27*(3), 185–190.

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). *dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R'* (Version 0.1.0) [Computer software].

*Haslam, C. Y., & Hamilton, R. J. (2010). Investigating the Use of Integrated Instructions to Reduce the Cognitive Load Associated with Doing Practical Work in Secondary School Science. *International Journal of Science Education, 32*(13), 1715–1737.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65.

*Holliday, W. G., & Harvey, D. A. (1976). Adjunct labeled drawings in teaching physics to junior high school students. *Journal of Research in Science Teaching, 13*(1), 37–43.

Hu, L., Chen, G., Li, P., & Huang, J. (2021). Multimedia Effect in Problem Solving: A Meta-Analysis. *Educational Psychology Review, 33*(4), 1717–1747.

*Kiili, K. (2006). Towards a Participatory Multimedia Learning Model. *Education and Information Technologies, 11*(1), 21–32.

*Kim, D., & Gilman, D. A. (2008). Effects of Text, Audio, and Graphic Aids in Multimedia Instruction for Vocabulary Learning. *Educational Technology & Society, 11*(3), 114–126.

*Kourilsky, M., & Wittrock, M. C. (1986). *Verbal and Imaginal Strategies in the Teaching of Economics.* (ERIC Technical Report No. ED 272 414; pp. 1–44). University of California, Los Angeles.

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher, 49*(4), 241–253.

*Kuo, M.-L. A., & Hooper, S. (2004). The Effects of Visual and Verbal Coding Mnemonics on Learning Chinese Characters in Computer-Based Instruction. *Educational Technology Research and Development, 52*(3), 23–38.

*Laney, J. D. (1990). Generative teaching and learning of cost-benefit analysis: An empirical investigation. *Journal of Research & Development in Education, 23*(3), 136–144.

*Lenzner, A., Schnotz, W., & Müller, A. (2013). The Role of Decorative Pictures in Learning. *Instructional Science: An International Journal of the Learning Sciences, 41*(5), 811–831.

*Leopold, C., Doerner, M., Leutner, D., & Dutke, S. (2015). Effects of Strategy Instructions on Learning from Text and Pictures. *Instructional Science: An International Journal of the Learning Sciences, 43*(3), 345–364.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communication and Technology, 30*(4), 195–232.

Levin, J. R. (1981). On functions of pictures in prose. In F. J. Pirozzolo & M. C. Wittrock (Eds), *Neuropsychological and cognitive processes in reading* (pp. 203–228). Academic Press.

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds), *The psychology of illustrations: Basic research* (Vol. 1, pp. 1–50).

Lortie-Forgues, H., & Inglis, M. (2019). Rigorous Large-Scale Educational RCTs Are Often Uninformative: Should We Be Concerned? *Educational Researcher, 48*(3), 158–166.

Lüdecke, D. (2019). *esc: Effect Size Computation for Meta Analysis* (Version 0.5.1) [Computer software]. https://CRAN.R-project.org/package=esc

*Mastropieri, M. A., Scruggs, T. E., McLoone, B., & Levin, J. R. (1985). Facilitating learning disabled students' acquisition of science classifications. *Learning Disability Quarterly, 8*(4), 299–309.

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist, 63*(8), 760–769.

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd edn, pp. 43–71). Cambridge University Press.

*Montare, A., Elman, E., & Cohen, J. (1977). Words and Pictures: A Test of Samuel's Findings. *Journal of Literacy Research, 9*(3), 269–285.

*Moore, P. J., & Skinner, M. J. (1985). The Effects of Illustrations on Children's Comprehension of Abstract and Concrete Passages. *Journal of Research in Reading, 8*(1), 45–56.

Nesbit, J. C., & Adesope, O. O. (2006). Learning With Concept and Knowledge Maps: A Meta-Analysis. *Review of Educational Research, 76*(3), 413–448.

Noetel, M., Griffith, S., Delaney, O., Harris, N. R., Sanders, T., Parker, P., del Pozo Cruz, B., & Lonsdale, C. (2022). Multimedia Design for Learning: An Overview of Reviews With Meta-Meta-Analysis. *Review of Educational Research, 92*(3), 413–454.

*Noh, T., & Scharmann, L. C. (1997). Instructional influence of a molecular-level pictorial presentation of matter on students' conceptions and problem-solving ability. *Journal of Research in Science Teaching, 34*(2), 199–217.

*O'Keefe, E. J., & Solman, R. T. (1987). The influence of illustrations on children's comprehension of written stories. *Journal of Reading Behavior, 19*(4), 353–377.

*Öz, H., & Efecioglu, E. (2015). Graphic Novels: An Alternative Approach to Teach English as a Foreign Language. *Journal of Language and Linguistic Studies, 11*(1), 75–90.

Paivio, A. (1979). *Imagery and Verbal Processes* (1st edn). Psychology Press.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press.

Payne, W. (2022). Dual coding theory in the Classroom: An exploration of teacher's perspectives across two English primary schools. *My College*.

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction, 3*(3), 227–238.

*Peng, C.-Y., & Levin, J. R. (1979). Pictures and Children's Story Recall: Some Questions of Durability. *Educational Communication and Technology: A Journal of Theory, Research, and Development, 27*(1), 39–44.

Perry, T., Lea, R., Jørgensen, C. R., Cordingley, P., Shapiro, K., & Youdell, D. (2021). *Cognitive Science in the Classroom*. Education Endowment Foundation (EEF).

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*(10), 991–996.

*Pressley, M., Pigott, S., & Bryant, S. L. (1982). Picture Content and Preschoolers' Learning from Sentences. *Educational Communication and Technology: A Journal of Theory, Research, and Development, 30*(3), 151–161.

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prevention Science, 23*(3), 425–438.

Readence, J. E., & Moore, D. W. (1981). A meta-analytic review of the effect of adjunct pictures on reading comprehension. *Psychology in the Schools, 18*(2), 218–224.

Reinwein, J. (2012). Does the Modality Effect Exist? And if So, Which Modality Effect? *Journal of Psycholinguistic Research, 41*(1), 1–32.

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review, 7*(3), 216–237.

Rey, G. D., Beege, M., Nebel, S., Wirzberger, M., Schmitt, T. H., & Schneider, S. (2019). A Meta-analysis of the Segmenting Effect. *Educational Psychology Review, 31*(2), 389–419.

Richter, J., Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review, 17*, 19–36.

*Riding, R. J., & Douglas, G. (1993). The effect of cognitive style and mode of presentation on learning performance. *British Journal of Educational Psychology, 63*(2), 297–307.

Rieber, L. P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research, 5*(4), 431–444.

*Rotbain, Y., Marbach-Ad, G., & Stavy, R. (2006). Effect of Bead and Illustrations Models on High School Students' Achievement in Molecular Genetics. *Journal of Research in Science Teaching, 43*(5), 500–529.

*Samuels, S. J. (1967). Attentional process in reading: The effect of pictures on the acquisition of reading responses. *Journal of Educational Psychology, 58*(6), 337–342.

*Saunders, R. J., & Solman, R. T. (1984). The Effect of Pictures on the Acquisition of a Small Vocabulary of Similar Sight-Words. *British Journal of Educational Psychology, 54*, 265–275.

*Scheiter, K., Gerjets, P., & Schuh, J. (2010). The Acquisition of Problem-Solving Skills in Mathematics: How Animations Can Aid Understanding of Structural Problem Features and Solution Procedures. *Instructional Science: An International Journal of the Learning Sciences, 38*(5), 487–502.

Schneider, S., Beege, M., Nebel, S., & Rey, G. D. (2018). A meta-analysis of how signaling affects learning with media. *Educational Research Review, 23*, 1–24.

Schroeder, N. L., & Cenkci, A. T. (2018). Spatial Contiguity and Spatial Split-Attention Effects in Multimedia Learning Environments: A Meta-Analysis. *Educational Psychology Review, 30*(3), 679–701.

Schroeder, N. L., & Cenkci, A. T. (2020). Do measures of cognitive load explain the spatial split-attention principle in multimedia learning environments? A systematic review. *Journal of Educational Psychology, 112*(2), 254–270.

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and Constructing Concept Maps: A Meta-Analysis. *Educational Psychology Review, 30*(2), 431–455.

***Scruggs, T. E., Mastropieri, M. A., & Levin, J. R. (1985)**. Increasing Content-Area Learning: A Comparison of Mnemonic and Visual-Spatial Direct Instruction. *Learning Disabilities Research, a Publication of the Division for Learning Disabilities, 1*(1), 18–31.

***Sherman, J. L. (1976)**. Contextual information and prose comprehension. *Journal of Literacy Research, 8*(4), 369–379.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014a). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General, 143*, 534–547.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014b). p-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results. *Perspectives on Psychological Science, 9*(6), 666–681.

***Snyder, A. E. (2012)**. *The Effects of Graphic Organizers and Content Familiarity on Second Graders' Comprehension of Cause/Effect Text* [Doctoral Dissertation, Columbia University].

***Stebner, F., Höffler, T. N., Kühl, T., Wirth, J., & Ayres, P. (2017)**. The role of process information in narrations while learning with animations and static pictures. *Computers & Education, 104*, 34–48.

Sumeracki, M. (2019). Six Strategies for Effective Learning: A Summary for Teachers. *The Learning Scientists*. https://www.learningscientists.org/blog/2019/11/28-1

Sundararajan, N., & Adesope, O. (2020). Keep it Coherent: A Meta-Analysis of the Seductive Details Effect. *Educational Psychology Review, 32*(3), 707–734.

Sutton, A. J. (2009). Publication bias. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds), *The Handbook of Research Synthesis and Meta-Analysis* (2nd edn, pp. 435–452). Russell Sage Foundation.

***Swanson, H. L. (2015)**. Cognitive strategy interventions improve word problem solving and working memory in children with math disabilities. *Frontiers in Psychology, 6*, 1–13.

***Swanson, H. L., Lussier, C. M., & Orosco, M. J. (2015)**. Cognitive Strategies, Working Memory, and Growth in Word Problem Solving in Children With Math Difficulties. *Journal of Learning Disabilities, 48*(4), 339–358.

Sweller, J. (2010). Cognitive load theory: Recent theoretical advances. In J. L. Plass, R. Moreno, & R. Brünken (Eds), *Cognitive load theory* (pp. 29–47). Cambridge University Press.

Thomas, J., Graziosi, S., Brunton, J., Ghouze, Z., O'Driscoll, P., Bond, M., & Koryakina, A. (2022). *EPPI-Reviewer: Advanced software for systematic reviews, maps and evidence synthesis* (Version 4.14.2.0) [Computer software]. EPPI Centre, UCL Social Research Centre.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20*(3), 375–393.

Vekiri, I. (2002). What Is the Value of Graphical Displays in Learning? *Educational Psychology Review, 14*(3), 261–312.

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika, 60*(3), 419–435.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1–48.

Vogt, F., & Rogalla, M. (2009). Developing Adaptive Teaching Competency through coaching. *Teaching and Teacher Education, 25*(8), 1051–1060.

***Wachter, L. N. (1993)**. *An Investigation of the Effects of Hierarchical Concept Mapping as a Prefatory Organizer on Fourth-grade Students' Comprehension and Retention of Expository Prose* [Doctoral Thesis, Pennsylvania State University].

Wang, M. C. (1980). Adaptive instruction: Building on diversity. *Theory Into Practice, 19*(2), 122–128.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686.

Xie, H., Wang, F., Hao, Y., Chen, J., An, J., Wang, Y., & Liu, H. (2017). The more total cognitive load is reduced by cues, the better retention and transfer of multimedia learning: A meta-analysis and two meta-regression analyses. *PLOS ONE, 12*, e0183884.

Zhang, Y., Guo, X., Pi, Z., & Yang, J. (2021). Learning by Drawing in STEM: A Meta-Analysis. *2021 Tenth International Conference of Educational Innovation through Technology (EITT)*, 83–87.

*Zheng, R. Z., Yang, W., Garcia, D., & McCadden, E. P. (2008). Effects of Multimedia and Schema Induced Analogical Reasoning on Science Learning. *Journal of Computer Assisted Learning, 24*(6), 474–482.

Appendix

A1. Literature searches

Search terms

Intervention:

visual* OR picture* OR image* OR drawing* OR pictorial OR illustration* OR multimedia OR animation* OR "dual coding" OR diagram* OR icon* OR "concept map" OR "graphic organizer"

Population:

pupil* OR grade* OR school* OR classroom* OR nursery OR preschool OR kindergarten OR "early years" OR "sixth form" OR "further education"

Outcome:

learning OR comprehension OR retention OR transfer OR "problem solving" OR knowledge OR achievement OR recall

Design:

random* OR experiment* OR intervention* OR "control group*" OR "control condition*" OR "treatment arm*" OR "treatment condition*" OR "treatment group*" OR posttest* OR "post-test*"

Databases

EconLit (EBSCO)
Education Resources Information Center (ERIC) (EBSCO)
Education Abstracts (EBSCO)
EPPI-Centre database of education research
PsycINFO (OVID)
PsycARTICLES

Reviews and meta-analyses

Alpizar et al. (2020)
Berney & Betrancourt (2016)
Carney & Levin (2002)
Ginns (2005, 2006)
Guo, McTigue et al. (2020)
Hu et al. (2021)
Levie & Lentz (1982)
Nesbit & Adesope (2006)
Readence & Moore (1981)
Rey (2012)
Rey et al. (2019)
Richter et al. (2016)
Schneider et al. (2018)
Schroeder & Cenkci, (2018)
Schroeder et al. (2018)
Sundararajan & Adesope (2020)
Vekiri (2002)
Xie et al. (2017)
Zhang et al. (2021)

A2. Included studies and outcomes

Table S1. Description of the 56 studies included in the meta-analysis

Authors (year)	Country	Learning outcomes	Intervention	Total N	Grade/Year as reported (target population)	Subject (topic or skill)
Acha (2009)	Spain	Immediate / delayed recall (2 outcomes)	Picture + written text, Computer-based	135	Grades 3-4	Second languages (learning English vocabulary)
Alvermann (1981)	USA	Immediate / delayed recall (4 outcomes)	Graphic organiser + descriptive or comparative text	114	Grade 10	Science & Technology (human biology)
Alvermann & Boothby (1986)	USA	Delayed transfer (6 outcomes)	Graphic organiser, 7 or 14 days instruction	24	Grade 4	Humanities (American history)
Ardasheva et al. (2018)	USA	Immediate understanding (1 outcome)	Pictures + written text	174	Grade 7 (English learners)	Second languages (ESL science reading comprehension)
Babaei & Izadpanah (2019)	Iran	Immediate understanding (1 outcome)	Advance organiser with multimedia annotations (pictures + related keywords)	128	High school (grade not reported; female learners)	Second languages (English listening comprehension)
Barnitz & Speaker Jr (1991)	United States	Immediate understanding and retention (2 outcomes)	Picture + poem	11	Grade 7 (English learners)	Second languages (ESL comprehension)
Bobis et al. (1993)	Australia	Immediate retention (1 outcome)	Written instructions + diagram	45	Grade 4	Mathematics (geometry; paper folding)
Brandt et al. (2001)	Belgium	Immediate understanding (3 outcomes)	Concept map + drawings	124	Final year students in secondary schools (grade not reported)	Science & Technology (electrochemistry)
ChanLin (2001)	Taiwan*	Immediate understanding, retention and transfer (8 outcomes)	Pictures and animations, computer-based	357	Grade 8 and 9	Science & Technology (physics)

Authors (year)	Country	Learning outcomes	Intervention	Total N	Grade/Year as reported (target population)	Subject (topic or skill)
Coleman et al. (2018)	United States	Immediate retention and comprehension (12 outcomes)	Science text + diagrams	213	Grade 4 (children who could read independently)	Science & Technology (water cycle and circulatory system)
DeRose (1976)	United States	Immediate retention (1 outcome)	Textbook passage + experimenter-provided illustration or verbal summary	192	Grade 5	English & Literacy (reading comprehension: industrial revolution)
Desiron et al. (2018)	United States	Immediate retention (1 outcome)	Written text + representational pictures	101	Grades 2, 3, 4	Science & Technology (tsunamis and drought)
Dexter (2010)**	United States	Immediate retention and transfer (3 outcomes)	Semantic map + diagram	62	Grade 8 (social studies inclusion classroom)	Humanities (history – feudalism in the middle ages)
Diana & Webb (1997)	United States	Immediate and delayed retention (4 outcomes)	Verbal presentation and discussion + map with factual icons	258	Grade 6	Humanities (ancient history)
Dolan (1994)	United	Delayed retention (8 outcomes)	Written text + black and white illustrations	193	Secondary (Hong Kong Form 3; UK Year 3)***	English & Literacy (reading comprehension: adventure story)
Dunham & Levin (1979)	United States	Immediate retention (1 outcome)	Pictures from story	160	Kindergarten, Grade 1	English & Literacy (reading comprehension: narrative passages)
Ehlers-Zavala (1999)	Chile	Immediate understanding (2 outcomes)	Detailed illustrations	129	Junior and senior high (ages 16-18; grade not reported)	Second languages (English reading comprehension: Nettie's Trip South)
Feldman (1985)	United States	Immediate understanding, retention and transfer (4 outcomes)	Original picture from basal reader	96	Grade 1	English & Literacy (reading comprehension: basal readers)

Authors (year)	Country	Learning outcomes	Intervention	Total N	Grade/Year as reported (target population)	Subject (topic or skill)
Gambrell and Jawitz (1993)	United States	Immediate, mixed outcomes (2 outcomes)	Text-relevant illustrations	120	Grade 4 (Selected based on reading comprehension, cognitive and reading ability.)	English & Literacy (reading comprehension: mystery story)
Gerjets et al. (2009)	Germany	Immediate, mixed outcomes (2 outcomes)	Animation – hypermedia learning environment, computer-based,	118	Grades 10 and 11	Mathematics (probability theory)
Goldberg (1974)	United States	Immediate, retention (2 outcomes)	sentences from social science/ science texts + pictures	216	Grade 5	English & Literacy (reading comprehension)
Gonzalez (2017)	United States	Immediate, mixed outcomes (1 outcome)	Written text + causal diagram	99	Grade 5	Science & Technology (chemical processes: oxidation/chemical reactions)
Griffin et al. (1991)	United States	Delayed, retention (1 outcome)	Passage of text or graphic organiser reflecting hierarchical structure of text	28	Grade 5 and 6 (with learning disabilities)	Science & Technology (fossil fuels)
Gruhn et al. (2020)	The Nether-lands	Immediate and delayed, retention (2 outcomes)	Written word and definition + picture, computer based,	189	Grade 4	English & Literacy (vocabulary learning)
Haring & Fry (1979)	United States	Immediate and delayed, retention (16 outcomes)	written text + pictures	150	Grades 4 and 6	English & Literacy (reading comprehension: story)
Haslam & Hamilton (2010)	New Zealand	Immediate understanding and retention (2 outcomes)	Written instructions + colour-coded, text-integrated diagrams	96	Secondary school (grade not reported)	Science & Technology (practical lesson on electrical circuits)
Holliday & Harvey (1976)	Canada	Immediate mixed outcomes (1 outcome)	Written text + labelled diagram	61	Grade 9	Science & Technology (density and pressure)

Authors (year)	Country	Learning outcomes	Intervention	Total N	Grade/Year as reported (target population)	Subject (topic or skill)
Kiili (2006)	Finland	Immediate understanding (2 outcomes)	Computer-based. Written text + interactive pictures and animations	167	Grade not reported – ages 10-12, elementary school	Science & Technology (human immune system)
Kim & Gilman (2008)	South Korea	Immediate and delayed retention (4 outcomes)	Computer-based; words and definitions + pictures and animations	172	Middle school (age 14, grade not reported)	Second languages (vocabulary learning)
Kourilsky & Wittrock (1986)	United States	Immediate understanding (2 outcomes)	Verbal instruction and discussion + diagrams (graphs)	83	High school seniors (grade not reported)	Humanities (economics)
Kuo & Hooper (2004)	United States	Delayed and immediate retention (4 outcomes)	Chinese characters + English translations + Picture, computer-based	92	High school (grade not reported)	Second languages (learning Chinese characters)
Laney (1990)	United States	Immediate transfer (1 outcome)	Verbal instruction and discussion + diagrams and pictures	111	Grade 3	Humanities (economics)
Lenzner et al. (2013)	Germany	Immediate, understanding (2 outcomes)	Computer-based; written text + decorative or instructional pictures	194	Grades 7 and 8	Science & Technology (light and shadow)
Leopold et al. (2015)	Germany	Immediate retention (3 outcomes)	Text + pictures	108	Grade 11	Science & Technology (water molecules)
Mastropieri et al. (1985)	United States	Immediate retention (2 outcomes)	Written text + mnemonic pictures or illustrations	36	Grades 7 and 8 (learning disabled students)	Science & Technology (mineral classifications)
Montare et al. (1977)	United States	Immediate retention (2 outcomes)	Word + picture	30	Kindergarten	English & Literacy (word acquisition)
Moore & Skinner (1985)	Australia	Immediate mixed (2 outcomes)	Reading passages + illustrations	55	Grade 6	English & Literacy (reading comprehension)

Authors (year)	Country	Learning outcomes	Intervention	Total N	Grade/Year as reported (target population)	Subject (topic or skill)
Noh & Scharmann (1997)	Korea	Delayed, understanding (2 outcomes)	Textbook and verbal instruction + diagrams of concepts	98	Grade 11 (Two all-boys declared science major class groups)	Science & Technology (chemistry)
O'Keefe & Solman (1987)	Australia	Immediate retention (6 outcomes)	Text + story pictures	28 and 32	Grade 5 and Grade 3	English & Literacy (reading comprehension: Flat Stanley)
Oz & Efecioglu (2015)	Turkey	Immediate mixed (1 outcome)	Plain text of play or graphic novel of play	56	Grade 10	Second languages (English reading comprehension: Macbeth graphic novel)
Peng and Levin (1979)	United States	Immediate and delayed retention (4 outcomes)	Verbal narration + pictures, or verbal narration + written text	64	Grade 2	English & Literacy (reading comprehension: narrative stories)
Pressley et al. (1982)	Canada	Immediate retention (2 outcomes)	Verbal sentence + picture	36 and 30	Preschool	English & Literacy (listening comprehension)
Riding and Douglas (1993)	United Kingdom	Immediate mixed (1 outcome)	Written text + picture	59	Year 11	Science & Technology (car brake systems)
Rieber (1989)	United States	Immediate transfer (8 outcomes)	Computer-based; on screen text + diagrams or animations	192	Grades 4, 5 and 6	Science & Technology (Physics, Newtonian mechanics)
Rotbain et al. (2006)	Israel	Delayed mixed (2 outcomes)	Verbal instruction + illustrations	258	Grade 11 and 12 (majoring in biology)	Science & Technology (molecular genetics)
Samuels (1967)	United States	Immediate retention (2 outcomes)	Written words + pictures	52	Grade 1	English & Literacy (reading words)
Saunders and Solman (1984)	Australia	Immediate retention (1 outcome)	Written words + pictures	70	kindergarten	English & Literacy (reading words)
Scheiter et al. (2010)	Germany	Immediate transfer (4 outcomes)	Computer-based; text and algebraic formulas + pictures and animations	32	Grade 9	Mathematics (algebra word problems)

Authors (year)	Country	Learning outcomes	Intervention	Total N	Grade/Year as reported (target population)	Subject (topic or skill)
Scruggs et al. (1985)	United States	Immediate retention (1 outcome)	Written text + mnemonic pictures	36	Grades 7, 8 and 9 (learning disabled students)	Science & Technology (minerals of North America)
Sherman (1976)	United States	Immediate retention (12 outcomes)	Passages of text + contextual pictures	144	Grade not reported - high school juniors	English & Literacy (reading comprehension: concrete and abstract paragraphs)
Snyder (2012)	United States	Immediate understanding and retention (2 outcomes)	Written text + graphic organiser	80	Grade 2	English & Literacy (reading comprehension: cause and effect)
Stebner et al. (2017)	Germany	Immediate understanding (3 outcomes)	Computer-based; on-screen text (with/without verbal narration) + pictures or animations	162	Grades 7 and 8	Science & Technology (surfactants and washing)
Swanson (2015)	United States	Delayed transfer (4 outcomes)	Written text and verbal explanations + diagrams (completed by pupils)	204	Grade 3 (with and without maths difficulties)	Mathematics (word problems)
Swanson et al. (2015)	United States	Delayed transfer (6 outcomes)	Written text and verbal explanations + diagrams (completed by pupils)	192	Grades 2 and 3 (with and without maths difficulties)	Mathematics (word problems)
Wachter (1993)	United States	Comprehension (4 outcomes)	Written passages + concept map (complete or partially complete)	120	Grade 4 (stratified by reading comprehension ability)	Humanities (history of Egypt)
Zheng et al. (2008)	United States	Immediate, retention and transfer (1 outcome)	Computer-based; interactive multimedia with and without verbal analogy	89	Grade 4	Science & Technology (water systems and electrical circuits)

Note. *Not explicitly reported, inferred from authors' location.

**We reviewed the doctoral dissertation only, there is also a published version available: Dexter, D. (2012). Using Graphic Organizers to Teach Content Area Material to Students with Learning Disabilities. International Journal for Research in Learning Disabilities, 1(1), 79-105.

***Reported as Year 3 in a secondary school (translates as current UK Year 9).

A great teacher changes the future every day. They can be the critical factor in a child's success, especially for those who have had a tough start in life. At Ambition Institute, we support teachers and school leaders at every stage of their careers, helping them to keep getting better. We are a charity providing training and professional development based on the most rigorous research and evidence about what really works. Together, we're shaping the future of education to give every child the best start in life.

Learn more

Get in touch today to find out how you can nominate your staff, or yourself, for one of our programmes.

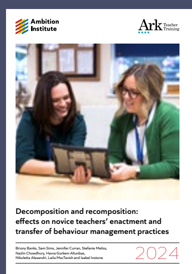
ambition.org.uk
info@ambition.org.uk
020 3846 5828
@Ambition.Inst
Ambition Institute

Ambition Institute
2 Bridge Wharf,
156 Caledonian Road,
London, N1 9UU

Discover more

Evidence is at the heart of what we do at Ambition Institute. Through our research we want to push forward our collective knowledge of effective professional development and share what works, to turn evidence into action. We aim to answer important questions faced by the sector, help teachers and leaders to address disadvantage, and show the impact of evidence-informed teaching.

This research explored whether advance organisers can help teachers learn and understand new information during online professional development.



This research explored whether breaking down (decomposing) and then recombining teaching practices into a different sequence of teaching (recomposing) helped initial teacher trainees apply what they learnt to new contexts.

Ambition Institute