Education
Endowment
Foundation

The Sutton Trust-Education Endowment Foundation Teaching & Learning Toolkit

December 2016

https://educationendowmentfoundation.org.uk/resources/teaching-learning-toolkit

Durham University

THE SUTTON TRUST

The Teaching & Learning Toolkit

The Sutton Trust-EEF Teaching and Learning Toolkit is an accessible summary of educational research. It provides guidance for teachers and schools on how to use their resources to improve the attainment of disadvantaged pupils. The Toolkit currently covers 34 topics, each summarised in terms of the average impact on attainment, the strength of the supporting evidence and the cost.

The Toolkit is a live resource that will be updated on a regular basis as findings from EEF-funded projects and other high-quality research become available. We would welcome suggestions for new topics to be included in future editions. If you have a topic suggestion, or any other comments or questions about the Toolkit, please contact Danielle Mason at danielle.mason@eefoundation.org.uk.

Why is research useful?

We know that the relationship between spending and pupil outcomes is not simple. Per pupil spending increased by 85% between 1997 and 2011, but improvements in pupil outcomes were marginal on most measures. At the school level, it is clear that different ways of spending school budgets can have very different impacts on pupil attainment, and choosing what to prioritise is not easy. Even once a decision to implement a particular strategy has been made, there are a wide variety of factors which determine its impact. We believe that educational research can help schools get the maximum 'educational bang for their buck', both in terms of making an initial choice between strategies, and in implementing a strategy as effectively as possible.

One particular spending decision which research can inform is how to spend the Pupil Premium. Introduced in 2010, the aim of the Pupil Premium is to raise attainment among disadvantaged children. It provides additional funding to schools for disadvantaged pupils to ensure they benefit from the same educational opportunities as pupils from wealthier families. In the 2015-16 financial year the Pupil Premium was worth £935 per eligible child in secondary schools and £1320 per eligible child in primary schools. If the Pupil Premium is to succeed in achieving its ambitious goals, the choices that schools make in allocating the money are of vital importance.

A range of teaching and learning approaches were selected for analysis and inclusion in the Toolkit. The choice of approaches was based

on: *i)* approaches commonly mentioned in connection with education policy, *ii)* suggestions from schools, and *iii)* approaches with a strong evidence of effectiveness not covered by either previous criterion.

Months' impact

Months' impact is estimated in terms of the additional months' progress you can expect pupils to make as a result of an approach being used in school, taking average pupil progress over a year as a benchmark.

For example, the 'feedback' strand in the Toolkit shows that improving the quality of feedback provided to pupils has an average impact of eight months. This means that pupils in a class where high quality feedback is provided can be expected to make, on average, eight months more progress over the course of a year compared to another class of pupils who are performing at the same level at the start of the year. At the end of the year the average pupil in a class of 25 pupils in the feedback group would now be equivalent to the 6th best pupil in the control class, having made 20 months progress over the year, compared to an average of 12 months in the other class.

These impact estimates are based on ‘effect sizes’ reported in British and international data (see table below). Effect sizes are quantitative measures of the impact of different approaches on learning. The Toolkit prioritises effect sizes derived from systematic reviews of research and quantitative syntheses of data such as meta-analyses of experimental studies. Approaches are only included in the Toolkit if there is a quantifiable evidence base which can be used to derive effect sizes. For more information about the Toolkit methodology please view the Toolkit Technical Appendices.

Most approaches included in the Toolkit tend to have very similar average impacts on pupils with different characteristics. However, where the research summarised suggests that an approach has a different average impact on the learning of pupils from disadvantaged backgrounds compared to the learning of their peers, the Toolkit’s ‘headline’ average impact figure refers to the former.

Months' Progress	Effect size from...	to...	Description
	-0.01	0.01	Very low or no impact
1	0.02	0.09	Low impact
2	0.10	0.18	Low impact
3	0.19	0.26	Moderate impact
4	0.27	0.35	Moderate impact
5	0.36	0.44	Moderate impact
6	0.45	0.52	High impact
7	0.53	0.61	High impact
8	0.62	0.69	High impact
9	0.70	0.78	Very high impact
10	0.79	0.87	Very high impact
11	0.88	0.95	Very high impact
12	0.96	>1.0	Very high impact

Cost

Cost estimations are based on the approximate cost of implementing an approach in a class of 25 pupils. Estimates commonly include the cost of additional resources, and training or professional development if required. For more information about the Toolkit methodology please view the Toolkit Technical Appendices.

Cost	Description
£	Very low: up to about £2,000 per year per class of 25 pupils, or less than £80 per pupil per year.
££	Low: £2,001 to £5,000 per year per class of 25 pupils, or up to about £200 per pupil per year
£££	Moderate: £5,001 to £18,000 per year per class of 25 pupils, or up to about £700 per pupil per year.
££££	High: £18,001 to £30,000 per year per class of 25 pupils, or up to £1,200 per pupil.
£££££	Very high: over £30,000 per year per class of 25 pupils, or over £1,200 per pupil.

Evidence

The toolkit presents a rating of evidence security for each approach - a 'padlock' security rating. These evidence ratings are based on: the *quantity* of evidence available (i.e. the number of systematic reviewsor meta-analyses and the number of primary studies which they synthesise); the methodological *quality* of the available evidence; and the *consistency* of estimated impact across the systematic reviews and meta-analyses that have been synthesised. For more information about the Toolkit methodology please view the Toolkit Technical Appendices.

Rating	Description
1 padlock	Very limited: Quantitative evidence of impact from single studies, but with effect size data reported or calculable. No systematic reviews with quantitative data or meta-analyses located.
2 padlocks	Limited: At least one meta-analysis or systematic review with quantitative evidence of impact on attainment or cognitive or curriculum outcome measures.
3 padlocks	Moderate: Two or more rigorous meta-analyses of experimental studies of school age students with cognitive or curriculum outcome measures.
4 padlocks	Extensive: Three or more meta-analyses from well-controlled experiments mainly undertaken in schools using pupil attainment data with some exploration of causes of any identified heterogeneity.
5 padlocks	Very Extensive: Consistent high quality evidence from at least five robust and recent meta-analyses where the majority of the included studies have good ecological validity and where the outcome measures include curriculum measures or standardised tests in school subject areas.

Who wrote the Toolkit?

The Toolkit was originally commissioned by the Sutton Trust and produced as the *'Pupil Premium Toolkit'* by Durham University in May 2011. The Sutton Trust-EEF Teaching and Learning Toolkit has been developed from this initial analysis, since the Education Endowment Foundation's launch in 2011.

The Toolkit has been written by Professor Steve Higgins, Dr Maria Katsipataki, Dr Berenice Villanueva-Aguilera (School of Education, Durham University), Professor Rob Coe (CEM Centre, Durham University), Dr Lee Elliot Major (The Sutton Trust), and Robbie Coleman, Peter Henderson and Danielle Mason (Education Endowment Foundation).

Full reference: Higgins, S., Katsipataki, M., Villanueva-Aguilera, A.B.V., Coleman, R., Henderson, P., Major, L.E., Coe, R. & Mason, D. (2016) The *Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.* December 2016. London: Education Endowment Foundation. https://educationendowmentfoundation.org.uk/resources/teaching-learning-toolkit

Arts participation

Low impact for low cost, based on moderate evidence.

+2

Arts participation is defined as involvement in artistic and creative activities, such as dance, drama, music, painting, or sculpture. It can occur either as an additional part of the curriculum or as extra-curricular activities. Participation may be organised as regular weekly or monthly activities or more intensive programmes such as summer schools or residential courses.

How effective is it?

Overall, the impact of arts participation on academic learning appears to be positive but low. Improved outcomes have been identified in English, mathematics and science learning. Benefits have also been found in both primary and secondary schools, though on average greater effects have been identified for younger learners.

In some cases, specific arts activities have been linked with benefits on particular outcomes. For example, there is some evidence of a positive link between music and spatial awareness. Wider benefits on attitudes to learning and well-being have also consistently been reported.

How secure is the evidence?

There are a number of systematic reviews and meta-analyses which have found small benefits for arts participation. However, these effects vary according to the type of approach and the age group targeted, so are hard to generalise.

A recent systematic review conducted for the EEF found no individual studies that passed a high benchmark for their security and convincingly demonstrated that arts participation had an impact on attainment. This Toolkit strand synthesizes meta-analyses, and gives a general picture of the pattern of findings in the literature.

What are the costs?

Costs vary considerably from junior drama groups with small annual subscriptions (about £20), through organised dance groups for young people at about £5 per session to high quality music tuition at about £35 per hour. Costs are estimated at £150 per year, though it should be noted that some activities would be considerably more expensive (e.g. nearer £1,500 for individual music tuition). Overall costs are estimated as low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. The research evidence shows a wide range of effects from the programmes studied, suggesting that achieving learning gains from arts programmes is not straightforward.
2. Benefits for learning appear to be more achievable with younger learners, with some promising evidence supporting the academic impact of programmes which develop skills in music performance in particular.
3. Arts-based approaches may offer a route to re-engage older learners in school.
4. The transfer of learning to the classroom is not automatic and needs further exploration. For example, how can you encourage pupils to apply their learning from arts participation to more formal contexts?

Technical Appendix: Arts participation

Definition

Arts participation is defined as involvement in artistic and creative activities, such as dance, drama, music, painting, and sculpture, either as an additional part of the curriculum or as extra-curricular activities. Participation may be organised as regular weekly or monthly activities or more intensive programmes such as summer schools or residential courses. Arts education and participation include a broad range of subjects including the traditional fine arts (e.g. visual arts, music, dance, performing arts, theatre and dance), modern dance and movement, poetry and creative writing, as well as teaching strategies which explicitly include arts elements such as drama-based pedagogy.

Search terms: Arts in education; arts/ fine arts/ performing arts participation, arts/fine arts/ performing arts; music education; drama education; dance education.

Evidence rating

There are seven meta-analyses, which have consistently found academic benefits for arts participation. However pooled effect sizes range widely from 0.03 to 0.77, nearly three quarters of a standard deviation and are therefore not consistent. Five of these have been published in the last 10 years. There are some indications of patterns of effect such as higher impact for younger children and for music studies but these are not consistent and vary according to the detail of the approach and the age group targeted. The quality of evaluation designs used by studies of arts participation has been criticised as insufficiently robust to draw causal inference in recent reviews. The effects may therefore be hard to generalise and the research not conclusive. Overall the evidence is rated as moderate.

Additional cost information

Costs vary considerably from junior drama groups with small annual subscriptions (about £20), through organised dance groups for young people at about £5 per session to high quality music tuition at about £35 per hour. Costs are estimated at £150 per year, though it should be noted that some activities would be considerably more expensive (e.g. nearer £1,500 for individual music tuition). The Act, Sing, Play programme, evaluated by the EEF, had financial costs relating to tutor training, travel, purchase of musical instruments, and other resources associated with delivering the programme estimated at £219 per pupil per year, though costs for a second year would be substantially lower. Overall costs are estimated as low.

References

Full references
Borman, G. D., Goetz, M. E., & Dowling, N. M. (2009). Halting the summer achievement slide: A randomized field trial of the KindergARTen Summer Camp. *Journal of Education for Students Placed at Risk*, 14(2), 133-147. http://dx.doi.org/10.1080/10824660802427652
Catterall, J. S., Dumais, S. A., & Hampden-Thompson, G. (2012). *The arts and achievement in at-risk youth: Findings from four longitudinal studies*. Washington: National Endowment for the Arts. http://www.nea.gov/research/Arts-At-Risk-Youth.pdf
*Conard, F. (1992). The arts in education and a meta-analysis. Doctoral disseration, Purdue University West Lafayette, IN http://docs.lib.purdue.edu/dissertations/AAI9229100/
Daykin, N., Orme, J., Evans, D., Salmon, D., McEachran, M. & Brain, S. (2008) The impact of participation in performing arts on adolescent health and behaviour: A systematic review of the

literature. *Journal of Health Psychology*, 13 (2). pp. 251-264. http://dx.doi.org/10.1177/1359105307086699

Haywood, S., Griggs, J., Lloyd, C., Morris, S., Kiss, Z. & Skipp, A. (2015) *Creative Futures: Act, Sing, Play Evaluation report and Executive summary June 2015* London: Education Endowment Foundation https://educationendowmentfoundation.org.uk/uploads/pdf/Act__Sing__Play.pdf

*Lee, B. K., Patall, E. A., Cawthon, S. W., & Steingut, R. R. (2014). The Effect of Drama-Based Pedagogy on PreK–16 Outcomes A Meta-Analysis of Research From 1985 to 2012. *Review of Educational Research* 85 (1) 3-49 http://dx.doi.org/10.3102/0034654314540477

Lord, P., Sharp, C., Dawson, A., Mehta, P., White, R. and Jeffes, J. (2013). Evaluation of In Harmony: Year 1 Interim Report. Slough: NFER. http://www.nfer.ac.uk/nfer/publications/ACII01/ACII01.pdf

*Lewis C.P. (2004) *The Relation between Extracurricular Activities with Academic and Social Competencies in School Age Children: A Meta-Analysis.* Doctoral dissertation, Texas A&M University College Station, TX: Texas A&M University. http://oaktrust.library.tamu.edu/bitstream/handle/1969.1/2710/etd-tamu-2004B-SPSY-lewis.pdf?sequence=1&isAllowed=y

*Newman M., Bird K., Tripney J., Kalra N., Kwan I., Bangpan M., & Vigurs C (2010). *Understanding the Impact of Engagement in Culture and Sport: A Systematic Review of the Learning Impacts for Young People. Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3358

See, B. & Kokotsaki, D. (2015) *Impact of arts education on the cognitive and non-cognitive outcomes of school-aged children A review of evidence London*: Education Endowment Foundation.

*Shulruf, B. (2010). Do extra-curricular activities in schools improve educational outcomes? A critical review and meta-analysis of the literature. *International Review of Education*, 56(5-6), 591-612. http://dx.doi.org/10.1007/s11159-010-9180-x

*Standley, J.M. (2008). Does Music Instruction Help Children Learn to Read? Evidence of a Meta-Analysis. Update. *Applications of Research in Music Education*, 27.1, pp. 17-32. http://dx.doi.org/10.1177/8755123308322270

*Winner, E., & Cooper, M. (2000). Mute Those Claims: No Evidence (Yet) for a Causal Link between Arts Study and Academic Achievement. *Journal of Aesthetic Education*, 34(3/4), 11. http://www.jstor.org/stable/3333637

*studies included in the summary of effects

Technical Appendix: Arts participation

Summary of effects		
Meta-analyses	**FSM effect size**	**Overall effect size**
Conard, 1992 (reading)		0.48
Conard, 1992 (maths)		0.29
Conard, 1992 (writing)		0.77
Lee et al. 2014		0.43
Lewis, 2004		0.20
Newman et al., 2010 (secondary science)		0.06
Newman et .al., 2010, (secondary English)		0.05
Newman et al., 2010 (secondary mathematics)		0.03
Newman et al., 2010 (prim/EY cognitive)		0.45
Shulruf, 2010 (GPA)		0.17
Shulruf, 2010 (maths)		0.05
Shulruf, 2010 (reading)		0.05
Standley, 2008		0.32
Winner & Cooper, 2000 (overall)		0.10
Recent single studies		
Borman et al. 2009		0.40
Haywood et al. 2015[1] (EEF evaluation: Music compared with drama on maths)		0.00
Haywood et al. 2015[1] (EEF evaluation: Music compared with drama on literacy)		0.05
Weighted mean effect size		**0.17**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Conard, 1992	This study is a meta-analysis of previous experimental studies which examined the effect of creative dramatics on the acquisition of cognitive skills. The areas of research investigated included: the achievement of students experiencing creative dramatic techniques as compared to traditional methods, the impact of study and sample characteristics on outcomes, and the effects of research and methodological features on outcomes. Refined meta-analysis methodology was used that weighted each study independently, thus accounting for varying sizes of groups used in the individual studies. A mean effect size of 0.48 was found for studies in which creative dramatics was used as an instructional technique. Creative dramatics tended to be more effective at the pre-school and elementary level than at the secondary level. Both regular and remedial students appear to benefit from, and enjoy participating in creative dramatics. Studies that used students in private schools produced larger effect sizes than those that used public school students. More detailed documentation of the different types of creative drama treatments is needed. Specifically, studies should include exactly what was done, how it was done, and how the effects were measured. Measurement characteristics, such as, reliability and validity, and other details of the dependent measures were often not reported in the creative dramatics research literature. Detailed reporting of study characteristics facilitates research synthesis. Qualitative reviews were combined with the quantitative analysis. The qualitative data greatly enhanced the results of the meta-analysis, and added further insights in interpreting

[1] Comparison between music and drama: not included in pooled effect

	the findings.
Lee et al. 2014	The President's Committee on the Arts and Humanities report heartily supported arts integration. However, the President's Committee called for a better understanding of the dimensions of quality and best practices. One promising arts integration method is drama-based pedagogy (DBP). A comprehensive search of the literature revealed 47 quasi-experimental DBP intervention studies conducted since 1985. The literature showed that designs were generally weak for making causal inferences and that outcomes other than achievement were infrequently studied. A meta-analysis of this research suggested that DBP has a positive, significant impact on achievement outcomes in educational settings. Effects were strongest when the intervention (a) was led by a classroom teacher or researcher rather than a teaching artist, (b) included more than five lessons, and (c) was integrated into English language arts or science curriculum compared to other domains. Positive effects across psychological and social outcomes were found. Implications for policy and practice are discussed.
Lewis, 2004	There has been a growing discussion in the fields of education and psychology about the relationship between social skill proficiency and academic excellence. However, the presence of extracurricular involvement as promoting both academic and social development has not been thoroughly explored. The most recent literature syntheses and meta-analyses on extracurricular activity participation were conducted in the 1980.s. An updated review and quantitative look at the participation literature is due. The purpose of this study is to integrate participation studies from the 1990s and give summative information as to the impact of extracurricular activity participation on various educational and psycho-social characteristics. Of the 164 identified studies, 41 were included in these meta-analyses. The current analyses produced 6 different activity categories: general extracurricular activity, sports, work and vocational activities, performing arts, pro-social activities, and community-based activities. The current meta-analysis suggests student outcomes were significantly related to general extracurricular activity and pro-social activity participation. General activities and pro-social activities had the most impact on academic achievement, while performing arts and pro-social activities. Participants reported the largest effect on identity and self-esteem related outcomes. Sports and related activities (i.e. Cheerleading) were not as strongly linked to academic achievement indicators as anticipated and student workers had more negative outcomes than any other activity participants. In conclusion, the best outcomes for children and adolescents are brought about through well-built, developmentally appropriate structured activities. Moreover, the academic and social profits of extracurricular activities that have been examined in this study can be used to inform program planning and implementation.
Newman et al. 2010	A set of systematic reviews were undertaken to examine the impact of young people's engagement (in each sector: sport, arts, MLA, heritage) on their learning. This systematic review aims to examine the impact of young people's participation in the arts. A subset of the studies used research designs and outcome measures that allowed for the findings to be translated into a common metric of effect sizes and so meta-analysed. The main aim of the meta-analysis was to combine results (where appropriate) across studies, therefore we have used data that could be translated into a common metric of effect sizes (Hedges' g). Most studies reported multiple learning outcomes; where possible, effect sizes were calculated for all outcomes reported in these studies. When compared to non-participation in structured arts activities: Participation in structured arts activities improves academic attainment in secondary school aged students. Participation in such activities could increase their academic attainment scores by 1% and 2%, on average, above that of non-participants (all other things being equal). Participation in structured arts activities improves pre-school and primary school aged children's early literacy skills. This result is based on narrative numerical synthesis and thus we are unable to estimate the size of any positive effect.

	Participation in structured arts activities improves young people's cognitive abilities (based on various measures of intelligence). Participation of young people in such activities could increase their cognitive abilities test scores by 16% and 19%, on average, above that of non-participants (all other things being equal). Participation in structured arts activities improves young people's transferable skills. Participation of young people in such activities could increase their transferable skills test scores by 10% and 17%, on average, above that of non-participants (all other things being equal).
Shulruf, 2010	Secondary schools tend to sponsor a large number of extra-curricular activities (ECA) yet little is known about their contribution to students' educational outcomes. This meta-analysis aims to determine what it is about ECA participation that supports positive educational outcomes. Furthermore, this study challenges the theoretical assumptions about the benefits of participation in ECA. 29 studies (all except for one based on data collected in the United States) met the search criteria for inclusion in the analysis. Most effect sizes on academic achievements yielded from non-specific ECA, academic clubs and journalism were small, as were participation in performing arts, sports and leadership activities on a range of educational outcomes. Although the results show associations between participation in ECA and educational outcomes, causal effects could not be confirmed. It is concluded that the lack of evidence supporting the causal effects, and thus the common theoretical assumptions about the effects of ECA on educational outcomes, is due to methodology limitations in these studies.
Standley, 2008	This meta-analysis of 30 studies using a variety of music interventions to affect reading skills resulted in a moderately strong, significant, overall effect size of d = .32. When music activities incorporate specific reading skills matched to the needs of identified children (d = .44) or contingent music is used to reinforce reading behaviour (d = .66), benefits are large. The music activities that pair alphabet recognition with phonetic patterns, incorporate word segmentation and sound blending skills, and promote rapid decoding skills are effective in enhancing reading instruction and require little transfer to the assessment methodology. Benefits are greater when the special music reading activities are added to an existing music education curriculum than when replacing it. All schedules of intervention are equally effective regardless of whether daily, intense, short-term, or weekly periodic intervention spread across the school year.
Winner & Cooper, 2000	We searched exhaustively for all relevant studies (published in English) that appeared from 1950 to 1998. We were unable to find any experimental studies that provided a test of which causal mechanism might underlie academic improvement as a function of arts study. Thus, the research we review below tells us only (1) whether there is a correlation between arts study and academic achievement; and (2) whether academic achievement improves when students are exposed to the arts. The research that has been carried out on this question has not been designed to reveal the mechanism underlying any improvement shown by the experimental studies. Effect sizes ranged from r=.04 - r=.08. A mean effect size of r=.05 (d=.10) was found on overall or compostive academic achievement.

Aspiration interventions

Very low or no impact for moderate cost based on very limited evidence.

0

Aspirations are what children and young people hope to achieve for themselves in the future. Raising aspirations is often believed to be an effective way to motivate pupils to work harder so as to achieve the steps necessary for later success. A number of approaches to raising aspirations have been tried across three broad areas:

1. Interventions that focus on parents and families;
2. Interventions that focus on teaching practice;
3. Out-of-school interventions or extra-curricular activities, sometimes involving peers and mentors.

Approaches that seek to raise aspirations are very diverse and may aim to improve learners' self-esteem, self-efficacy or self-belief, or to develop motivation and engagement.

How effective is it?

On average, interventions which aim to raise aspirations appear to have little to no positive impact on educational attainment. This may seem counterintuitive – and it should be noted that the relationship between aspirations and attainment is complex and not fully understood – but there appear to be three main explanations.

First, evidence suggests that most young people actually have high aspirations, implying that much underachievement results not from low aspiration itself but from a gap between the aspirations that do exist and the knowledge and skills that are required achieve them. As a result, it may be more helpful to focus on raising attainment more directly in the first instance. Second, where pupils do have lower aspirations, it is not clear that any targeted interventions consistently succeed in raising their aspirations. Third, where aspirations begin low and are successfully raised by an intervention, it is not clear that an improvement in learning necessarily follows. In programmes which do raise attainment, it is unclear whether raising aspirations can be credited for the learning gains rather than the additional academic support or increased parental involvement.

How secure is the evidence?

Generally the evidence base on aspiration is very weak. More rigorous studies are required, particularly focusing on pupil-level rather than school-level interventions. There are two systematic reviews of aspiration interventions, some of which include quantitative data. These indicate that the relationship between aspirations and attainment is complex, but that there is no evidence of a clear causal connection between learning, changing aspirations, and attitudes to school.

There are no meta-analyses of interventions to raise aspirations that report impact on attainment or learning. This lack of evidence does not mean that impact is not achievable, but should make schools cautious as to how they make any investment of time or resources in this area.

The majority of studies come from the USA. There has been little robust research on the impact of aspiration interventions in English schools.

What are the costs?

Overall, the costs are estimated as moderate. Costs vary widely, and are hard to estimate precisely. After school programmes typically cost about £5 to £10 per session, so a 20-week programme once per week would cost a maximum of £200 per pupil. The costs of parental involvement programmes also vary, but are typically between £200 per child per year when the school covers the staffing costs, and up to about £850 per child per year for family support involving a full-time support worker. Mentoring approaches in

the USA have been estimated at $900 per student per year or about £560.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. The relationship between aspirations and attainment is not straightforward. In general, approaches to raising aspirations have not translated into increased learning.
2. A key reason for this may be that most young people have high aspirations for themselves. As a result, it is more important to keep these on track by ensuring that students have the knowledge and skills to progress towards them.
3. The attitudes, beliefs and behaviours that surround aspirations in disadvantaged communities are diverse so generalisations should be avoided.
4. Effective approaches almost always have a significant academic component, suggesting that raising aspirations in isolation will not be effective.
5. Have you considered how you will monitor the impact on attainment of any interventions or approaches?

Aspiration interventions Toolkit references

Full references
Cummings, C., Laing, K., Law, J., McLaughlin, J., Papps, I., Todd, L., & Woolner, P. (2012). Can Changing Aspirations And Attitudes Impact On Educational Attainment? York: Joseph Rowntree Foundation http://www.jrf.org.uk/publications/aspirations-att...
Doyle, M. & Griffin, M. (2012). Raised aspirations and attainment? A review of the impact of Aimhigher (2004–2011) on widening participation in higher education in England, London Review of Education, 10(1), 75-88. http://www.dx.doi.org/10.1080/14748460.2012.659060
Gorard, S., See, B. H., & Davies, P. (2012). The impact of attitudes and aspirations on educational attainment and participation. York: Joseph Rowntree Foundation http://www.jrf.org.uk/sites/files/jrf/education-yo...
Herrera, C., Grossman, J. B., Kauh, T. J., Feldman, A. F. and McMaken, J. (2007). Making a difference in schools: the Big Brothers Big Sisters school-based mentoring impact study. Philadelphia: Public/Private Ventures https://www.bigsister.org/bigsister/file/Making%20...
Kintrea, K, St.Clair, R., and Houston, M. (2011). The influence of parents, places and poverty on educational attitudes and aspirations. York: Joseph Rowntree Foundation. http://www.jrf.org.uk/publications/influence-paren...
Lupton, R., and Kintrea, K. (2011) Can community-based interventions on aspirations raise young people's attainment? Social Policy and Society, 10 (3), 321-335. http://dx.doi.org/10.1017/S1474746411000054
Strand, S., & Winston, J. (2008). Educational aspirations in inner city schools. Educational Studies, 34(4), 249-267. http://www.dx.doi.org/10.1080/03055690802034021

Summary of effects	
Study	***Effect size***
No meta-analyses of impact of raising aspirations on learning outcomes. (A systematic review, Cummings et al. (2012), report a range of effects on attainment 0.17 to 0.45 for parental involvement; 0.09 to 0.22 for mentoring and from 0.03 to 0.09 for extra-curricular activities. However these effects are associated with other influences on learning such as parental involvement in reading or academic mentoring.)	
Indicative effect size	**0.0**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Cummings et.al, (2012).	This review set out to establish whether there were interventions that could be scaled up to address the attainment gap for socio-economically disadvantaged children and young people by changing a particular set of attitudes. These attitudes were the aspirations to do well at school and to aim for advanced education, the sense that one's own actions can change one's life, and the giving of value to schooling and school results, referred to as aspirations, locus of control and valuing school.

Behaviour interventions

Moderate impact for moderate cost, based on extensive evidence.

+4

Behaviour interventions seek to improve attainment by reducing challenging behaviour, including aggression, violence, bullying, substance abuse and general anti-social activities. Three broad categories of behaviour interventions can be identified:

1. Universal programmes which seek to improve behaviour and generally take place in the classroom;
2. More specialised programmes which are targeted at students with either behavioural issues or behaviour and academic problems;
3. School level approaches to developing a positive school ethos or improving discipline which also aim to support greater engagement in learning. It should also be noted that other approaches, such as parental involvement programmes, are often associated with reported improvements in school ethos or discipline, but are not included in this summary which is limited to interventions that focus directly on behaviour (see instead Parental involvement).

How effective is it?

Evidence suggests that behaviour interventions can produce large improvements in academic performance along with a decrease in problematic behaviours, though estimated benefits vary widely across programmes. Effect sizes are larger for targeted interventions matched to specific students with particular needs or behavioural issues, than for universal interventions or whole school strategies.

The majority of studies report higher impact with older pupils. Different treatment approaches, such as behavioural, cognitive and social skills for aggressive and disruptive behaviour, seem to be equally effective. Parental and community involvement programmes are often associated with reported improvements in school ethos or discipline so are worth considering as alternatives to direct behaviour interventions.

School-level behaviour approaches are often associated with improvements in attainment, but the evidence of a causal link to learning is lacking. There is some anecdotal evidence about the benefits of reducing problematic behaviour of disruptive pupils on the attainment of their classmates, but this is an understudied dimension in evaluations of behaviour programmes.

How secure is the evidence?

Overall, it is clear that reducing challenging behaviour in schools can have a direct and lasting effect on pupils' learning. This is based on a number of meta-analyses based on randomised controlled studies of interventions in schools.

Some caution in interpreting findings is needed as the majority of the meta-analyses on behaviour interventions focus on pupils diagnosed with specific emotional or behavioural disorders. One meta-analysis of an anger management intervention shows a positive effect on behaviour but an overall negative effect on learning. This implies both that careful targeting and evaluation is important, and also that it is possible to reduce problematic behaviour without improving learning. Further research is also needed to investigate links between universal approaches to improving behaviour and learning

What are the costs?

There are no specific costs reported in the studies summarised here. Costs will be highly dependent on the type of intervention. Teacher-led behavioural interventions in the classroom are the least costly, but the least effective (estimated at £20 per pupil per year). One to one support is more expensive, but more effective (about £40 per hour, or £640 per pupil for 15 sessions). Overall, costs are estimated as moderate.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Targeted interventions for those diagnosed or at-risk of emotional or behavioural disorders produce the greatest effects.
2. Programmes of two to six months seem to produce more long-lasting results.
3. The wide variation in impact suggests that schools should look for programmes with a proven track record of impact.
4. Have you considered what training and professional development is required for the programmes?
5. Have you explored how to involve parents or communities in behaviour programmes? On average they show higher effects.

Behaviour interventions
Toolkit references

Full references
Brigman, C. & Campbell, C. (2003). Helping Students Improve Academic Achievement and School Success Behaviour. Professional School Counselling, 7, 91-98.
Chitiyo, M., Makweche-Chitiyo, P., Park, M., Ametepee, L.K. & Chitiyo, J. (2011). Examining the Effect of Positive Behaviour Support on Academic Achievement of Students with Disabilities. Journal of Research in Special Educational Needs, 11:3, 171-177.
Flay, B.R., Allred, C.G., & Ordway, N. (2001). Effects of the Positive Action Program on Achievement and Discipline: Two-Matched Control Comparisons. Prevention Science, 2:2, 71-89.
Gansle, K.A. (2005). The Effectiveness of School-Based Anger Interventions and Programs: A Meta-Analysis. Journal of School Psychology, 43, 321-341.
Gonzales, J. (2004). Rational Emotive Therapy With Children and Adolescents: A Meta-Analysis. Journal of Emotional and Behavioral Disorders Winter 2004 vol. 12 no. 4 222-235.
McEvoy, A., & Welker, R. (2000). Antisocial Behaviour, Academic Failure and School Climate: A Critical Review. Journal of Emotional and Behavioural Disorders, 8:3, 130- 140.
Quinn, M.M., Kavale, K.A., Mathur, S.R., Rutherford, R.B., Jr. & Forness, S.R. (1999). The Effectiveness of School-Based Anger Interventions and Programs: A Meta-Analysis. Journal of Emotional and Behavioural Disorders, 7:1, 54-64.
Reddy, L.A., Newman, E., De Thomas, C.A., Chun, V. (2009). Effectiveness of School-Based Prevention and Intervention Programmes for Children and Adolescents with Emotional Disturbance: A Meta-Analysis. Journal of School Psychology, 47, 77-99.
Sander, J.P., Patall, E.A., Amoscato, L.A., Fisher, A.L., & Funk, C. (2012). A Meta-Analysis of the Effects of Juvenile Delinquency Interventions on Academic Outcomes. Children and Youth Services Review, 34, 1695-1708.
Solomon, B.G., Klein, S.A., Hintze, J.M., Cressey, J.M., & Peller, S.L. (2012). A Meta-Analysis of School-Wide Positive Behaviour Support: An Explanatory Study using Single-Case Synthesis. Psychology in the Schools, 49:2, 105-121.
Wilson, S.J., & Lipsey, M.W. (2007). School-Based Interventions for Aggressive and Disruptive Behaviour. Update of a Meta-Analysis. American Journal of Preventive Medicine, 33, 130-143.

Behaviour interventions
Toolkit references

Summary of effects	
Study	***Effect size***
Chitiyo et al. 2011 (positive behaviour support for pupils with disabilities)	0.87 (on academic achievement)
Gansle, 2005 (anger management)	-0.11 (on academic outcomes)
Gonzalez et al. 2004 (rational emotive therapy)	0.49 (on GPA)
Losel & Beelmann, 2003	0.38 (all outcomes)
Quinn et al. 1999 (emotional disorder)	0.05 (on academic achievement)
Reddy et al. 2009 (emotional disturbance -intervention)	1.78 (on general academic skills)
Reddy et al. 2009 (emotional disturbance -prevention)	0.28 (on general academic skills)
Sander et al. 2012 (juvenile delinquency)	0.02 (on academic achievement)
Wilson & Lipsey, 2007 (aggressive and disruptive)	0.22 (on school performance)
Weighted mean effect size	**0.30**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Chitiyo et al., 2011	Students who engage in challenging behaviour compromise the fundamental ability of schools to educate children. Consequently, teachers face the daunting task of designing effective strategies to promote positive educational outcomes for their students. Since the 1997 Individuals with Disabilities Education Act amendments, the use of positive behaviour supports (PBS) to address the behavioural needs of children challenged by disabilities has expanded. There is evidence to support the utility of PBS in reducing challenging behaviour among students. However, successful schools are also gauged by the academic achievement of their students. Hence, it is important to examine the extent to which behavioural outcomes are related to academic outcomes. The purpose of this paper is to examine the extent to which PBS interventions aimed at reducing challenging behaviour result in corresponding improvement in academic achievement. A meta-analysis of extant research indicated a positive correlation of 0.40 between improvement in problem behaviour and academic achievement.
Gansle, 2005	Twenty peer-reviewed journal articles that described outcomes of interventions that took place in school settings and either focused on anger or included anger as a dependent variable were meta-analysed. No differences in outcomes were found for group comparisons by school setting, special education status, entrance criteria, or treatment agents. Although 60% of articles discussed its importance, only two articles actually measured treatment integrity. Across outcomes, the weighted mean effect size of the interventions post treatment was determined to be .31. The largest effects were found for anger and externalizing behaviours, internalizing, and social skills, with mean effect sizes of .54, .43, and .34 respectively. Weighted mean effect sizes for follow-up studies were also calculated, but given the small number of studies that reported follow-up effects, those must be interpreted with caution. The results of this meta-analysis are discussed as they relate to research, practice, and intervention with children.
Gonzalez et al., 2004	This article systematically reviews the available research on rational emotive behavioural therapy (REBT) with children and adolescents. Meta-analytic procedures were applied to 19 studies that met inclusion criteria. The overall mean weighted effect of REBT was positive and significant. Weighted zr effect sizes were also computed for five outcome categories: anxiety, disruptive

	behaviours, irrationality, self-concept, and grade point average. In terms of magnitude, the largest positive mean effect of REBT was on disruptive behaviours. Analyses also revealed the following noteworthy findings: (a) there was no statistical difference between studies identified low or high in internal validity; (b) REBT appeared equally effective for children and adolescents presenting with and without identified problems; (c) non-mental health professionals produced REBT effects of greater magnitude than their mental health counterparts; (d) the longer the duration of REBT sessions, the greater the impact, and (e) children benefited more from REBT than adolescents. The findings are discussed in terms of several important limitations along with suggestions for future research.
Quinn et al., 1999	Many programs designed for youth with Emotional or Behavioural Disorders (EBD) include a social skill training component. Using quantitative methods of meta-analysis, the finding from 35 studies investigating the effects of social skills interventions for students with EBD were synthesized. The pooled mean effect size (ES) was 0.199 from which the average student with EBD would be expected to gain a modest eight percentile ranks on outcome measures after participating in a social skill training program. Studies were further grouped and analysed according to different variables (e.g. similarities of the intervention, participants and assessment procedures). Slightly greater ES were found for interventions focused on teaching and measuring specific social skills (e.g. cooperating or social problem solving) compared to more global interventions. Several pertinent issues for reviewing the results of this research synthesis are addressed.
Reddy et al., 2009	The present study evaluated the effectiveness of school-based prevention and intervention programs for children and adolescents at-risk for and with emotional disturbance. Published outcome studies (k=29) from December, 1988, to March, 2006, including 1405 children and adolescents were reviewed. Each investigation was coded on several variables describing the child, parent, and teacher samples, as well as reported outcome results. The overall mean weighted effect size was 1.00 at post-test and 1.35 at follow-up. Mean weighted ESs were 0.42 for between-subjects design studies, 0.87 for within-subjects design studies, and 1.87 for single-subject design studies. Prevention programs yielded a mean weighted ES of 0.54 and intervention programs produced a mean weighted ES of 1.35. Findings for specific outcome foci are presented and implications are discussed.
Sander et al., 2012	This meta-analysis examined the effects of juvenile delinquency interventions on academic outcomes. After retrieving over 250 reports, 15 reports met inclusion criteria and provided 134 effect sizes (92 unadjusted and 42 adjusted) based on 20 separate samples in a variety of settings, including school, community, and juvenile justice settings. Heterogeneity of the samples, generally weak research designs, and the absence of control conditions in many recovered reports was a limitation in the existing research. Overall, there were limited positive effects of juvenile delinquency interventions on academic outcomes. The lack of theory driven or empirically supported academic interventions was notable. Studies with the weakest designs produced the largest effects on academic achievement, and school attendance outcomes were enhanced only for older adolescents. The implications of findings for future research and policy are discussed.
Wilson & Lipsey, 2007	Research about the effectiveness of school-based psychosocial prevention programs for reducing aggressive and disruptive behaviour was synthesized using meta-analysis. This work updated previous work by the authors and further investigated which program and student characteristics were associated with the most positive outcomes. Two hundred forty-nine experimental and quasi-experimental studies of school-based programs with outcomes representing aggressive and/or disruptive behaviour were obtained. Effect sizes

	and study characteristics were coded from these studies and analysed. Positive overall intervention effects were found on aggressive and disruptive behaviour and other relevant outcomes. The most common and most effective approaches were universal programs and targeted programs for selected/indicated children. The mean effect sizes for these types of programs represent a decrease in aggressive/disruptive behaviour that is likely to be of practical significance to schools. Multicomponent comprehensive programs did not show significant effects and those for special schools or classrooms were marginal. Different treatment modalities (e.g., behavioural, cognitive, social skills) produced largely similar effects. Effects were larger for better-implemented programs and those involving students at higher risk for aggressive behaviour. Schools seeking prevention programs may choose from a range of effective programs with some confidence that whatever they pick will be effective. Without the researcher involvement that characterizes the great majority of programs in this meta-analysis, schools might be well-advised to give priority to those that will be easiest to implement well in their settings.

Block scheduling

Very low or no impact for very low or no cost, based on limited evidence.

0

Block scheduling is an approach to school timetabling in secondary schools. It typically means that pupils have fewer classes (4-5) per day, for a longer period of time (70-90 minutes). The three main types of block schedules found in the research are:

4x4 block scheduling: 4 blocks of extended (80–90 minute) classes each day, covering the same 4 subjects each day. Students take 4 subjects over 1 term, and 4 different subjects in the following term.A/B block scheduling: 3 or 4 blocks of extended (70–90 minute) classes each day, covering the same 3 or 4 subjects on alternating days. Students take 6 or 8 subjects each term.Hybrid: a hybrid of traditional models and 3/4-class-per-day approaches. Students have 5 classes per day, of between 60 and 90 minutes.

How effective is it?

There is no consistent pattern in the evidence. A 2010 systematic review concluded that the 4x4 pattern seemed to produce higher overall achievement than traditional schedules, though this may mask differences between subjects. More detailed analysis suggests that in science the A/B block scheduling approach resulted in higher results than traditional schedules (two to five months of additional progress). In mathematics and English the evidence was unclear with studies showing both better and worse results for any type of block scheduling compared with traditional scheduling.

The evidence suggests that how teachers use the time they are allocated is more important than the length of lesson or the schedule of lessons, and hence that the introduction of block scheduling is unlikely to raise attainment by itself. It may also be that when different timetable patterns are introduced, the changes will only be beneficial if teachers alter the way they teach to get the best from the time allocation. Teachers and students often perceive that timetabling changes are beneficial, especially when it appears to increase one to one interaction. However, these perceptions are not clearly linked with improved learning outcomes.

How secure is the evidence?

There are two recent meta-analyses which have looked at the evidence of the impact of timetabling and scheduling changes on students' learning but these rely on a small number of studies which have limited security.

Timetabling mainly affects secondary schools, though the time spent on different areas of the curriculum is also relevant at primary level. The research has mainly looked at impact on mathematics, English and science.

What are the costs?

The costs of making alterations to the timetable are mainly in terms of organisational effort and time and involve minimal financial outlay.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Timetabling changes alone are not sufficient to improve learning.
2. Teachers need to alter the way that they teach, and should plan and organise different kinds of learning activities to obtain benefits.
3. Have timetabling changes been matched to curriculum goals and teaching and learning objectives (such as longer lessons for science experiments)?
4. Have you considered how longer lessons may provide opportunities for other promising approaches, such as improving the amount of feedback that students get from the teacher or from each other?

Technical Appendix: Block scheduling

Definition

Block scheduling is one approach to school timetabling in secondary schools. It typically means that pupils have fewer classes (4-5) per day, for a longer period of time (70-90 minutes). The three main types of block schedules found in the research are:

- 4x4 block scheduling: 4 blocks of extended (80–90 minute) classes each day, covering the same 4 subjects each day. Students take 4 subjects over 1 term, and 4 different subjects in the following term.
- A/B block scheduling: 3 or 4 blocks of extended (70–90 minute) classes each day, covering the same 3 or 4 subjects on alternating days. Students take 6 or 8 subjects each term.
- Hybrid: a hybrid of traditional models and 3/4-class-per-day approaches. Students have 5 classes per day, of between 60 and 90 minutes.

Block scheduling and timetabling changes refer to alterations to lessons within the existing length of the school day, rather than approaches which seek to extend the school day or the school year (see the 'Extending School Time' Toolkit strand for an overview of these other approaches).

Search terms: school timetabling; timetable alternation; block scheduling.

Evidence rating

There are two meta-analyses published within the last 10 years drawing on a total of 46 studies, which have looked at the quantitative evidence of the impact of timetabling and scheduling changes on students' learning. There are some further correlational studies that investigate this topic. Effects overall tend to be small. Timetabling is mainly an issue for secondary schools, though the time spent on different areas of the curriculum is also relevant at primary level. The research has mainly looked at impact on mathematics, English and science. The small number of underlying studies, the variation in focus and the overall quality of the underlying studies meant that neither meta-analysis reported an overall pooled effect. Overall the evidence is therefore limited.

Additional cost information

The costs of making alterations to the timetable are mainly in terms of organisational effort and time and involve minimal financial outlay.

References

Full references
*Dickson K., Bird K., Newman M. & Kalra N. (2010). *Effect of Block Scheduling on Academic Achievement in Secondary Schools: A Systematic Review of Evidence*. The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), Institute of Education, University of London. http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2476
Dexter, K. M., Tai, R. H., & Sadler, P. M. (2006). Traditional and block scheduling for college science preparation: A comparison of college science success of students who report different high school scheduling plans. *The High School Journal*, 89(4), 22-33. http://dx.doi.org/10.1353/hsj.2006.0007
Dostal, J. R. (2010). *Alternative scheduling models and their effect on science achievement at the high school level (*Doctoral dissertation, University of Nebraska). http://www.dostal.khs.kearneypublicschools.org/modules/groups/homepagefiles/profile/771434/71601

/File/Dissertation-FINAL%20DRAFT.pdf
Gruber, C.D. & Onwuegbuzie, A.J. (2001). Effects of Block Scheduling on Academic Achievement among High School Students. *The High School Journal,* 84.4, 32-42. http://dx.doi.org/10.1353/hsj.2001.0010
Lewis, C.W., Winokur, M.A., Cobb, R.B., Gliner, G.S. & Schmidt, J. (2005). Block Scheduling in the High School Setting: A Synthesis of Evidence-Based Research*. A report prepared for MPR Associates, Inc., Berkeley, CA.
Trenta, L., & Newman, I. (2002). Effects of a high school block scheduling program on students: A four-year longitudinal study of the effects of block scheduling on student outcome variables. *American Secondary Education*, 54-71. http://www.jstor.org/stable/41064590
Veal, W.R. & Flinders, D.J. (2001). How Block Scheduling Reform Effects Classroom Practice. *The High School Journal,* 84.4 pp 21-31. http://dx.doi.org/10.1353/hsj.2001.0012
Zepeda, S.J. & Mayers, R.S. (2006). An Analysis of Research on Block Scheduling. *Review of Educational Research*, 76.1 pp 137-170. http://dx.doi.org/10.3102/00346543076001137

* Studies marked with an asterisk are included in the summary of effects

Summary of effects		
Meta-analyses	***FSM effect size***	***Overall effect size***
Dickson et al. 2010 on achievement		0.11,
Dickson et al. 2010 on mathematics		-0.02
Dickson et al. 2010 science		0.20
Lewis et al, 2005 on mathematics		-0.10
Lewis et al, 2005 on English		-0.17
Lewis et al, 2005 on science		-0.12
Weighted mean effect size		**0.00**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Dickson et al. 2010	Block scheduling is one approach to school scheduling. It typically means that students have fewer classes (4-5) per day, for a longer period of time (70-90 minutes). There are three main types of block schedule investigated in this review, comprising the following: • **4 x 4:** four blocks of 80–90 minute classes in one day, with students taking four subjects in one term • **A/B:** classes of 70-90 minutes each for 3/4 different subjects on every alternating day • **hybrid:** five classes per day, between 55 and 90 minutes in length The in-depth review asks the following: Does block scheduling result in higher levels of student attainment than traditional scheduling? Studies used different measures of academic achievement across different academic subjects. These included test results in Mathematics, English, Science, exam scores or average grade scores across different subjects. Sub-questions were also asked in the in-depth review and these investigated whether the effect of block scheduling varied by type of block schedule and type of subject(s) taught. Only 12 of the 14 studies included in the in-depth review provided the data necessary for statistical meta-analysis to assess the effectiveness of different types of block scheduling on academic achievement. The 12 studies were considered to be of medium weight of evidence and two were considered to be of low weight of evidence, overall, for this review. Where we were able to combine data to

	produce summary effect sizes, we found that 4 x 4 block scheduling resulted in higher cross-subject achievement than traditional schedules. However, the outcome average cross-subject achievement could conceal worsening performance in some subjects and better performance in others. For single subject outcomes: In Science, A/B block scheduling resulted in higher results than traditional schedules. In Mathematics and English, the evidence was unclear, with studies showing both better and worse results for block scheduling compared with traditional scheduling. There is not conclusive evidence in this review to support the introduction of policy guidance on the use of block scheduling in secondary schools. Findings do not indicate that participating in block schedules would produce negative outcomes for pupils across subjects, but the findings on positive effects are not strong enough to recommend their implementation.
Lewis et al. 2005	The purpose of this study was to produce a systematic review and synthesis of evidence based research on the effect of block scheduling on student achievement in United States High-schools. This report provides a brief introduction to block scheduling, chronicles the search strategies used to locate the final literature set, and describes the processes employed to code the studies on outcome, intervention, and methodological criteria using the What Works Clearinghouse (WWC) framework. In addition, findings, conclusions, and recommendations are discussed for the studies that merited inclusion into the block scheduling evidence base.

Collaborative learning

Moderate impact for very low cost, based on extensive evidence.

+5

Collaborative or cooperative learning can be defined as learning tasks or activities where students work together in a group small enough for everyone to participate on a collective task that has been clearly assigned. This can be either a joint task where group members do different aspects of the task but contribute to a common overall outcome, or a shared task where group members work together throughout the activity.

Some collaborative learning approaches also get mixed ability teams or groups to work in competition with each other, in order to drive more effective collaboration. There is a very wide range of approaches to collaborative and cooperative learning involving different kinds of organisation and tasks, but this summary does not include Peer tutoring, which is reviewed separately.

How effective is it?

The impact of collaborative approaches on learning is consistently positive, but it does vary so it is important to get the detail right. Effective collaborative learning requires much more than just sitting pupils together and asking them to work together; structured approaches with well-designed tasks lead to the greatest learning gains. There is some evidence that collaboration can be supported with competition between groups, but this is not always necessary, and can lead to learners focusing on the competition rather than the learning it aims to support. Approaches which promote talk and interaction between learners tend to result in the best gains.

How secure is the evidence?

Evidence about the benefits of collaborative learning has been found consistently for over 40 years and a number of systematic reviews and meta-analyses of research studies have been completed. In addition to direct evidence from research into collaborative learning approaches, there is also indirect evidence where collaboration has been shown to increase the effectiveness of other approaches such as mastery learning or digital technology. It appears to work well for all ages if activities are suitably structured for learners' capabilities and positive evidence has been found across the curriculum. Not all of the specific approaches to collaborative learning that are adopted by schools have been evaluated so it is important to evaluate any new initiative in this area.

What are the costs?

The direct costs involved are very low, though professional development is advisable. Estimated costs for a class of 25 pupils are about £500 or £20 per pupil per year, plus the costs of monitoring and evaluating impact of adopting the approach. Overall the costs are estimated as very low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Pupils will need support and practice to work together; this does not happen automatically.
2. Tasks need to be designed carefully so that working together is effective and efficient, otherwise some pupils will try to work on their own.
3. Competition between groups can be used to support pupils in working together more effectively within their group, though over-use of competition can focus learners on the competition rather than succeeding in their learning so it needs to be used cautiously.
4. It is particularly important to encourage lower achieving pupils to talk and articulate their thinking in collaborative tasks as they may contribute less.
5. Have you considered the professional development required to support effective use of these approaches?

Collaborative learning Toolkit references

Education Endowment Foundation

Full references
Blatchford, P., Kutnick, P., Baines, E., & Galton, M. (2003). Toward a social pedagogy of classroom group work. International Journal of Educational Research, 39(1), 153-172.
Cohen, E.G. (1993). Restructuring the Classroom: Conditions for Productive Small Groups. Review of Educational Research, 64:1, 1-35.
Gillies, R. M. (2003). Structuring cooperative group work in classrooms. International Journal of Educational Research, 39(1), 35-49.
Gillies, R. M., & Boyle, M. (2010). Teachers' reflections on cooperative learning: Issues of implementation. Teaching and Teacher Education, 26(4), 933-940.
Igel, C. C. (2010). The effect of cooperative learning instruction on K-12 student learning: A meta-analysis of quantitative studies from 1998 to 2009. PhD Thesis presented to the Faculty of the Curry School of Education University of Virginia (UMI No. AAT 3435906).
Johnson, D.W., Maruyama, G., Johnson, R., & Nelson, D. (1981). Effects of Cooperative, competitive and individualistic goal structures on Achievement: A meta-analysis. Psychological Bulletin, 89:1, 47-62.
Kyndt, E., Raes, E., Lismont, B., Timmers, F., Dochy, F., & Cascallar, E. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings?. Educational Research Review. 10:133-149.
Kutnick, P., & Blatchford, P. (2014). Groups and Classrooms. In Effective Group Work in Primary School Classrooms (pp. 23-49). Springer Netherlands.
Nunnery, J. A., Chappell, S., & Arnold, P. (2013). A meta-analysis of a Cooperative Learning Models Effects on Student Achievement in Mathematics. Cypriot Journal of Educational Sciences, 8(1), 34-48.
Romero C.C. (2009). Cooperative learning instruction & science achievement for secondary and early post-secondary students: A systematic Review. Dissertation, Colorado State University.
Slavin, R. E. (1992). When and why does cooperative learning increase achievement? Theoretical and empirical perspectives. In R. Hertz-Lazarowitz & N. Miller (Eds.), Interaction in cooperative groups: The theoretical anatomy of group learning (pp. 145–173). New York: Cambridge Univerity Press.
Stoner, D. A. (2004). The Effects of Cooperative Learning Strategies on Mathematics Achievement Among Middle-grades Students: A meta-analysis (Doctoral dissertation, University of Georgia).

Summary of effects	
Study	***Effect size***
Igel, 2010	0.44
Johnson et.al. 1981 (co-op v individualistic)	0.78
Johnson et.al. 1981 (co-op v competitive)	0.78
Johnson et.al. 2000 (learning together)	0.91
Johnson et.al. 2000 (group investigation)	0.62
Johnson et.al. 2000 (academic controversy)	0.86
Johnson et.al. 2000 (jigsaw groups)	0.09
Johnson et.al. 2000 (student-team achievement)	0.28
Johnson et.al. 2000 (cooperative integrated read & composition)	0.18
Johnson et.al. 2000 (team assisted individualization)	0.19
Kyndt et al. 2013	0.54
Nunnery, Chappell & Arnold, 2013	0.16
Puzio & Colby, 2013 (on reading comprehension)	0.20
Romero, 2009	0.40
Stoner, 2004	0.13
Weighted mean effect size	**0.41**

Collaborative learning Toolkit references

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses and abstracts	
Study	***Abstract***
Igel, 2010	Cooperative instruction is one of the most theoretically-grounded, popular, and misunderstood of the instructional strategies. Grounded within social-psychology and learning theory, properly specified cooperative instruction requires design elements such as positive interdependence and individual accountability that go beyond basic group-mediated instruction. Despite its popularity and a large corpus of literature, practitioners and researchers alike often confuse cooperative instruction with less stringent forms of group-mediated instruction. The present study clarifies this distinction, and meta-analyzes the results of twenty rigorous studies on the effect of cooperative interventions on K-12 student learning. The meta-analysis employs rigorous selection criteria to maintain internal validity and newly developed statistical adjustments to account for analytic errors found throughout much of the primary research base. Findings reveal a moderate overall effect (0.44) for cooperative interventions with differential estimates across a range of moderators. These finding are placed within the context of the larger corpus of research on cooperative learning and its implications for practitioners discussed.
Johnson et.al., 1981	We reviewed 122 studies and compared the relative effectiveness of cooperation, cooperation with intergroup competition, interpersonal competition, and individualistic goal structures in promoting achievement and productivity in North American samples. These studies yielded 286 findings. Three meta-analysis procedures were used: voting method, effect-size method, and z-scores method. The results of the meta-analyses indicate (a) that cooperation is considerably more effective than interpersonal competition and individualistic efforts, (b) that cooperation with intergroup competition is also superior to interpersonal competition and individualistic efforts, and (c) that there is no significant difference between interpersonal competitive and individualistic efforts. Through multiple regression, a number of potentially mediating variables for these results are identified.
Johnson et.al. 2000	Cooperative learning is one of the most widespread and fruitful areas of theory, research, and practice in education. Reviews of the research, however, have focused either on the entire literature which includes research conducted in non-educational settings or have included only a partial set of studies that may or may not validly represent the whole literature. There has never been a comprehensive review of the research on the effectiveness in increasing achievement of the methods of cooperative learning used in schools. An extensive search found 164 studies investigating eight cooperative learning methods. The studies yielded 194 independent effect sizes representing academic achievement. All eight cooperative learning methods had a significant positive impact on student achievement. When the impact of cooperative learning was compared with competitive learning, Learning Together (LT) promoted the greatest effect, followed by Academic Controversy (AC), Student-Team- Achievement-Divisions (STAD), Teams-Games-Tournaments (TGT), Group Investigation (GI), Jigsaw, Teams-Assisted-Individualization (TAI), and finally Cooperative Integrated Reading and Composition (CIRC). When the impact of cooperative lessons was compared with individualistic learning, LT promotes the greatest effect, followed by AC, GI, TGT, TAI, STAD, Jigsaw, and CIRC. The consistency of the results and the diversity of the cooperative learning methods provide strong validation for its effectiveness.
Kyndt et al., 2013	One of the major conclusive results of the research on learning in formal learning settings of the past decades is that cooperative learning has shown to evoke clear positive effects on different variables. Therefore this meta-analysis has two principal aims. First, it tries to replicate, based on recent studies, the research about the main effects of cooperative learning on three categories of outcomes: achievement, attitudes and

	perceptions. The second aim is to address potential moderators of the effect of cooperative learning. In total, 65 articles met the criteria for inclusion: studies from 1995 onwards on cooperative learning in primary, secondary or tertiary education conducted in real-life classrooms. This meta-analysis reveals a positive effect of cooperative learning on achievement and attitudes. In the second part of the analysis, the method of cooperative learning, study domain, age level and culture were investigated as possible moderators for achievement. Results show that the study domain, the age level of the students and the culture in which the study took place are associated with variations in effect size.
Nunnery, Chappell & Arnold, 2013	This study synthesizes the mathematics achievement impacts observed in randomized Studies of the Student Teams Achievement Divisions cooperative learning model. A total of 15 randomized studies were retrieved from the extant literature. Analyses of d= +0.16. Cohen's d effect size estimates indicated an overall Statistically significant positive effect of. Effect size Estimates were also examined for between- class Heterogeneity to ascertain whether there were Differences in effects for younger children in Elementary settings versus adolescent children in Secondary settings. These analyses indicated That cooperative learning had a much stronger effect on student achievement for adolescent children than for younger children.
Puzio & Colby, 2013	We conducted a meta-analysis on the effectiveness of cooperative and collaborative learning to support enhanced literacy outcomes. Interventions considered were provided in regular education settings (i.e., not pull-out instruction) with students from Grades 2 through 12. Reviewing more than 30 years of literacy research, we located 18 intervention studies with 29 study cohorts. Included studies primarily used standardized assessments to report on students' reading, vocabulary, or comprehension achievement, which we analyzed separately. Overall, students had significantly higher literacy achievement scores when instructional interventions utilized cooperative and collaborative activity structures. The overall weighted mean effect sizes ranged from 0.16 to 0.22 ($p < .01$) with more than 94% of the point estimates being positive. Because cooperative or collaborative learning was always one of multiple intervention components, it was impossible to estimate the unique, added effects of cooperative and/or collaborative learning. Although the small number of eligible studies precludes any claims about the effectiveness of specific forms of grouping and the circumstances under which programs have more impact, our findings suggest that cooperative and collaborative grouping was a core component of effective literacy interventions, particularly at the elementary level.
Romero, 2009	A systematic review of 2,506 published and unpublished citations identified in a literature search on science outcomes associated with cooperative learning in secondary and early post-secondary science classrooms between 1995 and 2007 was conducted. The goal of this review was to determine what impact cooperative learning had on science achievement of students compared to traditional instruction. A tri-level screening and coding process was implemented and identified 30 original, empirical studies that met the inclusionary criteria while yielding an overall effect size estimate. The minimum methodological criteria for inclusion were as follows: (a) the study utilized a treatment/control design, (b) cooperative learning was the intervention, and the control group experienced traditional instruction, (c) the subjects in included studies were secondary or early-post-secondary students, (d) the study was performed in a science classroom, and (e) student achievement was the outcome measure. This meta-analysis describes the main effect of cooperative learning; additionally, a variety of moderator analyses were conducted in order to determine if particular study and participant characteristics influenced the effect of the intervention. The results of this review indicate that cooperative learning improves student achievement in science. The overall mean effect size was .308, a medium effect (Cohen, 1988). Moderator analyses on study participant characteristics gender and ability level were inconclusive based on the small number of studies in which data on these characteristics were disaggregated. If the intervention was structured

	in a particular fashion, the effect on student achievement was greater than that for an unstructured intervention. The intervention showed a greater effect on student achievement in biology classes than in other science disciplines. Studies performed using cluster randomized or quasi-experimental without subject matching methodologies showed a greater effect on student achievement in science than studies that used the quasi-experimental with subject matching methodology. Implications for teacher education policy and recommendations for improvements in methodological practices and reporting are given.
Stoner, 2004	The purpose of this study was to examine the existing body of literature and through the use meta-analysis determine the effect of cooperative learning strategies on the mathematics achievement of middle-grades students, grades 4-8. A collection of 25 quantitative studies produced an effect size which indicated that cooperative learning strategies have a positive effect] on the mathematics achievement of middle-grades students. Through correlational analysis, the current study examined relationships between the duration of the studies and effect size of the studies. Also examined was the duration of the studies and grade 4 and grade 8 NAEP mathematics proficiency scores for 2003. Correlation Tables as well as scatter plots for each correlation were provided for visual examination. Also examined were the location of the studies; the particular method of data analysis that each study used; and the dependent outcome measure of each of the studies. Conclusions and recommendations for further research were provided.

Digital technology

Moderate impact for moderate cost, based on extensive evidence.

+4

The use of digital technologies to support learning. Approaches in this area are very varied, but a simple split can be made between:

1. Programmes for students, where learners use technology in problem solving or more open-ended learning, and
2. Technology for teachers such as interactive whiteboards or learning platforms.

How effective is it?

Overall, studies consistently find that digital technology is associated with moderate learning gains (on average an additional four months). However, there is considerable variation in impact. Evidence suggests that technology should be used to supplement other teaching, rather than replace more traditional approaches. It is unlikely that particular technologies bring about changes in learning directly, but different technology has the potential to enable changes in teaching and learning interactions, such as by providing more effective feedback for example, or enabling more helpful representations to be used or simply by motivating students to practise more.

There is some evidence that it is more effective with younger learners and studies suggest that individualising learning with technology (one to one laptop provision, or individual use of drill and practice) may not be as helpful as small group learning or collaborative use of technology. There is clear evidence that it is more beneficial for areas like writing rather than spelling or mathematics practice rather than problem solving.

How secure is the evidence?

There is extensive evidence across age groups and for most areas of the curriculum which shows positive impact on learning. However, the variation in effects and the range of technologies available suggest that it is important to evaluate the impact on learning when technology is used. The pace of technological change means that evidence is usually about yesterday's technology rather than today's, but average impacts have remained consistent for some time, implying that general messages are likely to remain relevant.

What are the costs?

The costs of investing in new technologies are high, but they are already part of the society we live in and most schools are already equipped with computers and interactive whiteboards. The evidence suggests that schools rarely take into account or budget for the additional training and support costs which are likely to make the difference to how well the technology is used. Expenditure is estimated at £300 per pupil for equipment and technical support and a further £500 per class (£20 per pupil) for professional development and support. Costs are therefore estimated as moderate.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Effective use of technology is driven by learning and teaching goals rather than a specific technology: technology is not an end in itself.
2. Are you clear about how you expect the introduction of technology to improve learning? New technology does not automatically lead to increased attainment.
3. Technology should support pupils to work harder, for longer or more efficiently to improve their learning.
4. Motivation to use technology does not always translate into more effective learning, particularly if the use of the technology and the learning outcomes are not closely aligned.
5. Teachers need support and time to learn to use new technology effectively. This involves more than just learning how to use the technology; it should include support to understand how it can be used for learning.

Digital technology Toolkit references

Full references
Bayraktar S. (2000). A Meta-Analysis of the Effectiveness of Computer Assisted Instruction in Science Education. Journal of Research on Technology in Education, 42.2: 173-188.
Camnalbur & Erdogan (2010). A Meta Analysis on the Effectiveness of Computer-Assisted Instruction: Turkey Sample, 8: 497-505.
Cheung & Slavin (2011). The Effectiveness of Education Technology for Enhancing Reading Achievement: A Meta-Analysis. Best Evidence Encyclopedia. www.bestevidence.org.
Crook, C., Harrison, C., Farrington-Flint, L., Tomás, C., Underwood, J. (2010). The Impact of Technology: Value-added classroom practice Final report Coventry: Becta.
Higgins, S. (2003). Does ICT Improve Learning and Teaching in Schools? Nottingham: British Educational Research Association.
Kulik, J. (2003). Effects of Using Instructional Technology in Elementary and Secondary Schools: What Controlled Evaluation Studies Say Arlington, VA: SRI International.
Liao, Y.C. (2007). Effects of Computer-Assisted Instruction on Students' Achievement in Taiwan: A Meta-Analysis. Computers and Education, 48.2 pp 216-233.
Li & Ma (2010). A Meta-analysis of the Effects of Computer Technology on School Students' Mathematics Learning. Educational Psychology Review, 22.3: 215-243.
Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. US Department of Education.
Moran, J., Ferdig, R. E., Pearson, P. D., Wardrop, J., & Blomeyer, R. L. (2008). Technology and reading performance in the middle-school grades: A meta-analysis with recommendations for policy and practice. Journal of Literacy Research, 40(1), 6-58.
Morphy P. & Graham S. (2012). Word Processing Programs and Weaker Writers/Readers: A Meta-Analysis of Research Findings, Reading and Writing, 25, 641-678.
Onuoha, C. O. (2007). Meta-analysis of the effectiveness of computer-based laboratory versus traditional hands-on laboratory in college and pre-college science instructions.
Pearson, D.P., Ferdig, R.E., Blomeyer, R.L. & Moran, J. (2005). The Effects of Technology on Reading Performance in the Middle-School Grades: A Meta-Analysis With Recommendations for Policy Naperville, Il: University of Illinois/North Central Regional Educational Laboratory.
Rosen, Y., & Salomon, G. (2007). The differential learning achievements of constructivist technology-intensive learning environments as compared with traditional ones: A meta-analysis. Journal of Educational Computing Research,36(1), 1-14.
Sandy-Hanson A.E. (2006). A Meta-Analysis of the Impact of Computer Technology versus Traditional Instruction on Students in Kindergarten through Twelfth Grade in the United States: A Comparison of Academic Achievement, Higher Order Thinking Skills, Motivation, Physical Outcomes and Social Skills. Doctoral Dissertation, Howard University
Seo, Y. J., & Bryant, D. P. (2009). Analysis of studies of the effects of computer-assisted instruction on the mathematics performance of students with learning disabilities. Computers & Education, 53(3), 913-928.
Strong, G. K., Torgerson, C. J., Torgerson, D., & Hulme, C. (2011). A systematic meta- analytic review of evidence for the effectiveness of the 'Fast ForWord'language intervention program. Journal of Child Psychology and Psychiatry, 52(3), 224-235.
Tamim R.M., Bernard R.M., Borokhovski E., Abrami P.C., & Schmid R.F. (2011). What Forty Years of Research Says about the Impact of Technology on Learning: A Second-Order Meta-Analysis and Validation Study. Review of Educational Research, 81, 4-28.
Tokpah, C. L. (2008). THE EFFECTS OF COMPUTER ALGEBRA SYSTEMS ON STUDENTS'ACHIEVEMENT IN MATHEMATICS (Doctoral dissertation, Kent State University).
Torgerson, C.J. & Elbourne, D. (2002). A systematic review and meta-analysis of the effectiveness of information and communication technology (ICT) on the teaching of spelling. Journal of Research in Reading, 25, 129-143.
Torgerson, C, & Zhu, D. (2003). A systematic review and meta-analysis of the effectiveness of ICT on literacy learning in English, 5-16. In: Research Evidence in Education Library. London: EPPI-Centre,

Social Science Research Unit, Institute of Education.
Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. Journal of Educational Computing Research, 34(3), 229-243.
Waxman, H., Connell, M., & Gray, J. (2002). A quantitative synthesis of recent research on the effects of teaching and learning with technology on student outcomes.
Waxman, H.C., Lin, M. & Michko, G.M. (2003). A Meta-Analysis of the Effectiveness of Teaching and Learning With Technology on Student Outcomes A Meta-Analysis of the Effectiveness of Teaching and Learning With Technology on Student Outcomes December 2003. Technology. Naperville, Illinois.

Summary of effects	
Study	***Effect size***
Bayraktar, 2000 (science)	**0.27**
Camnalbur & Erdogan , 2010 (in Turkey)	**1.05**
Cheung & Slavin, 2011(on mathematics	**0.15**
Christmann & Badgett, 2003	**0.34**
Li & Ma 2010 (on mathematics)	**0.71**
Liao, 2007 (in Taiwan)	**0.55**
Pearson, 2005 (on reading)	**0.49**
Sandy-Hanson, 2006 (general academic)	**0.28**
Tamim et al., 2011 (general academic)	**0.35**
Torgeson & Elbourne, 2002 (on spelling)	**0.37**
Torgeson & Zhu, ,2003(on reading)	**-0.05**
Torgeson & Zhu, 2003(on spelling)	**0.02**
Torgeson & Zhu, 2003 (on writing)	**0.89**
Waxman, Lin & Michko, 2003 (cognitive outcomes	**0.44**
Indicative effect size	**0.28**

For more information about the effect sizes in the Toolkit, click **here**.

Meta-analyses abstracts	
Study	***Abstract***
Bayraktar S. (2000)	This meta-analysis investigated how effective computer-assisted instruction (CAI) is on student achievement in secondary and college science education when compared to traditional instruction. An overall effect size of 0.273 was calculated from 42 studies yielding 108 effect sizes, suggesting that a typical student moved from the 50th percentile to the 62nd percentile in science when CAI was used. The results of the study also indicated that some study characteristics such as student-to-computer ratio, CAI mode, and duration of treatment were significantly related to the effectiveness of CAI. (Keywords: academic achievement, computer-assisted instruction, instructional effectiveness, meta-analysis, science education.)
Camnalbur & Erdogan (2010)	Studies focusing on the effectiveness of computer-assisted instruction have been growing recently in Turkey. In this research, quantitative studies comparing the effectiveness of computer- assisted instruction to traditional teaching method and conducted between 1998 and 2007 are studied by meta analysis. Seventy eight studies that have eligible data were combined with meta analytical methods by coding protocol from the 422 master's and doctoral degree and 124 articles. As a result for the study, the effect size of computer-assisted instruction method for academic achievement calculated 1.048. This is large scale according to Thalheimer and Cook, large and Cohen, Welkowitz and Ewen (2000). Recommendations were made based on the results of the study.

	Keywords
Cheung & Slavin (2011)	The present review examines research on the effects of technology use on reading achievement in K-12 classrooms. Unlike previous reviews, this review applies consistent inclusion standards to focus on studies that met high methodological standards. In addition, methodological and substantive features of the studies are investigated to examine the relationship between education technology and study features. A total of 85 qualified studies based on over 60,000 K-12 participants were included in the final analysis. Consistent with previous reviews of similar focus, the findings suggest that education technology generally produced a positive, though small, effect (ES=+0.16) in comparison to traditional methods. However, the effects may vary by education technology type. In particular, the types of supplementary computer-assisted instruction programs that have dominated the classroom use of education technology in the past few decades are not producing educationally meaningful effects in reading for K-12 students. In contrast, innovative technology applications and integrated literacy interventions with the support of extensive professional development showed somewhat promising evidence. However, too few randomized studies for these promising approaches are available at this point for firm conclusions.
Li & Ma (2010)	This study examines the impact of computer technology (CT) on mathematics education in K-12 classrooms through a systematic review of existing literature. A meta- analysis of 85 independent effect sizes extracted from 46 primary studies involving a total of 36,793 learners indicated statistically significant positive effects of CT on mathematics achievement. In addition, several characteristics of primary studies were identified as having effects. For example, CT showed advantage in promoting mathematics achievement of elementary over secondary school students. As well, CT showed larger effects on the mathematics achievement of special need students than that of general education students, the positive effect of CT was greater when combined with a constructivist approach to teaching than with a traditional approach to teaching, and studies that used non- standardized tests as measures of mathematics achievement reported larger effects of CT than studies that used standardized tests. The weighted least squares univariate and multiple regression analyses indicated that mathematics achievement could be accounted for by a few technology, implementation and learner characteristics in the studies. Keywords
Liao (2007)	A meta-analysis was performed to synthesize existing research comparing the effects of computer- assisted instruction (CAI) versus traditional instruction (TI) on students? achievement in Taiwan. Fifty- two studies were located from our sources, and their quantitative data was transformed into effect size (ES). The overall grand mean of the study-weighted ES for all 52 studies was 0.55. The results suggest that CAI is more effective than TI in Taiwan. In addition, two of the seventeen variables selected for this study (i.e., statistical power, and comparison group) had a statistically significant impact on the mean ES. The results from this study suggest that the effects of CAI in instruction are positive over TI. The results also shed light on the debate of learning from media between Clark and Kozma.
Pearson (2005)	This article reports the results of a meta-analysis of 20 research articles containing 89 effect sizes related to the use of digital tools and learning environments to enhance literacy acquisition. Results (weighted effect size of 0.489) demonstrate that technology can have a positive effect on reading comprehension, but little research has focused on the effect of technology on metacognitive, affective, and dispositional outcomes. We conclude that although there is reason to be optimistic about using technology in middle-school literacy programs, there is also reason to encourage the research community to redouble its emphasis on digital learning environments for students in this age range and to broaden the scope of the interventions and outcomes they study. Learning
Tamim et al.,	This research study employs a second-order meta-analysis procedure to sum-

(2011)	marize 40 years of research activity addressing the question, does computer technology use affect student achievement in formal face-to-face classrooms as compared to classrooms that do not use technology? A study-level meta- analytic validation was also conducted for purposes of comparison. An extensive literature search and a systematic review process resulted in the inclusion of 25 meta-analyses with minimal overlap in primary literature, encompassing 1,055 primary studies. The random effects mean effect size of 0.35 was significantly different from zero. The distribution was heteroge- neous under the fixed effects model. To validate the second-order meta- analysis, 574 individual independent effect sizes were extracted from 13 out of the 25 meta-analyses. The mean effect size was 0.33 under the random effects model, and the distribution was heterogeneous. Insights about the state of the field, implications for technology use, and prospects for future research are discussed.
Torgeson & Elbourne, (2002)	Recent Government policy in England and Wales on Information and Communica-tion Technology (ICT) in schools is heavily influenced by a series of non-randomised controlled studies. The evidence from these evaluations is equivocal with respect to the effect of ICT on literacy. In order to ascertain whether there is any effect of ICT on one small area of literacy, spelling, a systematic review of all randomised controlled trials (RCTs) was undertaken. Relevant electronic databases (including BEI, ERIC, Web of Science, PsycINFO, The Cochrane Library) were searched. Seven relevant RCTs were identified and included in the review. When six of the seven studies were pooled in a meta-analysis there was an effect, not statistically significant, in favour of computer interventions (Effect size =0.37, 95% confidence interval=70.02 to 0.77, p=0.06). Sensitivity and sub-group analyses of the results did not materially alter findings. This review suggests that the teaching of spelling by using computer software may be as effective as conventional teaching of spelling, although the possibility of computer-taught spelling being inferior or superior cannot be confidently excluded due to the relatively small sample sizes of the identified studies. Ideally, large pragmatic randomised controlled trials need to be undertaken.
Torgeson & Zhu (2003)	N/A
Waxman, Lin & Michko (2003)	To estimate the effects of teaching and learning with technology on students' cognitive, affective, and behavioral outcomes of learning, 282 effect sizes were calculated using statistical data from 42 studies that contained a combined sample of approximately 7,000 students. The mean of the study-weighted effect sizes averaging across all outcomes was .410 ($p < .001$), with a 95-percent confidence interval (CI) of .175 to .644. This result indicates that teaching and learning with technology has a small, positive, significant ($p < .001$) effect on student outcomes when compared to traditional instruction. The mean study-weighted effect size for the 29 studies containing cognitive outcomes was .448, and the mean study-weighted effect size for the 10 comparisons that focused on student affective outcomes was .464. On the other hand, the mean study-weighted effect size for the 3 studies that contained behavioral outcomes was -.091, indicating that technology had a small, negative effect on students' behavioral outcomes. The overall study-weighted effects were constant across the categories of study characteristics, quality of study indicators, technology characteristics, and instructional/teaching characteristics. Learning

Early years intervention

+5

Moderate impact for very high costs, based on extensive evidence.

Early years or early childhood interventions are approaches that aim to ensure that young children have educationally based pre-school or nursery experiences which prepare for school and academic success, usually through additional nursery or pre-school provision. Many of the researched programmes and approaches focus on disadvantaged children. Some also offer parental support. The research summarised here looks at general or multi-component programmes and approaches.

For more information about the impact of different aspects of early years provision please see the Early Years Toolkit.

How effective is it?

Overall, the evidence suggests that early years and pre-school intervention is beneficial. On average, early years interventions have an impact of five additional months' progress, and appear to be particularly beneficial for children from low income families.

Once early years provision is in place, efforts to improve the quality of provision, for example by training staff, appear to be more promising than simply increasing the quantity of provision by providing extra hours in the day, or by changing the physical environment of early years settings.

In most studies, the impact on attainment tends to wear off over time, though impact on attitudes to school tends to be more durable. There is no established amount of time over which the fade takes place; rather, there is a pattern of decline over time. Early years and pre-school interventions are therefore not sufficient to close the gap in attainment for disadvantaged children.

How secure is the evidence?

There are a number of systematic reviews and meta-analyses which have looked at the impact of early childhood intervention. Most of these are from the USA, however, where children tend to start school at a later age.

Evaluations of Sure Start in the UK do not show consistent positive effects and indicate that some caution is needed when generalising from exceptionally successful examples. However, overall the evidence supporting early childhood intervention is robust.

What are the costs?

Understandably the costs are high, as adult/child ratios in pre-school provision tend to be higher than in school classes and family interventions have similar high costs. The average cost per child of a Sure Start Local Programme was £1,300 in2009/10, so the estimates are in the region of £1,000-£2,000 per child. The average annual cost of sending a child over the age of two to a nursery is about £5800. Overall, the costs are estimated as very high.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. High quality provision is essential with well-qualified and well-trained staff.
2. High quality provision is likely to be characterised by the development of positive relationships between staff and children and by engagement of the children in activities which support pre-reading, the development of early number concepts and non-verbal reasoning..
3. Extended attendance (1 year +) and starting early (i.e. at 3 years old) is more likely to have an impact than shorter sessions starting later, which on average produce much lower gains.
4. Disadvantaged children benefit from good-quality programmes, especially where these include a mixture of children from different social backgrounds, and a strong educational component.

Technical Appendix: Early Years Interventions

Definition

Early years or early childhood interventions are approaches that aim to ensure that young children have educationally based pre-school or nursery experiences which prepare for school and academic success, usually through additional nursery or pre-school provision.

Search terms:

Early years, pre-school, nursery, pre-kindergarten, pre-k, foundation stage

Evidence rating

There are eleven meta-analyses included in the summary, with six conducted in last 10 years. The pooled effects from these syntheses range from 0.15 to 0.55 so provide a consistent estimate of effect (within 0.40 of a standard deviation), though more recent analyses of immediate impact have tended to be lower (0.15 to 0.32). Some variation is consistently explained by moderator analyses. The estimate of longterm benefit is consistent (0.53 and 0.55). A number of the meta-analyses include experimental and quasi-experimental studies which are not well controlled. Overall the evidence is rated as extensive.

Additional cost information

Information	Source
Cost of sending a child over the age of two to a nursery	Family and Childcare Trust, Childcare Costs Survey 2015 http://www.familyandchildcaretrust.org/sites/default/files/files/Childcare%20cost%20survey%202015%20Final.pdf
Cost of Sure Start Local Programme	Department for Education, National evaluation of Sure Start local programmes: An economic perspective http://www.ness.bbk.ac.uk/cost-effectiveness/documents/DFE-RR073_full.pdf

References

Full references
*Anderson, L.M., Shinn, C., Fullilove, M.T., Scrimshaw, S.C., Fielding, J.E., Normand, J., Carande-Kulis, V.G. (2003). The Effectiveness of Early Childhood Development Programs: A Systematic Review. *American Journal of Preventative Medicine* 24, 32-46. http://dx.doi.org/10.1016/S0749-3797(02)00655-4
Bakermans-Kranenburg, M.J., van IJzendoorn, M.H., Bradley, R.H. (2005). Those Who Have, Receive: The Matthew Effect in Early Childhood Intervention in the Home Environment. *Review of Educational Research*, 75. 1, 1-26. http://dx.doi.org/10.3102/00346543075001001
Barnett, W.S. (2011). Effectiveness of Early Educational Intervention *Science* 333, 975-978. http://dx.doi.org/10.1126/science.1204534
*Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2008). Meta-Analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, 112.3, 579–620. http://www.tcrecord.org/content.asp?contentid=15440
Campbell, F.A. & Ramey, C.T. (1994). Effects of Early Intervention on Intellectual and Academic Achievement: A Follow-up Study of Children from Low-Income Families. *Child Development* 65.2, 684-698. http://dx.doi.org/10.1111/j.1467-8624.1994.tb00777.x
*Chambers, B., Cheung, A., Slavin, R., Smith, D., & Laurenzano, M. (2010). Effective early childhood

education programmes: a best-evidence synthesis. Reading: CfBT Education Trust. http://cfbt.hs.llnwd.net/e1/~/media/cfbtcorporate/files/research/2010/r-early-childhood-programmes-synthesis-2010.pdf

*Gilliam, W. & Zigler, E.F. (2000). A critical meta-analysis of all evaluations of state-funded preschool from 1977 to 1998: implications for policy, service delivery and program evaluation. *Early Childhood Research Quarterly*, 15(4), 441-473. http://dx.doi.org/10.1016/S0885-2006(01)00073-4

*Gorey, K.M. (2001). Early childhood education: A meta-analytic affirmation of the short- and long-term benefits of educational opportunity. *School Psychology Quarterly*, 16(1), 9-30. http://dx.doi.org/10.1521/scpq.16.1.9.19163

*Karoly, L., Kilburn, R., Cannon, J.S. (2005). *Early Childhood Interventions: Proven Results, Future Promise*. Rand Corporation. http://www.rand.org/pubs/monographs/MG341.html

La Paro, K.M. & Pianta, R.C. (2000). Predicting Children's Competence in the Early School Years: A Meta-Analytic Review. *Review of Educational Research*, 70.4, 443-484. http://dx.doi.org/10.3102/00346543070004443

*Lewis, R.J. & Vosburgh, W.T. (1988). Effectiveness of Kindergarten Intervention Programs: A Meta-Analysis. *School Psychology International*, 9(4), 265-275. http://dx.doi.org/10.1177/0143034388094004

*Manning, M., Hommel, & Smith (2010). A meta-analysis of the effects of early years developmental prevention programs in at-risk populations on non-health outcomes in adolescence Childr*en and Youth Services Review* 32, 506–519. http://dx.doi.org/10.1016/j.childyouth.2009.11.003

Melhuish, E., Belsky, J., Leyland, A.H., Barnes, J (2008). /FR/027 London: University of London, Institute for the Study of Children, Families and Social Issues. http://www.ness.bbk.ac.uk/impact/documents/41.pdf

Melhuish, E. Belsky, J. MacPherson, K., Cullis, A. (2011). *The quality of group childcare settings used by 3-4 year old children in Sure Start Local Programme areas and the relationship with child outcomes* Research Report DFE-RR068 London Department for Education. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/182027/DFE-RR068.pdf

*Nelson, G., Westhues, A., & MacLeod, J. (2003). A meta-analysis of longitudinal research on preschool prevention programs for children. *Prevention & Treatment*, 6(1), 31a. http://dx.doi.org/10.1037/1522-3736.6.1.631a

*Shager, H. M., Schindler, H. S., Magnuson, K. A., Duncan, G. J., Yoshikawa, H., & Hart, C. M. (2013). Can research design explain variation in Head Start research results? A meta-analysis of cognitive and achievement outcomes. *Educational Evaluation and Policy Analysis*, 35.1, 76-95 http://dx.doi.org10.3102/0162373712462453

* Washington State Institute for Public Policy (2013) *Benefit-Cost Results State and district early childhood education programs* http://www.wsipp.wa.gov/BenefitCost?topicId=4

Summary of effects		
Meta-analysis	FSM effect size	Effect size
Anderson et al., 2003		0.35
Camilli et al. 2010		0.23
Chambers et al. 2010 (literacy) (mathematics)		0.15 0.17
Gilliam & Zigler, 2000		(NPE)
Gorey, 2001 (estimate on long term impact)		0.55
Karoly, Kilbirn & Cannon, 2005		0.28
La Paro & Pianta (2000)		0.51
Lewis & Vosburgh, 1988		0.41
Manning et al. 2010 (on adolescent education)		0.53
Nelson et al. 2003		0.52

Shager et al. 2013	0.27	0.27
Washington State Institute 2013	0.31	0.32
Indicative effect size (weighted mean)		**0.38**

For more information about the effect sizes in the Toolkit, click here.

Meta-analysis	Abstract
Anderson et al. 2003	Early childhood development is influenced by characteristics of the child, the family, and the broader social environment. Physical health, cognition, language, and social and emotional development underpin school readiness. Publicly funded, centre-based, comprehensive early childhood development programs are a community resource that promotes the well-being of young children. Programs such as Head Start are designed to close the gap in readiness to learn between poor children and their more economically advantaged peers. Systematic reviews of the scientific literature demonstrate effectiveness of these programs in preventing developmental delay, as assessed by reductions in retention in grade and placement in special education.
Camilli et al. 2008	**Background/Context:** There is much current interest in the impact of early childhood education programs on pre-schoolers and, in particular, on the magnitude of cognitive and affective gains. **Purpose/Objective/Research Question/Focus of Study:** Because this new segment of public education may require substantial resources, accurate descriptions are required of the potential benefits and costs of implementing specific preschool programs. To address this issue comprehensively, a meta-analysis was conducted for the purpose of synthesizing the outcomes of comparative studies in this area. **Population/Participants/Subjects:** A total of 123 comparative studies of early childhood interventions were analyzed. Each study provided a number of contrasts, where a contrast is defined as the comparison of an intervention group of children with an alternative intervention or no intervention group. **Intervention/Program/Practice:** A prevalent pedagogical approach in these studies was direct instruction, but inquiry-based pedagogical approaches also occurred in some interventions. No assumption was made that nominally similar interventions were equivalent. **Research Design:** The meta-analytic database included both quasi-experimental and randomized studies. A coding strategy was developed to record information for computing study effects, study design, sample characteristics, and program characteristics.
Chambers et al. 2010	This report systematically reviews research on the outcomes of programmes that teach young children in a group setting before they begin reception. Study inclusion criteria included the use of randomised or matched control groups, and study duration of at least 12 weeks. Studies included valid measures of language, literacy, phonological awareness, mathematical, and/or cognitive outcomes that were independent of the experimental treatments. A total of 40 studies, evaluating 29 different programmes met these criteria for outcomes assessed at the end of preschool and/or reception/ kindergarten. The review concludes that on academic outcomes at the end of preschool and/or reception, six early childhood programmes showed strong evidence of effectiveness and five had moderate evidence of effectiveness. Of the 29 programmes reviewed, eight are available for implementation in the UK. A few longitudinal studies have followed their subjects into secondary school, and even adulthood. These studies show

	that comprehensive programmes, from a cognitive developmental perspective rather than a solely academic focus, had better long-term effects on social adjustment outcomes such as reductions in delinquency, welfare dependency, and teenage pregnancy, and increases in educational and employment levels.
Gilliam & Zigler 2000	The number of state-funded preschool programs for low-income children has increased dramatically over the past few decades, and recent research has indicated that these programs vary considerably along a variety of dimensions. By 1998 only 13 of the current 33 state preschool programs (which serve children 3 to 5, provide some form of classroom-based educational service, and are primarily funded and administered at the state level) had completed a formal evaluation of the program's impact on child outcomes. This paper presents a critical meta-analytic review of these evaluations, providing measures of standardized effects for all significant impacts to facilitate comparisons across differing domains of outcome and evaluative methods. Although several methodological flaws in these studies are identified, the pattern of overall findings may offer modest support for positive impacts in improving children's developmental competence in a variety of domains, improving later school attendance and performance, and reducing subsequent grade retention. Significant impacts were mostly limited to kindergarten and first grade; however, some impacts were sustained several years beyond preschool. The results of these studies were similar to evaluations of other large-scale preschool programs for low-income children, such as Head Start. Modest outcome goals are warranted for preschool programs serving low-income children, for example, the promotion of school readiness. Suggestions are presented for improved preschool and early intervention program evaluation.
Gorey 2001	Some scholars who emphasize the heritability of intelligence have suggested that compensatory preschool programs, designed to ameliorate the plight of socioeconomically or otherwise environmentally impoverished children, are wasteful. They have hypothesized that cognitive abilities result primarily from genetic causes and that such environmental manipulations are ineffective. Alternatively, based on the theory that intelligence and related complex human behaviors are probably always determined by myriad complex interactions of genes and environments, the present meta-analytic study is based on the assumption that such behaviors can be both highly heritable and highly malleable. Integrating results across 35 preschool experiments and quasi-experiments, the primary findings were: (a) preschool effects on standardized measures of intelligence and academic achievement were statistically significant, positive and large; (b) cognitive effects of relatively intense educational interventions were significant and very large, even after 5 to 10 years, and 7 to 8 of every 10 preschool children did better than the average child in a control or comparison group; and (c) cumulative incidences of an array of personal and social problems were statistically significantly and substantially lower over a 10- to 25- year period for those who had attended preschool (e.g., school drop-out, welfare dependence, unemployment, poverty, criminal behavior). The need for a very large, well-controlled, national experiment to either confirm or refute these provocative, review-generated findings is discussed.
Karoly et al. 2005	PNC asked the RAND Corporation to prepare a thorough, objective review and synthesis of current research that addresses the potential for interventions of various forms in early childhood to improve outcomes for participating children and their families. In particular, we consider: • the potential consequences of not investing additional resources in the

	lives of children—particularly disadvantaged children— prior to school entry • the range of early intervention programs, focusing on those that have been rigorously evaluated • the demonstrated benefits of interventions with high-quality evaluations and the features associated with successful programs • the returns to society associated with investing early in the lives of disadvantaged children.
La Paro, K.M. & Pianta, R.C. 2000	School readiness screenings are prevalent throughout the United States. Although readiness encompasses a multitude of components, readiness assessments generally focus on measuring and predicting children's pre-academic skills and behaviors and are often the basis for placement and programming decisions. However, no quantitative estimates of effect sizes exist for the relations between preschool or kindergarten academic/cognitive and social~behavioral assessments and early school outcomes. This review presents the results of a meta-analysis of cross-time relations of academic/ cognitive and social~behavioral assessments from preschool to second grade. Results from 70 longitudinal studies that reported correlations between academic/cognitive and social/behavioral measures administered in preschool or kindergarten and similar measures administered in first and second grade were included in the analysis. Academic/cognitive assessments predicting similar outcomes showed moderate effect sizes across both time spans; effect sizes were small for social/behavioral predictors of early school social out- comes. Effect sizes varied considerably across individual studies and sam- ples. Findings are discussed in terms of assessment and conceptualization of school readiness, the role of school and classroom experiences in contribut- ing to individual differences in school outcomes, and the importance of a quantitative estimate of effect size for early education policy and practice.
Lewis & Vosburgh 1988	Psychologists and educators continue to design and implement kindergarten intervention programs unsubstantiated by previous research. The present study used meta-analysis procedures to examine the effects of kindergarten intervention programs on variables related to school success. The meta-analysis was performed on 444 effect sizes derived from 65 previous studies involving 3194 kindergarten children. The mean effect size of 0.434 indicated that test scores obtained by the treatment groups were raised from the 50th to the 67th percentile in relation to the control groups. Strong to moderate positive effects were demonstrated on all measured variables related to school success. As predicted the effect sizes from highly structured approaches (M= 0.517) were larger than those from less structured approaches (M= 0.298, t= 4.671, df=386, p< 0.001). In general there was no significant difference found between various levels of parental involvement (F= 0.244, df= 2.385, p> 0.05). However, when only the long-term effects were compared, a significant difference was found between the programs with active parental involvement (M= 0.521) and those without (M= 0.362, t= 2.067, df= 134, p<0.05). Strong effects were found on studies based on behavioral (M= 0.523) psycho-educational (M= 0.497) and stage referenced (M= 0.355) theories. The lack of research to support kindergarten programs based on maturational theories is discussed. The positive results of this meta-analysis should encourage program planners and policy makers to support the widespread implementation of structured early intervention and prevention programs at the kindergarten level.
Manning et al. 2010	We present the results of a meta-analytic review of early developmental prevention programs (children aged 0–5: structured preschool programs, center-based developmental day care, home visitation, family support services and parental education) delivered to at-risk populations on non-

	health outcomes during adolescence (educational success, cognitive development, social–emotional development, deviance, social participation, involvement in criminal justice, and family well-being). This review improves on previous meta-analyses because it includes a more comprehensive set of adolescent outcomes, it focuses on measures that are psychometrically valid, and it includes a more detailed analysis of program moderator effects. Seventeen studies, based on eleven interventions (all US-based) met the ten criteria for inclusion into the analysis. The mean effect size across all programs and outcomes was 0.313, equivalent to a 62% higher mean score for an intervention group than for a control group. The largest effect was for educational success during adolescence (effect size 0.53) followed by social deviance (0.48), social participation (0.37), cognitive development (0.34), involvement in criminal justice (0.24), family well-being (0.18), and social–emotional development (0.16). Programs that lasted longer than three years were associated with larger sample means than programs that were longer than one year but shorter than three years. More intense programs (those with more than 500 sessions per participant) also had larger means than less intense programs. There was a marginally significant trend for programs with a follow-through component into the early primary school years (e.g. preschool to Grade 3) to have more positive effects than programs without a follow-through. We conclude that the impact of well-conducted early development programs on quality of life in adolescence can be substantial for social policy purposes.
Nelson et al. 2003	The objectives of this research were to determine the effectiveness of preschool prevention programs for disadvantaged children and families in the short-term (preschool), medium-term (K-8), and the long-term (high school and beyond) and to identify factors that moderate program success. Meta-analysis was used to examine the effect sizes (d) of different outcome domains of 34 preschool prevention programs that had at least one follow-up assessment when the children were in school. While cognitive impacts resulting from these programs were greatest during the preschool period (d=.52), they were still evident during K-8 (d=.30). Social-emotional impacts on children were similar at K-8 (d=.27) and high school and beyond (d=.33), as were parent family wellness impacts at preschool (d=.33) and K-8 (d=.30). As predicted, cognitive impacts during the preschool time period were greatest for those programs that had a direct teaching component in preschool. Also as predicted, cognitive impacts during the K-8 time period were greatest for those programs that had a follow through educational component in elementary school. The longer the intervention for children, the greater were the impacts on preschool cognitive outcomes and child social-emotional outcomes at K-8; and the more intense the intervention for children, the greater were the impacts on preschool cognitive outcomes and parent-family outcomes at K-8. The largest impacts on preschool cognitive outcomes and child social-emotional and parent-family outcomes at K-8 were found for those programs that served predominantly African-American children. These results indicate that preschool prevention programs do have positive short-, medium-, and long-term impacts on several outcome domains. The findings were discussed in terms of contemporary trends in and future directions for policies and preschool prevention programs for children and families.
Shager et al. 2013	This meta-analysis explores the extent to which differences in research design explain the heterogeneity in program evaluation findings from Head Start impact studies. We predicted average effect sizes for cognitive and achievement outcomes as a function of the type and rigor of research

	design, quality and timing of dependent measure, activity level of control group, and attrition. Across 28 evaluations, the average program-level effect size was .27. About 41 percent of the variation in impacts across evaluations can be explained by the measures of research design features, including the extent to which the control group experienced other forms of early care or education, and 11 percent of the variation within programs can be explained by features of the outcomes.
Washington State Institute for Public Policy 2013	The WSIPP benefit-cost analysis examines the monetary value of programs or policies to determine whether the benefits from the program exceed its costs. Pre-kindergarten funded by states or school districts that is universal or targets low-income students.

Extending school time

Low impact for moderate cost, based on moderate evidence.

+2

Research has focused on three main approaches to extending school time:

1. Extending the length of the school year;
2. Extending the length of the school day;
3. Providing additional time for targeted groups of pupils either before or after school. There are examples of the school year being extended by up to five additional weeks or the school day being extended to 12 hours long. This summary focuses on extending core school time and the use of targeted before and after school programmes, particularly to support disadvantaged or low attaining pupils. Other approaches to increasing learning time are included in other sections of the Toolkit, such as Homework, Early Years Intervention and Summer Schools.

How effective is it?

Overall, the evidence indicates that, on average, pupils make two additional months' progress per year from extended school time or the targeted use of before and after school programmes. There is some evidence that disadvantaged pupils benefit disproportionately, making approximately two and a half months' additional progress. There are also often wider benefits for low-income students in terms of attendance at school, behaviour and relationships with peers.

After school programmes that support and encourage children academically while providing stimulating environments and activities are more likely to have an impact on attainment. To be successful, any increases in school time should be supported by both parents and staff, and extreme increases (e.g. more than nine hours of schooling per day) do not appear to be as effective. The research also indicates that attracting and retaining pupils in before and after school programmes is harder at secondary level than at primary level.

How secure is the evidence?

The evidence is moderately secure. Decisions to lengthen the school year or school day are often one component of wider approaches to school reform. This makes attributing any learning gains to additional time difficult. Gains are not consistent across studies, indicating that additional time may be used ineffectively. Discrete or targeted programmes are more likely to have been evaluated robustly, though even here there is substantial variation in impact.

Most of the evaluations of extended school time come from the USA, where enthusiasm for extended school time has outpaced the research base, indicating the need for more rigorous evaluations with outcome measures that demonstrate impact on learning. Evidence from the UK is relatively scarce.

What are the costs?

Overall, costs are estimated as moderate. The average cost of teaching a pupil is about £2,500 a year (£13 per day) in primary school and about £3,500 a year (£18 per day) in secondary. Extending the school year by two weeks would therefore require about £260 per pupil per year for primary schools and about £360 per pupil per year for secondary. Estimates suggest after school clubs cost, on average, £7 per session per pupil. A weekly session would therefore cost £273 per pupil over the course of a 39-week school year. The use of well-qualified and trained staff may increase these cost estimates.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Planning to get the most from the extra time is important. It should meet pupils' needs and build on their capabilities.
2. After school programmes with a clear structure, a strong link to the curriculum, and well-qualified and well-trained staff are more clearly linked to academic benefits.
3. After school programmes could give the opportunity to carry out some more intensive tuition (see entries for one to one or small group tuition).
4. Enrichment activities can have an impact on attainment, but the link is not well-established and the impact of different interventions can vary a great deal (see entries for sports or arts participation).
5. Have you explored how the quality of teaching and learning during school time can be improved? It might be cheaper and more efficient to attempt to use existing time more effectively before considering extending the school day.

Extending school time
Toolkit references

Full references
Baker, D. P., Fabrega, R., Galindo, C., & Mishook, J. (2004). Instructional time and national achievement: Cross-national evidence. *Prospects*, 34(3), 311-334. http://www.dx.doi.org/10.1007/s11125-004-5310-1
Cooper, H., Valentine, J. C., Charlton, K., & Melson, A. (2003). The effects of modified school calendars on student achievement and on school and community attitudes. Review of Educational Research, 73(1), 1-52. http://www.dx.doi.org/10.3102/00346543073001001
Crawford, (2011). Meta-Analysis of the Impact of After-School Programs on Students Reading and Mathematics Performance. Denton, Texas. UNT Digital Library.
Durlak J.A. & Weissberg, R.P. (2007). The Impact of After-School Programs that Promote Personal and Social Skills. Chicago: Collaborative for Academic, Social, and Emotional Learning (CASEL).
Durlak, J.A., Weissberg, R.P., Pachan, M. (2007). A Meta-Analysis of After-School Programs That Seek to Promote Personal and Social Skills in Children and Adolescents The Impact of After-School Programs that Promote Personal and Social Skills. American Journal of Community Psychology 45:294–309. Chicago: CASEL.
Fashola O.S (1998). Review of Extended Day and After-School Programs and their Effectiveness. CRESPAR Report No. 24. Center for Research on the Education of Students Placed at Risk.
Grossman, J.B., Price, M.L., Fellerath, V., Jucovy, L.Z., Kotloff, L.J., Raley, R. & Walker, K.E. (2002). Multiple Choices After School: Findings from the Extended-Service Schools Initiative. Philadelphia: Public/Private Ventures.
Kidron, Y., & Lindsay, J. (2014). The effects of increased learning time on student academic and nonacademic outcomes: Findings from a meta-analytic review. (ED-IES-12-C-0005:REL 2014–015). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center or Education Evaluation and Regional Assistance, Regional Educational Laboratory Appalachia. http://ies.ed.gov/ncee/edlabs
Lauer P.A., Akiba M., Wilkerson S.B., Apthorp H.S., Snow D., & Martin-Glenn M.L. (2006). Out-of-School Time Programs: A Meta-Analysis of Effects for at-risk Students. Review of Educational Research, 76, 275-313. http://www.dx.doi.org/10.3102/00346543076002275

Summary of effects		
Meta-analyses	***FSM effect size***	**Overall effect size**
Baker et.al. 2004 (international comparison)		0.12 (maths in the UK)
Cooper et al., 2003 (district level comparison)	0.24 0.19	0.06 (with comparison group) 0.11 (with well-matched controls
Crawford, 2011		0.40
Durlak & Weissberg, 2007		0.16
Fashola, 1998		NPE
Kidron & Lindsay, 2014	0.02 (literacy) 0.04 (maths)	-0.04 (literacy) 0.03 (maths)
Lauer, Akiba & Wilkerson, 2006		0.07
Scott-Little et al., 2002		NPE
Zief et al. 2006		0.08 (on GPA) 0.03 (on reading)
Weighted mean effect size		**0.11**

For more information about the effect sizes in the Toolkit, click here.

Extending school time
Toolkit references

Meta-analyses abstracts	
Study	***Abstract***
Cooper et.al. (2003).	This review synthesizes studies of the effects of modifying the academic calendar in Grades K-12 to do away with the long summer break while not increasing the length of the school year. The synthesis indicated that the quality of evidence on modified calendars is poor. Within this weak inferential frame-work, the average effect size for 39 school districts was quite small, d = .06, favoring modified calendars. Studies that used statistical or matching controls revealed an effect size of d = .11. Modified calendars were associated with higher achievement for economically disadvantaged students. Students, parents, and staffs who participated in modified calendar programs were positive about their experiences. Policymakers can improve acceptance of modified calendars by involving communities in the planning and by providing quality inter-session activities.
Baker et.al. (2004).	This article examines what we know about the influence of instructional time on achievement, particularly from the perspective of national implementation of schooling and national levels of achievement in mathematics. The report is in four sections. The first section provides a brief introduction to the idea of instructional time as a fundamental educational resource in the implementation of mass compulsory schooling. The second section reviews some past research exploring the relationship between instructional time and achievement. This section ends with a focus on specific research about how the economic development of a country can condition the relationship between instructional time and mathematics achievement across national school systems. The third section presents some original analyses of cross-national data to further illustrate these ideas. The final section provides some policy recommendations.
Crawford, 2011	The purpose of this study employing meta-analysis was to assess the impact that after-school programs have on reading and mathematics outcomes. The participants in the primary studies were students in Grades K through 8; years 200 through 2009. The study utilized the theory of change as its theoretical basis. This meta-analysis used the effect size as the standard measure. It began with an overall Cohen's d of .40 for the impact that after-school programs have on reading and mathematics outcomes, and then proceeded to analyse three moderator variables: subject, time periods, and grade level. The findings of the meta-analysis, both overall and sub analyses, show that the independent variable, after-school programs, has an impact on the dependent variable, reading and mathematics. The overall results indicated that after-school programs are educationally significant in the areas of reading and mathematics combined. As for the moderator variable, the results for the areas of (a) subject (reading and mathematics), (b) time period (2000-2002, 2003-2005 and 2006-2009), and (c) grade (middle, and middle plus elementary combined), all indicated educationally significant results. The notable exception was the grade moderator, elementary. This study provides more information for researchers, practitioners and policy makers upon which to make practical research based decisions about after-school programs for the purpose of determining the applicability of such in their educational setting.
Durlak & Weissberg 2007	A meta-analysis of after-school programs (ASPs) that seek to enhance the personal and social development of children and adolescents indicated that youth improved in three general areas: feelings and attitudes, indicators of behavioral adjustment, and school performance. More specifically, significant increases occurred in youths' self-perceptions and bonding to school, their positive social behaviors, and in their school grades and level of academic achievement. At the same time, significant reductions occurred in problem behaviors and drug use. Substantial differences emerged between programs that used evidence-based approaches for skill training and those that did not. The former programs consistently produced significant improvements among participants in all of the above outcome areas (mean effect sizes ranged from 0.24 to 0.35), whereas the latter programs did not produce significant results in any

	outcome category. Our findings have two important implications for future research, practice and policy. The first is that ASPs should contain components to foster the personal and social skills of youth, because participants can benefit in multiple ways if these components are offered. The second is that such components are effective only if they use evidence-based approaches. When it comes to enhancing personal and social skills, successful programs are SAFE sequenced, active, focused and explicit.
Fashola 1998	This report identifies and reviews thirty-four programs that have been used as after-school programs by schools and/or communities, including extended day programs and some supplemental school programs that have potential for after-school usage. Five categories of programs are reviewed: • language arts after-school programs, • study skills programs, • academic programs in other curriculum areas, • tutoring programs for reading, and • community-based programs. The review discusses these programs in terms of their evidence of effectiveness for improving student outcomes and their evidence of replicability in other locations. The report also summarizes correlational research studies that have examined the effects of after-school programs. Based on the program evaluations and the correlational research, the report presents a set of components of effective after-school programs and presents recommendations for implementing these components. The report concludes that stronger evaluations of these and other current after-school programs must be conducted, and other well-designed programs need to be developed and evaluated, in order to produce after-school programs that can be considered to be effective and replicable for increasing student achievement or other student outcomes.
Kidron & Lindsay 2014	Interest in increased learning time programs delivered beyond the regular school day has grown (Stonehill et al., 2011). These programs provide additional instruction in English language arts, math, and other subjects and are meant to enhance students' academic interests and success (Redd et al., 2012). The most common approaches include out-of-school programs (before- and after-school and weekend programs); summer school; schools with longer school days, weeks, or years; and year-round schools. Numerous evaluations have tested the effects of such programs on students' academic knowledge, study skills, social skills, and motivation to learn. This meta-analysis examined more than 7,000 studies, sorted them by scientific rigor, and identified 30 that used research designs capable of yielding strong evidence about the outcomes of increased learning time. In some cases the 30 studies found that increased learning time programs had a positive effect on student outcomes; in other cases the studies found no positive effect. This suggests that no single increased learning time program fits the needs of all students. The information in this report should help practitioners decide how best to select and implement an increased learning time approach. The programs were found, for example, to improve academic outcomes when instruction was led by certified teachers. Ten studies reported that literacy instruction was delivered by certified teachers and found a statistically significant positive effect on literacy achievement. Five studies reported that math instruction was conducted by certified teachers and found a statistically significant positive effect on math achievement. In both cases, however, the effects were small. Programs that used a traditional instruction style (with the teacher responsible for the progression of activities and students following directions to complete tasks) improved academic outcomes in literacy (nine studies) and math (four studies). The effects were small for both subjects. Programs that used an experiential learning instruction style (such as hands-on, inquiry-based instruction) improved student social-emotional skill development (for example, self-confidence and self-management; four studies). Again, the effects were small. The findings also show that increased learning time can benefit students at risk of academic failure. Increased learning time improved the literacy achievement of students performing below standards (three studies). Increased learning time also promoted the social-emotional skill development (for example,

	emotional well-being and externalizing behavior) of students with attention deficit/hyperactivity disorder (three studies). Programs that targeted specific student subgroups (such as struggling readers) and used explicit instruction to teach well specified skills tended to show a positive effect on student outcomes. Practitioners who wish to use increased learning time programs might there-fore set goals and design activities based on a deep understanding of student needs and interests. Because this study examined the data one category at a time, it does not provide information on potential interactions among implementation features, such as how the effectiveness of experiential learning, might vary with teacher–student ratio or the frequency and duration of classes. As the evidence base grows, studies like this one will be able to assess the effects of increased learning time using multiple factors at the same time.
Lauer et al. 2006	Schools and districts are adopting out-of-school-time (OST) programs such as after-school programs and summer schools to supplement the education of low-achieving students. However, research has painted a mixed picture of their effectiveness. To clarify OST impacts, this synthesis examined research on OST programs for assisting at-risk students in reading and/or mathematics. Researchers analyzed 35 OST studies that employed control or comparison groups and met other inclusion criteria. Meta-analyses indicated small but statistically significant positive effects of OST on both reading and mathematics student achievement and larger positive effect sizes for programs with specific characteristics such as tutoring in reading. Whether the OST program took place after school or during the summer did not make a difference in effectiveness.
Scott-Little, Hamann & Jurs 2002	Funding for after-school programs has increased dramatically, and there has been a corresponding increase in the need for sound evaluations to document the quality and impact of the programs. A comprehensive search for after-school evaluations was completed in order to conduct a meta-evaluation of evaluation methodologies used and to synthesize the findings of the evaluations. Results of the meta-evaluation indicate that the after-school evaluation reports located for the study demonstrated moderate compliance with The Program Evaluation Standards established by the Joint Committee on Standards for Educational Evaluation but limited use of research designs that support causal conclusions and insufficient information to allow for meta-analysis of program effects. However, some tentative conclusions can be reached about the effectiveness of after-school programs. Overall, it appears that after-school programs may have positive impacts on participants, but more rigorous research designs are necessary to provide data that clearly document program effects.
Zief, Lauver & Maynard 2006	Campbell Collaboration review: An extensive search of the literature uncovered only five studies that met the inclusion criteria for this review. A logic model for understanding the mechanisms for changing outcomes as a result of participation in after-school programs suggests that higher grades might occur after changes in students' behaviors and social and emotional outcomes. However, our analysis shows stronger effects for improved grades than for the behaviors that could impact grades, like improved school attendance and decreased television viewing.

Feedback

+8

High impact for very low cost, based on moderate evidence.

Feedback is information given to the learner and/or the teacher about the learner's performance relative to learning goals. It should aim to (and be capable of) producing improvement in students' learning. Feedback redirects or refocuses either the teacher's or the learner's actions to achieve a goal, by aligning effort and activity with an outcome. It can be about the learning activity itself, about the process of activity, about the student's management of their learning or self-regulation or (the least effective) about them as individuals. This feedback can be verbal, written, or can be given through tests or via digital technology. It can come from a teacher or someone taking a teaching role, or from peers.

How effective is it?

Feedback studies tend to show very high effects on learning. However, it also has a very high range of effects and some studies show that feedback can have negative effects and make things worse. It is therefore important to understand the potential benefits and the possible limitations of the approach. In general, research-based approaches that explicitly aim to provide feedback to learners, such as Bloom's 'mastery learning', also tend to have a positive impact. Feedback has effects on all types of learning across all age groups. Research in schools has focused particularly on English, mathematics and, to a lesser extent, science.

Research evidence about feedback was part of the rationale for Assessment for Learning (AfL). One evaluation of AfL indicated an impact of half of a GCSE grade per student per subject is achievable, which would be in line with the wider evidence about feedback.

Other studies reporting lower impact indicate that it is challenging to make feedback work in the classroom. This has also been demonstrated in a recent EEF pilot study where teachers tried to apply the evidence on feedback through an action research approach.

How secure is the evidence?

There are a substantial number of reviews and meta-analyses of the effects of feedback. Educational (rather than psychological or theoretical) studies tend to identify positive benefits where the aim is to improve learning outcomes in reading or mathematics or in recall of information. The most recent meta-analysis of studies focusing on formative assessment in schools indicates the gains are more modest, suggesting an improvement of about three months' additional progress is achievable in schools or nearer four months' when the approach is supported with professional development.

What are the costs?

The costs of providing more effective feedback are not high. However it is likely to require sustained professional development to improve practice, and this includes active inquiry and evaluation. Overall, costs are estimated as under £80 per pupil and very low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Providing effective feedback is challenging. Research suggests that it should be specific, accurate and clear (e.g. “It was good because you...” rather than just “correct”); compare what a learner is doing right now with what they have done wrong before (e.g. “I can see you were focused on improving X as it is much better than last time’s Y…”); encourage and support further effort and be given sparingly so that it is meaningful; provide specific guidance on how to improve and not just tell students when they are wrong; and be supported with effective professional development for teachers.
2. Wider research suggests the feedback should be about complex or challenging tasks or goals as this is likely to emphasise the importance of effort and perseverance as well as be more valued by the pupils. Feedback can come from other peers as well as adults (see Peer tutoring).
3. Have you considered the challenge of implementing feedback effectively and consistently?
4. What professional development requirements are likely to be necessary for success?

Technical Appendix: Feedback

Definition

Feedback is information given to the learner and/or teacher about the learner's performance relative to learning goals or outcomes. It should aim to (and be capable of) producing improvement in students' learning. Feedback redirects or refocuses either the teacher's or the learner's actions to achieve a goal, by aligning effort and activity with an outcome. It can be about the output of the activity, the process of the activity, the student's management of their learning or self-regulation, or them as individuals. This feedback can be verbal or written, or can be given through tests or via digital technology. It can come from a teacher or someone taking a teaching role, or from peers.

Search terms: feedback; formative evaluation; assessment for learning; feedback interventions. corrective feedback.

Evidence rating

There are seven meta-analyses of feedback and feedback interventions which have consistently found high average effects of feedback on learning and academic performance. Only two of these have been conducted in the last 10 years. Many of the studies included are small scale studies from psychology which demonstrate theoretical principles, but which may be difficult to generalise to educational practice. Larger scale educational studies tend to have lower effects. The meta-analyses include a very wide range of effects. Overall the evidence is rated as moderate.

Additional cost information

The costs of providing more effective feedback are not high. However it is likely to require sustained professional development to improve practice, and this includes active inquiry and evaluation. Overall, costs are estimated as under £80 per pupil and very low.

References

Full references
*Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213-238. http://dx.doi.org/10.3102/00346543061002213
Bennett, R.E. (2011). Formative assessment: a critical review, *Assessment in Education: Principles, Policy & Practice*, 18: 1, 5-25. http://dx.doi.org/10.1177/001440298605300301
Black, P. & Wiliam, D. (2005). Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices. *Curriculum Journal*, 16, 249-261. http://dx.doi.org/10.1080/09585170500136218
Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability* 21 1, pp 5-31. http://dx.doi.org/10.1007/s11092-008-9068-5
Black P. & Wiliam, D. (1998). Assessment and classroom learning, *Assessment in Education*, 5, pp. 7-73. http://dx.doi.org/10.1080/0969595980050102
Bloom, B.S., Hastings, J.T. & Madaus, G.F. (eds.) (1971). Handbook on the Formative and Summative Evaluation of Student Learning New York: McGraw-Hill.
*Fuchs, L.S. & Fuchs, D. (1986). Effects of systematic formative evaluation A meta-analysis. Exceptional Children, 53.3 pp 199-208. http://dx.doi.org/10.1177/001440298605300301

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative Assessment and Writing. *The Elementary School Journal*, 115(4), 523-547. http://dx.doi.org/10.1086/681947
Hattie, J. and Timperley, H. (2007). The Power of Feedback. *Review of Educational Research* 77.1 pp 81–112. http://dx.doi.org/10.3102/003465430298487
*Kingston, N. & Nash, B. (2011). Formative Assessment: A Meta-Analysis and Call for Research. *Educational Measurement: Issues and Practice*. 30.4 pp 28-37. http://dx.doi.org/10.1111/j.1745-3992.2011.00220.x
* Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254. http://dx.doi.org/10.1037/0033-2909.119.2.254
Kulik, C. Kulik, J. & Bangert-Drowns, R. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, 60.2 pp 265-306. http://dx.doi.org/10.3102/00346543060002265
*Lysakowski, R.S., & Walberg, H.J. (1982). Instructional Effects of Cues, Participation, and Corrective Feedback: A Quantitative Synthesis. *American Educational Research Journal*, 19(4), 559-578. http://dx.doi.org/10.3102/00028312019004559
Smith, E. & Gorard, S. (2005) They don't give us our marks': the role of formative feedback in student progress. *Assessment in Education* 12. 1, pp. 21–38. http://dx.doi.org/10.1080/0969594042000333896
*Tenenbaum, G., & Goldring, E. (1989). A Meta-Analysis of the Effect of Enhanced Instruction: Cues, Participation, Reinforcement and Feedback, and Correctives on Motor Skill Learning. *Journal of Research and Development in Education,* 22(3), 53-64.

* Studies marked with an asterisk are included in the summary of effects

Summary of effects		
Study	***FSM effect size***	***Overall effect size***
Bangert-Drowns et al., 1991		0.26
Fuchs & Fuchs, 1986		0.72
Graham et al. 2015 (writing)		0.61
Kingston & Nash, 2011 (AfL)		0.20
Kluger & DeNisi, 1996		0.41
Lysakowski & Walberg, 1982		0.97
Tenenbaum & Goldring, 1989		0.72
Weighted mean effect size		**0.63**

For more information about the effect sizes in the Toolkit, click here.

Technical Appendix: Feedback

Education Endowment Foundation

Meta-analyses abstracts	
Study	***Abstract***
Bangert-Drowns et al. 1991	Feedback is an essential construct for many theories of learning and instruction and an understanding of the conditions for effective feedback should facilitate both theoretical development and instructional practice. In an early review of feedback effects in written instruction Kulhavy (1977) proposed that feedback's chief instructional significance is to correct errors. This error-correcting action was thought to be a function of presentation timing, response certainty and whether students could merely copy answers from feedback without having to generate their own. The present meta-analysis reviewed 58 effect sizes from 40 reports. Feedback effects were found to vary with for control for pre-search availability, type of feedback, use of pre-tests and type of instruction and could be quite large under optimal conditions. Mediated intentional feedback for retrieval and application of specific knowledge appears to stimulate the correction of erroneous responses in situations where its mindful (Solomon & Globerson, 1987) reception is encouraged.
Fuchs & Fuchs, 1986	While the aptitude treatment interaction (ATI) approach to educational measurement emphasizes establishing salient learner characteristics, systematic formative evaluation provides ongoing evaluation for instructional program modification. Systematic formative evaluation appears more tenable than ATI for developing individualized instructional programs. This meta-analysis investigates the effects of educational programs on student achievement. Twenty-one controlled studies generated 95 relevant effect sizes, with an average effect size of .72. The magnitude of effect size was associated with publication type, data evaluation methods, and use of behaviour modification. Findings indicate that unlike reported ATI approaches to individualization, systematic formative evaluation procedures reliably increase academic achievement. This suggests that, given an adequate measurement methodology, practitioners can inductively formulate successful individualized educational programs.
Graham et al. 2015	To determine whether formative writing assessments that are directly tied to everyday classroom teaching and learning enhance students' writing performance, we conducted a meta-analysis of true and quasi-experiments conducted with students in grades 1 to 8. We found that feedback to students about writing from adults, peers, self, and computers statistically enhanced writing quality, yielding average weighted effect sizes of 0.87, 0.58, 0.62, and 0.38, respectively. We did not find, however, that teachers' monitoring of students' writing progress or implementation of the 6 _ 1 Trait Writing model meaningfully enhanced students' writing. The findings from this meta-analysis provide support for the use of formative writing assessments that provide feedback directly to students as part of everyday teaching and learning. We argue that such assessments should be used more frequently by teachers, and that they should play a stronger role in the Next-Generation Assessment Systems being developed by Smarter Balanced and PARCC.
Kingston & Nash 2011	An effect size of about .70 (or .40–.70) is often claimed for the efficacy of formative assessment, but is not supported by the existing research base. More than 300 studies that appeared to address the efficacy of formative assessment in grades K-12 were reviewed. Many of the studies had severely flawed research designs yielding un-interpretable results. Only 13 of the studies provided sufficient information to calculate relevant effect sizes. A total of 42 independent effect sizes were available. The median observed effect size was .25. Using a random effects model, a weighted mean effect size of .20 was calculated. Moderator analyses suggested that formative assessment might be more effective in English language

	arts (ELA) than in mathematics or science, with estimated effect sizes of .32, .17, and .09, respectively. Two types of implementation of formative assessment, one based on professional development and the other on the use of computer-based formative systems, appeared to be more effective than other approaches, yielding mean effect size of .30 and .28, respectively. Given the wide use and potential efficacy of good formative assessment practices, the paucity of the current research base is problematic. A call for more high-quality studies is issued.
Kluger & De Nisi 1996	Since the beginning of the century, feedback interventions (FIs) produced negative—but largely ignored—effects on performance. A meta-analysis (607 effect sizes; 23,663 observations) suggests that FIs improved performance on average (d = .41) but that over '/3 of the FIs decreased performance. This finding cannot be explained by sampling error, feedback sign, or existing theories. The authors proposed a preliminary FI theory (FIT) and tested it with moderator analyses. The central assumption of FIT is that FIs change the locus of attention among 3 general and hierarchically organized levels of control: task learning, task motivation, and meta-tasks (including self-related) processes. The results suggest that FI effectiveness decreases as attention moves up the hierarchy closer to the self and away from the task. These findings are further moderated by task characteristics that are still poorly understood.
Lysakowski & Walberg 1982	To estimate the instructional effects of cues, participation, and corrective feedback on learning 94 effect sizes were calculated from statistical data in 54 studies containing a combined sample of 14,689 students in approximately 700 classes. The mean of the study-weighted effect size is .97, which suggest average percentiles on learning outcomes of 83 and 50 respectively, for experimental and control groups. The strong effects appeared constant from elementary level through college, and across socioeconomic levels, races, private and public schools, and community types. In addition the effects were not significantly different across the categories of methodological rigor such as experiments and quasi-experiments
Tenenbaum & Goldring 1982	Estimated the effect of enhanced instruction on motor skill acquisition in a meta-analysis of 15 studies that used 4–5 yr old children and 4th–21th graders in Israel. Ss exposed to enhanced instruction gained more qualified motor skills than over 75% of the Ss exposed to regular instruction in a variety of motor skills. Enhanced instruction used cues and explanations by the instructor to clarify the motor skill, encouraged Ss to actively participate in the task over 70% of the time, reinforced Ss' responses, and supplied ongoing feedback and correctives to ensure motor skill acquisition.

Homework (Primary)

Low impact for very low or no cost, based on limited evidence.

+2

Homework refers to tasks given to pupils by their teachers to be completed outside of usual lessons. Common homework activities in primary schools tend to be reading or practising spelling and number facts, but may also include more extended activities to develop inquiry skills or more directed and focused work such as revision for tests.

How effective is it?

It is certainly the case that schools whose pupils do homework tend to be more successful. However it is less clear that the homework is the reason why they are successful. A number of reviews and Meta-analyses have explored this issue. There is stronger evidence that it is helpful at secondary level [see Homework (secondary)], but there is much less evidence of benefit at primary level.

There is some evidence that when homework is used as a short and focused intervention it can be effective in improving students' attainment, but this is limited for primary age pupils. Overall the general benefits are likely to be modest if homework is more routinely set.

The quality of the task set appears to be more important than the quantity of work required from the pupil.

How secure is the evidence?

Homework has been extensively researched. There is a relatively consistent picture that pupils in schools which give more homework perform better, although for primary age pupils the difference is small. However, there are only a small number of studies which have investigated whether this relationship is due to the homework itself, rather than other school factors. These studies compare classes where homework is introduced to similar classes where homework is not given. They tend to show that homework can be beneficial, but this finding is less secure than the first, because of the smaller number of studies and the quality of the evidence.

What are the costs?

There are few costs associated with homework, though there are implications for staff time for preparation and marking. With younger children there may be additional resources required (such as reading books or games for children to take home). Overall costs are estimated as very low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Overall, homework in primary schools does not appear to lead to large increases in learning.
2. Effective homework is associated with greater parental involvement and support. How will you design homework to encourage parental engagement?
3. The broader evidence base suggests that short focused tasks or activities which relate directly to what is being taught, and which are built upon in school, are likely to be more effective than regular daily homework.
4. Have you made the purpose of homework clear to children?

Technical Appendix: Homework (primary)

Definition

Homework refers to tasks given to pupils by their teachers to be completed outside of usual lessons. Common homework activities in primary schools tend to be reading or practising spellings and number facts, but may also include extended activities to develop inquiry skills or more directed and focused work such as revision for tests.

Search terms: homework, homework clubs, home assignment, home reading, flipped learning.

Evidence rating

Homework has been extensively researched. However, studies have mainly looked at the correlation between homework and how well schools or pupils perform, especially for younger children. There is a relatively consistent finding that there is a positive association but that this is very small for primary age pupils. There are only a small number of studies which have investigated what happens when homework is introduced and comparison is made with classes where homework is not given, and there is very little evidence of this kind specifically for primary age pupils. The studies there are tend to show that homework can be beneficial, though the evidence is not secure. The variation in what is assigned as 'homework' and how this relates to what happens in school means the variation in reported impact between different studies is always likely to be large. There are two meta-analyses, one published in the last ten years, and one recent systematic review. There are no meta-analyses which specifically focus on homework for primary age pupils. Overall the evidence is rated as limited.

Additional cost information

There are few costs associated with homework, though there are implications for staff time for preparation and marking. With younger children there may be additional resources required (such as reading books or games for children to take home). Overall costs are estimated as very low.

References

Full references
Canadian Council on Learning (2009) A systematic review of literature examining the impact of homework on academic achievement Toronto: Canadian Council on Learning Learning http://edu.au.dk/fileadmin/edu/Udgivelser/SystematicReview_HomeworkApril27-2009.pdf
*Cooper, H., Robinson, J.C., Patall, E.A. (2006). Does Homework Improve Academic Achievement? A Synthesis of Research, *1987-2003. Review of Educational Research*, 76. 1 pp. 1-62. http://dx.doi.org/10.3102/00346543076001001
Dettmers, S., Trautwein, U., & Ludtke, O. (2009). The relationship between homework time and achievement is not universal: evidence from multilevel analyses in 40 countries. *School Effectiveness and School Improvement,* 20(4), 375-405. http://dx.doi.org/10.1080/09243450902904601
*Farrow, S., Tymms, P., & Henderson, B. (1999). Homeworkand attainment in primary schools, *British Educational Research Journal,* 25(3), 323-341. http://dx.doi.org/10.1080/0141192990250304
Gustafsson, J. (2013) Causal inference in educational effectiveness research: a comparison of three methods to investigate effects of homework on student achievement , School Effectiveness and School Improvement, 24:3, 275-295. http://dx.doi.org/10.1080/09243453.2013.806334
*Paschal, R.A., Weinstein, T. & Walberg, H.J. (1984). The effects of homework on learning: A quantitative synthesis. The Journal of Educational Research, 78:2, 97-104. Rønning, M. (2011). Who benefits from homework assignments? *Economics of Education Review*, 30, 55-64.

http://dx.doi.org/10.1080/00220671.1984.10885581
Trautwein, U. (2007). The homework-achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. *Learning and Instruction*, 17, 372–388. http://dx.doi.org/10.1016/j.learninstruc.2007.02.009
Trautwein, U., Schnyder, I., Niggli, A., Neumann, M., & Lüdtke, O. (2009). Chameleon effects in homework research: The homework–achievement. *Contemporary Educational Psychology*, 34, 77-88. http://dx.doi.org/10.1016/j.cedpsych.2008.09.001

* Studies marked with an asterisk are included in the summary of effects.

Summary of effects		
Study	***FSM effect size***	***Overall effect size***
Cooper et al., 2006 (correlation for elementary)		0.10
Paschal et al., 1984 (all ages)	0.15	0.36
Farrow et al. 1999 (correlation)		0.00
Median effect size		**0.10**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Cooper et al. 2006	In this article, research conducted in the US since 1987 on the effects of homework is summarized. Studies are grouped into four research designs. The authors found that all studies, regardless of type, had design flaws. However, both within and across design types, there was generally consistent evidence for a positive influence of homework on achievement. Studies that reported simple homework-achievement correlations revealed evidence that a stronger correlation existed a) in Grades 7-12 than in K-6 and b) when students rather than parents reported time on homework. No strong evidence was found for an association between the homework-achievement link and the outcome measure (grades as opposed to standardized tests) or the subject matter (reading as opposed to math). On the basis of these results and others, the authors suggest future research.
Paschal et al. 1984	This paper synthesizes empirical studies of homework and of various homework strategies on the academic achievement and attitude of elementary and secondary students. A computer search yielded 15 published and un published studies with explicit statistical results. Fifty-four characteristics of treatments, contexts, conditions, validity, and outcomes were coded for each study. About 85% of the effect sizes favored the homework groups. The mean effect size is .36 (probability less than .0001). Homework that was graded or contained teachers' comments produced stronger effects (.80).
Farrow et al. 1999	An analysis of data relating to homework in the final year of primary school is reported in the core areas of mathematics, English and science. Information was available on achievement levels, attitudes, developed ability, cultural capital and sex. Widespread variation in reported homework frequency emerged in the core subjects in primary schools, as did significant differences between girls and boys for mathematics and reading. The findings indicated that the highest test scores were achieved by those pupils who reported doing homework 'once a month' in each of the core subjects. Homework reported more frequently than 'once a month' was generally associated with lower attainment. Multilevel models that controlled for important variables did not lend support to the 'more is better' view of homework. Our contention is that the assumptions about the value of homework (largely derived from

	secondary school practice and experience) should not be automatically 'grafted on' to primary practice. More serious consideration should be given to the nature and frequency of homework setting in primary schools.

Homework (Secondary)

Moderate impact for very low or no cost, based on moderate evidence.

+5

Homework refers to tasks given to pupils by their teachers to be completed outside of usual lessons. Common homework activities may be reading or preparing for work to be done in class, or practising and completing tasks or activities already taught or started in lessons, but it may include more extended activities to develop inquiry skills or more directed and focused work such as revision for exams.

How effective is it?

On average, the impact of homework on learning is consistently positive (leading to on average five months' additional progress). However, beneath this average there is a wide variation in potential impact, suggesting that how homework is set is likely to be very important.

There is some evidence that homework is most effective when used as a short and focused intervention (e.g. in the form of a project or specific target connected with a particular element of learning) with some exceptional studies showing up to eight months' positive impact on attainment. Benefits are likely to be more modest, up to two to three months' progress on average, if homework is more routinely set (e.g. learning vocabulary or completing problem sheets in mathematics every day).

Evidence also suggests that how homework relates to learning during normal school time is important. In the most effective examples homework was an integral part of learning, rather than an add-on. To maximise impact, it is also appears to be important that students are provided with high quality feedback on their work (see Feedback).

Studies imply that there is an optimum amount of homework of between 1 and 2 hours per school day (slightly longer for older pupils), with effects diminishing as the time that students spend on homework increases.

How secure is the evidence?

Homework has been extensively studied. However, studies have mainly looked at the correlation between homework and how well schools perform. It is certainly the case that schools whose pupils do homework tend to perform well, but it is less clear that the homework is the reason why they are successful.

There are a smaller number of studies which have investigated what happens when homework is introduced and compared with classes where homework is not given. These studies tend to show that homework is beneficial, though the evidence is less secure.

What are the costs?

There are few costs associated with homework, though there are implications for staff time for preparation and marking. With younger children there may be additional resources required (such as reading books or games for children to take home). Overall costs are estimated as very low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Planned and focused activities are more beneficial than homework which is more regular but routine or not linked with what is being learned in class.
2. It should not be used as a punishment or penalty for poor performance.
3. A variety of tasks with different levels of challenge is likely to be beneficial.
4. The quality of homework is more important than the quantity. Pupils should receive specific and timely feedback on homework.
5. Have you made the purpose of homework clear to children (e.g. to increase a specific area of knowledge, or fluency in a particular area)?

Technical Appendix: Homework (secondary)

Definition

Homework refers to tasks given to pupils by their teachers to be completed outside of usual lessons. Common homework activities in secondary schools are completing tasks assigned in lessons, such as practicing further questions or problems in mathematics, or preparing for tasks in future lessons. It may also include routine course work or revising for tests and examinations. In some models of 'flipped learning', pupils prepare at home for classroom discussion and application tasks. It also includes activities such as 'homework clubs' where pupils are given the opportunity to complete their assigned tasks in school, usually at the end of the school day.

Search terms: homework, homework clubs, home assignment, home reading, flipped learning.

Evidence rating

Homework has been extensively studied and is a controversial topic. Studies have mainly looked at the correlation between homework and how well schools or pupils perform. There is a relatively consistent finding that there is a positive association but that this reduces when ability and home background are taken into account. There are a smaller number of studies which have investigated what happens when homework is introduced and comparison is made with classes where homework is not given, where the typical gain is as much as five or six months additional progress for secondary pupils. There are two meta-analyses, one published in the last ten years, and one recent systematic review. The variation in what is assigned as 'homework' and how this relates to what happens in school means the variation in reported impact between different studies is always likely to be large. Overall the evidence is rated as limited.

Additional cost information

There are few costs associated with homework, though there are implications for staff time for preparation and marking. Some tasks may require additional resources (such as books or materials for pupils to use at home). Overall costs are estimated as very low.

References

Full references
Canadian Council on Learning (2009) *A systematic review of literature examining the impact of homework on academic achievement* Toronto: Canadian Council on Learning Learning http://edu.au.dk/fileadmin/edu/Udgivelser/SystematicReview_HomeworkApril27-2009.pdf
*Cooper, H., Robinson, J.C., Patall, E.A. (2006). Does Homework Improve Academic Achievement? A Synthesis of Research, *1987-2003. Review of Educational Research*, 76. 1 pp. 1-62. http://dx.doi.org/10.3102/00346543076001001
Dettmers, S., Trautwein, U., & Ludtke, O. (2009). The relationship between homework time and achievement is not universal: evidence from multilevel analyses in 40 countries. *School Effectiveness and School Improvement,* 20(4), 375-405. http://dx.doi.org/10.1080/09243450902904601
Eren, O., & Henderson, D. J. (2011). Are we wasting our children's time by giving them more homework?. *Economics of Education Review*, 30(5), 950-961. http://dx.doi.org/10.1016/j.econedurev.2011.03.011
Farrow, S., Tymms, P., & Henderson, B. (1999). Homeworkand attainment in primary schools, *British Educational Research Journal,* 25(3), 323-341. http://dx.doi.org/10.1080/0141192990250304

Technical Appendix: Homework (secondary)

Gustafsson, J. (2013) Causal inference in educational effectiveness research: a comparison of three methods to investigate effects of homework on student achievement , School Effectiveness and School Improvement, 24:3, 275-295. http://dx.doi.org/10.1080/09243453.2013.806334
*Paschal, R.A., Weinstein, T. & Walberg, H.J. (1984). The effects of homework on learning: A quantitative synthesis. The Journal of Educational Research, 78:2, 97-104. Rønning, M. (2011). Who benefits from homework assignments? *Economics of Education Review*, 30, 55-64. http://dx.doi.org/10.1080/00220671.1984.10885581
Trautwein, U. (2007). The homework-achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. *Learning and Instruction*, 17, 372–388. http://dx.doi.org/10.1016/j.learninstruc.2007.02.009
Trautwein, U., Schnyder, I., Niggli, A., Neumann, M., & Lüdtke, O. (2009). Chameleon effects in homework research: The homework–achievement. *Contemporary Educational Psychology*, 34, 77-88. http://dx.doi.org/10.1016/j.cedpsych.2008.09.001
Tymms, P. B. and C. T. Fitz-Gibbon (1992). The relationship of homework to A-level results. *Educational Research*, 34(1): 3-19. http://dx.doi.org/10.1080/0013188920340l0

* Studies marked with an asterisk are included in the summary of effects

Summary of effects		
Study	***FSM effect size***	***Overall effect size***
Cooper et al., 2006 (secondary homework/ no homework/ RCT)		0.54
Cooper et al., 2006 (secondary correlational)		0.49
Paschal et al., 1984 (all age groups)	0.15	0.36
Weighted mean effect size		**0.44**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Cooper et al. 2006	In this article, research conducted in the US since 1987 on the effects of homework is summarized. Studies are grouped into four research designs. The authors found that all studies, regardless of type, had design flaws. However, both within and across design types, there was generally consistent evidence for a positive influence of homework on achievement. Studies that reported simple homework-achievement correlations revealed evidence that a stronger correlation existed a) in Grades 7-12 than in K-6 and b) when students rather than parents reported time on homework. No strong evidence was found for an association between the homework-achievement link and the outcome measure (grades as opposed to standardized tests) or the subject matter (reading as opposed to math). On the basis of these results and others, the authors suggest future research.
Paschal et al. 1984	This paper synthesizes empirical studies of homework and of various homework strategies on the academic achievement and attitude of elementary and secondary students. A computer search yielded 15 published and un published studies with explicit statistical results. Fifty-four characteristics of treatments, contexts, conditions, validity, and outcomes were coded for each study. About 85% of the effect sizes favored the homework groups. The mean effect size is .36 (probability less than .0001). Homework that was graded or contained teachers' comments produced stronger effects (.80).

Individualised instruction

+2

Low impact for very low cost, based on moderate evidence.

Individualised instruction provides different tasks for each learner and provides support at the individual level. It is based on the idea that all learners are different and therefore have different needs, so an individualised or personally tailored approach to instruction ought to be more effective, particularly in terms of the tasks and activities that pupils undertake and the pace at which they make progress through the curriculum. Examples of individualised education have been tried over the years in education, particularly in areas like mathematics where pupils can have individual sets of activities which they complete, often largely independently.

How effective is it?

Individualising instruction does not tend to be particularly beneficial for learners. One possible explanation for this is that the role of the teacher becomes too managerial in terms of organising and monitoring learning tasks and activities, without leaving time for interacting with learners or providing formative feedback to refocus effort. The average impact on learning tends overall to be low, and is even negative in some studies, appearing to delay progress by one or two months.

How secure is the evidence?

There have been a number of meta-analyses which have found broadly similar effects, and support the conclusion that individualising learning for whole classes is not beneficial for pupils' learning.

This finding is also supported by research from other connected fields, such as computer based learning, and Bloom's 'mastery learning', where students have instructions broken down into steps, receive feedback on their learning, and only move on when they have 'mastered' a particular step. In both fields, small group approaches appear to be more effective than individualised approaches.

The evidence is mostly drawn from secondary school studies and predominantly in mathematics, though there is also evidence from other curriculum subjects such as science, history and geography.

What are the costs?

The costs of implementing individualised learning are usually low, unless the approach uses technology (such as tutoring programmes or integrated learning systems). Estimated outlay for increased resourcing per pupils is £150 per year. Overall costs are therefore estimated as very low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Overall the evidence does not support approaches which individualise instruction at class level.
2. It is hard to identify exactly why individualised instruction is not more effective. It may be that in a classroom setting, learners receive less direct teaching, get less feedback or move at a slower pace when they manage their own learning progress with support (see Meta-cognition and self-regulation).
3. Individualised instruction runs the risk of the teacher managing diverse activities and learners, without sufficient time to work directly with each individual.
4. Have you considered small group or one to one settings as a more viable strategy?
5. Approaches to individualise learning activities supported by technology may provide learners with effective practice, however it is still important to ensure that learners receive direct instruction from a teacher when learning new content, or when they are not making progress.

Individualised instruction
Toolkit references

Full references
Aiello, N.C. & Wolfle, L.M. (1980). A Meta-Analysis of Individualized Instruction in Science Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 7-11, 1980). ERIC ED190404.
Bangert, R.L., Kulik, J.A., Kulik, C.C. (1983). Individualized Systems of Instruction in Secondary Schools. Review of Educational Research, 53.2. pp. 143-158.
Horak, V.M. (1981). A Meta-Analysis of Research Findings on Individualized Instruction in Mathematics. Journal of Educational Research, 74:4 p 249.
Lou, Y., Abrami, P.C., & d'Apollonia, S. (2001). Small Group and Individual Learning with Technology: A Meta-Analysis. Review of Educational Research 71(3), 449-521. doi: 10.3102/00346543071003449.
Slavin R.E. & Karweit, N.L. (1985). Effects of Whole Class, Ability Grouped, and Individualized Instruction on Mathematics American Educational Research Journal 22.3 pp. 351-367.
Willett, J.B., Yamashita, J.J. & R.D. Anderson (1983). "A Meta-Analysis of Instructional Systems Applied in Science Teaching." Journal of Research in Science Teaching 20(5):405-17.
Yeh, S.(2010). 'Understanding and addressing the achievement gap through individualized instruction and formative assessment'. Assessment in Education: Principles, Policy & Practice, 17: 2, 169-182.

Summary of effects	
Study	***Effect size***
Aiello & Lee, 1980 (science)	0.35
Bangert et al., 1983	0.10
Horak, 1981	-0.07
Willett et al., 1983 (science)	0.17
Indicative effect size	**0.10**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Aiello & Wolfle, (1980).	Reported are the results of a meta-analysis of 30studies of individualized instruction in science in which this method was compared with a, traditional lecture method of science "instruction. Studies analysed also included measurements from which effect sizes could be calculated. Five methods of individualized instruction were identified :(1) audio-tutorial instruction(AT), (2)computer-assisted instruction (CAI), (3) personalized system of instruction (PSI), (4) programmed instruction (PI), and (5)a combination category for studies containing characteristics of individualization but not easily identifiable as one of the previous four methods, On the basis of effect size, individualized instruction appeared to be more effective than the traditional lecture approach for all methods studied. Findings reported were termed preliminary indicating this study was not completed when reported.

Bangert et.al. (1983).	This meta-analytic synthesis of findings from 51 studies indicated that use of an individualized teaching system has only a small effect on student achievement in secondary school courses. This result was consistent across a variety of academic settings and research designs and held true for both published and unpublished studies. In addition, individualized teaching systems did not contribute significantly to student self-esteem, critical thinking ability, or attitudes toward the subject matter being taught. Findings from studies of individualized college teaching are strikingly different from these secondary school findings.
Horak, (1981).	The present study investigated the effects of individualized instruction on mathematics achievement at the elementary and secondary school levels. The meta-analysis technique developed by Glass was applied to the same sample of studies used by Schoen in his previous voting-method analysis of individualization. The analysis of the 129 effect sizes revealed important trends for the use of self-paced modular instruction in mathematics. This study is also significant in its comparison of the conclusions drawn from a voting-method analysis and Glass's meta-analysis technique.
Willett et.al. (1983).	This article is a report of a meta-analysis on the question: “What are the effects of different instructional systems used in science teaching?” The studies utilized in this meta-analysis were identified by a process that included a systematic screening of all dissertations completed in the field of science education since 1950, an ERIC search of the literature, a systematic screening of selected research journals, and the standard procedure of identifying potentially relevant studies through examination of the bibliographies of the studies reviewed. In all, the 130 studies coded gave rise to 341 effect sizes. The mean effect size produced over all systems was 0.10 with a standard deviation of 0.41, indicating that, on the average, an innovative teaching system in this sample produced one-tenth of a standard deviation better performance than traditional science teaching. Particular kinds of teaching systems, however, produced results that varied from this overall result. Mean effect sizes were also computed by year of publication, form of publication, grade level, and subject matter.

Learning styles

Low impact for very low cost, based on limited evidence.

+2

The idea underpinning learning styles is that individuals all have a particular approach to or style of learning. The theory is that learning will therefore be more effective or more efficient if pupils are taught using the specific style or approach that has been identified as their learning style. For example, pupils categorised as having a 'listening' learning style, could be taught more through storytelling and discussion and less through traditional written exercises.

How effective is it?

There is very limited evidence for any consistent set of learning 'styles' that can be used reliably to identify genuine differences in the learning needs of young people, and evidence suggests that it is unhelpful to assign learners to groups or categories on the basis of a supposed learning style.

Overall the evidence shows an average impact of 2 months progress for learning style interventions. However, given the limited evidence for the existence of 'learning styles', it is reasonable to conclude that these gains may be the result of pupils taking responsibility for their own learning (see Meta-cognition) or from teachers using a wider range of activities to teach the same content, rather than the result of different learning styles.

Learning preferences do change in different situations and over time and there is some evidence that cognitive preference and task type may be connected (for example, visualisation is particularly valuable for some areas of mathematics). However, studies where teaching activities are targeted towards particular learners based on an identified learning 'style' have not convincingly shown any major benefit, particularly for low attaining pupils. Impacts recorded are generally low or negative, and the evidence suggests that only one or two pupils in a class of 25 might benefit from this approach.

The lack of impact of learning styles has been documented at all stages of education but it is particularly important not to label primary age pupils or for them to believe that their lack of success is due to their learning style.

How secure is the evidence?

Overall the picture is consistent though rigorous research is limited. The evidence for the lack of impact (and in some cases detrimental effect) of using learning styles approaches has been shown in a number of studies. The lack of validity and reliability of learning styles tests has also been the focus of a number of reviews.

What are the costs?

The costs are very low, usually involving preparation of a greater range and variety of teaching and learning materials, though some of the available tests of learning styles require purchase. Typically, these cost about £5 per pupil, although it is important to be aware of the limitations of these tests, given the lack of evidence for the existence of learning styles noted above.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Learners are very unlikely to have a single learning style, so restricting pupils to activities matched to their reported preferences may damage their progress. This is especially true for younger learners in primary schools whose preferences and approaches to learning are still very flexible.
2. Labelling students as a particular kind of learner is likely to undermine their belief that they can succeed through effort and to provide an excuse for failure.
3. It appears to be more promising to focus on other aspects of motivation to engage pupils in learning activities.
4. It certainly appears to be beneficial to have different representations of ideas when developing understanding, but this does not demonstrate that individual learners have a learning style.
5. How are you encouraging pupils to take responsibility for identifying how they can succeed in their learning and develop their own successful strategies and approaches?

Technical Appendix: Learning styles

Definition

A ‘learning style’ is an individual's unique approach to learning based on their strengths, weaknesses, and personal preferences, often in relation to different modes of information (visual, auditory, tactile, etc.) or in relation to its organisation (e.g. abstract, concrete; wholist, serialist, etc.). Different models in the literature describe these on a continuum from fixed to malleable according to how they conceptualise a particular ‘style’.

The idea underpinning learning styles is that individuals all have a particular approach to or style of learning. The theory is that learning will therefore be more effective or more efficient if pupils are taught using the specific style or approach that has been identified as their learning style. For example, pupils categorised as having a ‘listening’ learning style, could be taught more through storytelling and discussion and less through traditional written exercises. Although this is intuitively appealing, a number of serious issues challenge this field. The first is the robustness of the concept of a learning ‘style’ and which particular version is adopted. Most concepts have not been able to demonstrate sufficient reliability and/or validity. The next major problem is that the ‘meshing’ hypothesis where individuals learn better when targeting their ‘style’ lacks empirical validation. Where positive findings have been found it seems more likely that this impact is due to encouraging learners to take responsibility for choosing a learning strategy or approach, or to teachers presenting the same information in different ways, thereby increasing the repetition of information or enabling connections to be made across different representations. More recent exploration of ‘cognitive’ styles or preferences have attempted to address these issues, but so far with little success.

Learning styles; learning preferences; cognitive styles; cognitive preferences.

Evidence rating

There are four meta-analysis of learning styles and modality preference approaches, three of which found very low effects (-0.03 to +0.14). The fourth, with a pooled effect of 0.67, has been criticised for the validity of the underlying model, for technical problems with the meta-analysis and potential bias in selection and inclusion criteria. None have been conducted in the last ten years. Overall the evidence-base is limited.

Additional cost information

The costs are low, usually involving preparation of a greater range and variety of teaching and learning materials, though some of the available tests of learning styles require purchase. Typically, these cost about £5 per pupil, though, as noted above, it is important to be aware of the limitations of these tests. Overall the costs are estimated as very low.

References

Full references
Arnold, R.D. (1968). Four Methods of Teaching Word Recognition to Disabled Readers. *The Elementary School Journal*, 68. 5 pp. 269-274. http://www.jstor.org/stable/1000662
Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning. A systematic and critical review*. London: Learning and Skills Research Centre.

*Garlinger, D.K. & Frank, B.M. (1986). Teacher-student cognitive style and academic achievement: a review and a mini-meta analysis. *Journal of Classroom Interaction* 21.2, 2-8. http://www.jstor.org/stable/23869505
*Kavale, K.A. & Forness, S.R. (1987). Substance Over Style: Assessing the Efficacy of Modality Testing and Teaching. Exceptional Children, 54.3 pp. 228-39. http://dx.doi.org/10.1177/001440298705400305
Kavale, K., Hirshoren, A., & Forness, S. (1998). Meta-analytic validation of the Dunn-and-Dunn model of learning-style preferences: A critique of what was Dunn. *Learning Disabilities Research and Practice,* 13, pp 75-80. http://dx.doi.org/10.3200/JOER.101.2.94-98
Kavale, K. A., & LeFever, G. B. (2007). Dunn and Dunn model of learning-style preferences: Critique of Lovelace meta-analysis. *The Journal of Educational Research*, 101(2), 94-97.. http://dx.doi.org/
*Lovelace, M.K. (2002). A meta-analysis of experimental research studies based on the Dunn and Dunn learning-style model. St. John's University (New York), School of Education and Human Services). ProQuest Dissertations and Theses, p 177. Retrieved from http://search.proquest.com/docview/275698679 (20/6/12).
Lovelace, M. K. (2005). Meta-analysis of experimental research based on the Dunn and Dunn model. The Journal of Educational Research, 98(3), 176-183. http://dx.doi.org/10.3200/JOER.98.3.176-183
Mayer, R.E. (2011). Does styles research have useful implications for educational practice? *Learning and Individual Differences* 21 pp 319-320. http://dx.doi.org/10.1016/j.lindif.2010.11.016
Martin, S. (2010). Teachers using learning styles: Torn between research and accountability? *Teaching and Teacher Education* 26 pp 1583-1591. http://dx.doi.org/10.1016/j.tate.2010.06.009
Pashler, H., McDaniel, M., Rohrer, D. & Bjork, R. (2008). Learning Styles: Concepts and Evidence. *Psychological Science in the Public Interest* 9.3 pp 106-119. http://dx.doi.org/10.1111/j.1539-6053.2009.01038.x
*Slemmer, D.L. (2002). *The effect of learning styles on student achievement in various hypertext, hypermedia and technology enhanced learning environments: a meta-analysis* Unpublished PhD dissertation, Boise State University, Boise, Idaho (ProQuest Dissertations and Theses).

* Studies marked with an asterisk are included in the summary of effects

Summary of effects		
Summary of effects	**FSM effect size**	**Overall effect Size**
Kavale & Forness, 1987		0.14
Garlinger & Frank, 1986		-0.03
Lovelace, 2002		0.67
Slemmer, 2002		0.13
Median effect size		**0.13**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Kavale & Forness, 1987	A literature search identified 39 studies assessing modality preferences and modality teaching. The studies, involving 3,087 disabled and nondisabled elementary/secondary level subjects, were quantitatively synthesized. Subjects receiving differential instruction based on modality preferences exhibited only modest gains.
Garlinger & Frank, 1986	Reviews the effects on academic achievement associated with matching students and teachers on field-dependent–independent dimensions of cognitive style. To integrate and clarify the current status of findings relevant

	to this issue, a narrative summary of 7 studies is provided, followed by a meta-analysis. Findings suggest that field-independent students show greater achievement when matched with similar teachers.
Lovelace, 2002 (see also Lovelace, 2005)	The purpose of this investigator was to conduct a quantitative synthesis of experimental studies in which this model had been utilized between 1980 and 2000. Of the 695 different citations elicited by the database and reference-section searches, 76 original research investigations met the established inclusion criteria. A total of 7196 participants from these experimental research investigations provided 168 individual effect sizes for this meta-analysis. Data from these investigations were collected, coded, and summarized. The mean effect-size values for students' achievement and attitudes calculated and interpreted by this meta-analysis provided evidence for increased achievement and improved attitudes when responsive instruction was provided for diagnosed learning-style preferences. Not enough data were available to calculate mean-effect size values for behavior. Three tests determined the heterogeneity of the included investigations. Therefore, independent variables that impacted upon effect sizes were examined using both inductive and deductive moderator searches. No significant difference was found between subset categories of twelve of seventeen independent variables for achievement or for sixteen of seventeen variables for attitude. Therefore, there were a total of six moderating variables. No publication bias was revealed by correlations between sample sizes and effect sizes and the calculation of a Fail Safe N statistic. Finally, the current investigation and the previous meta-analysis conducted by Sullivan (1993) and reported in the Journal of Educational Research (Dunn, Griggs, Olson, Gorman & Beasley, 1995) and the National Forum of Applied Educational Research Journal (Sullivan, 1996-7) were compared. The mean effect size results for achievement from the current and the previous meta-analysis were consistent or robust. Therefore, it can be strongly suggested that learning-styles responsive instruction would increase the achievement and improve the attitudes toward learning of all students. Although authors of both studies revealed heterogeneous findings, there were indications that the data were less variable in this investigation and fewer moderating variables were revealed. In summary, although several moderating variables influenced the outcome, the results of this investigation overwhelmingly supported the position that matching students' learning style preferences with complementary instruction improved both academic achievement and student attitudes toward learning. According to Cohen's (1977, 1988, 1992) definitions, all averages for r and d effect sizes for both the previous and the present meta-analysis were medium to large. The Dunn and Dunn Learning Style Model had a robust medium to large effect that was both practically and educationally significant.
Slemmer, 2002	To identify forms of technology or types of technology-enhanced learning environments that may effectively accommodate the learning needs of students, 48 studies were included in a meta-analysis to determine the effects of learning styles on student achievement within technology-enhanced learning environments. A total of 51 weighted effect sizes were calculated from these studies with moderator variables coded for five study characteristics, six methodology characteristics, and six program characteristics. This meta-analysis found that learning styles do appear to influence student achievement in various technology-enhanced learning environments, but not at an overall level of practical significance. The total mean weighted effect size for the meta analysis was zr = .1341. Although the

	total mean weighted effect size did not reach the established level of practical significance (zr = .16), the value was greater than zr = .10, which is the level generally established by researchers as having a small effect. Additional findings from the moderator variables included: (1) Articles published in journals were the only type of publication that produced a significant mean weighted effect size (zr = .1939). (2) Studies that reported t statistics produced one of the highest total mean weighted effect sizes (zr =.4936) of any of the moderator variables. (3) Studies that reported an F statistic with df = 1 in the numerator had a significant total mean weighted effect size (zr = .2125); while studies that reported an F statistic with df > 1 in the numerator had a non-significant total mean weighted effect size (zr = .0637). (4) When all of the students received the same technology-enhanced lesson, there was a significant difference in student achievement between students with different learning styles (zr = .2952). (5) Studies that used Witkin's learning styles measure indicated a significant interaction between students' learning style and technology-enhanced learning environments as measured by student achievement (zr = .1873), while none of the quadrant-based learning style models indicated a significant interaction. (6) As the duration of treatment increased, the findings of the studies increased in significance. In general, this study provided evidence that under some conditions, students interact differently with technology in technology-enhanced learning environments depending on their specific learning style and the type of technology encountered.

Mastery learning

Moderate impact for very low cost, based on moderate evidence.

+5

Mastery learning breaks subject matter and learning content into units with clearly specified objectives which are pursued until they are achieved. Learners work through each block of content in a series of sequential steps.

Students must demonstrate a high level of success on tests, typically at about the 80% level, before progressing to new content. Mastery learning can be contrasted with other approaches which require pupils to move through the curriculum at a pre-determined pace. Teachers seek to avoid unnecessary repetition by regularly assessing knowledge and skills. Those who do not reach the required level are provided with additional tuition, peer support, small group discussions, or homework so that they can reach the expected level.

How effective is it?

There are a number of meta-analyses which indicate that, on average, mastery learning approaches are effective, leading to an additional five months' progress over the course of a school year compared to traditional approaches. Unusually however, among the evidence reviewed in the Toolkit, the effects of mastery learning tend to cluster at two points with studies showing either little or no impact or an impact of up to six months' gain. This clear split and wide variation implies that making mastery learning work effectively is challenging.

Mastery learning appears to be particularly effective when pupils work in groups or teams and take responsibility for supporting each other's progress (see also Collaborative learning and Peer tutoring). It also appears to be important that a high level of success is set. When pupils work at their own pace, as opposed to working as a part of group or whole class, it appears to be much less effective (see also Individualised instruction). Mastery learning may also be more effective when used as an occasional or additional teaching strategy as the impact decreases for longer programmes of over 12 weeks or so. Schools may wish to consider using mastery learning for particularly challenging topics or concepts, rather than for all lessons.

Lower attaining pupils may gain more from this strategy than high attaining students, by as much as one or two months' progress, so mastery learning appears to be a promising strategy for narrowing the gap. However, it should be noted that teachers also need to plan carefully for how to manage the time of pupils who make progress more quickly.

How secure is the evidence?

Overall, the evidence base is judged to be of moderate security. There is a large quantity of research on the impact of mastery learning, though much of it is relatively dated and findings are not consistent. In addition, most meta-analyses examining mastery learning use statistical techniques which may inflate the overall effect size so some caution is needed in interpreting the average impact. Having noted these concerns, a more recent study in the US found that mastery learning approaches can increase learning by up to six months in maths for 13-14 year olds, which is consistent with several older studies.

In February 2015, the EEF published an evaluation of the Mathematics Mastery programme, based on two randomised controlled trials conducted in English schools. On average, pupils in classes where the approach was used made one additional month's progress compared to similar classes that did not. It is possible that this estimate is more relevant to English schools than some older studies. An alternative explanation is that the Mathematics Mastery programme did not include some of the features of programmes that were previously associated with higher impacts. For example, although additional support was provided to struggling students, classes did not delay starting new topics until a high level of proficiency had been reached by all pupils.

What are the costs?

Few additional resources are required to introduce a mastery learning approach. Professional development and additional support for staff is recommended, particularly in the early stages of setting up a programme. Estimates are less than £80 per pupil, indicating very low overall costs. Additional small group tuition and one to one support are also likely to be needed. This may not result in additional financial cost if schools use existing staff resources, but teachers should think carefully about the impact of this extra support in terms of the extra time and effort it will require.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Overall, mastery learning is a learning strategy with good potential, particularly for low attaining students.
2. Implementing mastery learning effectively is not straightforward, however, requiring a number of complex components and a significant investment in terms of design and preparation.
3. Setting clear objectives and providing feedback from a variety of sources so that learners understand their progress appear to be key features of using mastery learning effectively. A high level of success, at least 80%, should be required before pupils move on.
4. Incorporating group and team approaches where pupils take responsibility for helping each other within mastery learning appears to be effective.

Technical Appendix: Mastery Learning

Definition

Traditional teaching keeps time constant and allows pupils' 'mastery' of curriculum content to vary. Mastery learning keeps learning outcomes constant and but varies the time needed for pupils to become proficient or competent at these objectives. The mastery learning method breaks subject matter into blocks or units with predetermined objectives and specified outcomes. Learners must demonstrate mastery on unit tests, typically 80%, before moving on to new material. Any pupils who do not achieve mastery are provided with extra support through a range of teaching strategies such as more intensive teaching, tutoring, peer-assisted learning, small group discussions, or additional homework. Learners continue the cycle of studying and testing until the mastery criteria are met.

Some of the ideas behind mastery learning date back to American schools in the 1920's with the work of Washburne (1922, as cited in Block, 1971). A version of mastery learning was revived in the form of programmed instruction in the late 1950's based on the work of Skinner. It aimed to provide students with instructional materials that would let them move at their own pace and receive constant feedback on their level of mastery (see also Individualised instruction). During the 1960's Bloom's (1968) 'Learning for Mastery' led to a resurgence of interest from both researchers and practitioners. He is now generally acknowledged as the originator of the mastery model. Bloom argued that learners should not spend more time overall on activities to achieve proficiency. Although it may take longer in the early stages, he suggested learners would need less time to master more advanced material because of their higher levels of basic competence.

In terms of assessment and feedback, a number of aspects of mastery learning are similar to other contemporary approaches such as the use of initial diagnostic assessments like universal screening in Response to Intervention models (Mellard & Johnson, 2008). The use of formative assessments and tests to monitor pupils' progress systematically then give detailed feedback on what they need to do to close the gap between their current performance and the desired goal is similar to assessment for learning and feedback models (Black and Wiliam, 1998; Hattie & Timperley, 2007).

Mastery learning is therefore not a new approach, though different versions have been developed and used at different times. It is based on the belief that all pupils can learn when provided with appropriate activities and support. All pupils must achieve a pre-specified level of mastery on one unit before they can to progress to the next. Learners are also provided with specific feedback about their progress at regular intervals. This helps learners identify where they have been successful and where they have been less successful. Any objectives in the curriculum which have not been learned are given more time and more effort to achieve mastery.

Search terms:

Mastery learning, learning for mastery

Evidence rating

There are five meta-analyses included in the summary, but none of these have been conducted in the last 10 years. The pooled effects from these syntheses range from 0.04 to 0.60 so do not provide a consistent estimate of effect. This variation is not explained by moderator analyses. A number of the meta-analyses include experimental and quasi-experimental studies which are not well controlled. The pooled effects in the early studies are simple means or median values rather than weighted

models (fixed effect or random effects). In addition studies have not been adjusted for clustering. One recent study has not replicated these effects, but this was an experimental study rather than a randomised trial. Overall the evidence is rated as moderate.

Additional cost information

The main financial cost of implementing a mastery learning approach will be the cost of professional development. The average cost of professional development in EEF-funded programmes is well under £80 per pupil.

The average 'per pupil' cost of the Maths Mastery programme was estimated to be around £131 per year for primary school pupils and around £50 per year for secondary school pupils, in the first year, with per pupil costs likely to reduce in future years in both cases.

Additional one to one and small group support are also likely to be needed. Many schools will provide this support using existing staff and resources. Although this approach will not incur an additional financial cost, teachers should be aware of the cost in terms of extra time required. They must also think carefully about the activity they might have to stop doing in order to provide this additional support.

References

Full references
Block, J. (1971). *Mastery learning: Theory and practice*. New York: Holt, Rinehart, & Winston.
Bloom, B. (1968). Learning for mastery. *Evaluation Comment*,1(2), 1-5.
Bloom, B. (1971). *Mastery learning.* New York: Holt, Rinehart, & Winston.
*Guskey, T. R., & Pigott, T. D. (1988). Research on group-based mastery learning programs: A meta-analysis. *The Journal of Educational Research*, 197-216. http://dx.doi.org/10.1080/00220671.1988.10885824
Davis, D., & Sorrell, J. (1995). Mastery learning in public schools. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. http://teach.valdosta.edu/whuitt/files/mastlear.html
Guskey, T.R. (2007). Closing Achievement Gaps: Revisiting Benjamin S. Bloom's "Learning for Mastery. *Journal of Advanced Academics*. 19, 8-31. http://dx.doi.org/10.4219/jaa-2007-704
Hattie, J. (2008). *Visible Learning*. London: Routledge.
Jerrim, J., Austerberry, H., Crisan, C., Ingold, A., Morgan, C., Pratt, D., Smith, C. & Wiggins, M. (2015) *Mathematics Mastery Secondary Evaluation Report (February 2015) London*: EEF. https://educationendowmentfoundation.org.uk/uploads/pdf/Mathematics_Mastery_Secondary_(Final)1.pdf
*Kulik, C. L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, 60(2), 265-299. http://dx.doi.org/10.3102/00346543060002265
Miles, K. (2010). Mastery Learning and Academic Achievement. Ed.D. Dissertation, Walden University Document URL ProQuest Dissertations & Theses: http://search.proquest.com/docview/193327442.
*Slavin, R. E. (1987). Mastery learning reconsidered. *Review of Educational Research*, 57(2), 175-213. http://dx.doi.org/10.3102/00346543057002175
Slavin, R. E. (1990). Mastery learning re-reconsidered. *Review of Educational Research*, 300-302. http://www.jstor.org/stable/1170613
Slavin, R. E., & Karweit, N. L. (1984). Mastery learning and student teams: A factorial experiment in urban general mathematics classes. *American Educational Research Journal*, 21(4), 725-736.

Technical Appendix: Mastery Learning

http://dx.doi.org/10.3102/00028312021004725
Vignoles. A., Jerrim, J. & Cowan, C. (2015) *Mathematics Mastery Primary Evaluation Report (February 2015) London*: EEF. https://educationendowmentfoundation.org.uk/uploads/pdf/Mathematics_Mastery_Primary_(Final)1.pdf
*Waxman, H. C., Wang, M. C., Anderson, K. A., Walberg, H. J., & Waxman, C. (1985). Adaptive education and student outcomes: A quantitative synthesis. The Journal of Educational Research, 228-236. http://dx.doi.org/10.1080/00220671.1985.10885607
*Willett, J. B., Yamashita, J. J., & Anderson, R. D. (1983). A meta-analysis of instructional systems applied in science teaching. Journal of Research in Science Teaching, 20(5), 405-417. http://dx.doi.org/10.1002/tea.3660200505
What Works Clearinghouse (2012). *WWC review of the report: Mastery learning and student teams: A factorial experiment in urban general mathematics classes*. U.S. Department of Education, Institute of Education Sciences: http://files.eric.ed.gov/fulltext/ED535811.pdf

Summary of Effects		
Meta-analyses	***FSM effect size***	***Effect size***
Guskey & Piggott, 1988		0.60
Kulik, Kulik & Bangert-Drowns, 1990		0.52
Slavin, 1987		0.04
Willett et al. 1983 (for mastery learning)		0.59
Waxman et al. 1985	0.11	0.39
Indicative effect size (mean)		**0.44**
*Recent studies (*Not included in any of the meta-analyses)		
Vignoles et al. 2015		0.10
Jerrim et al. 2015	0.07	0.06
Miles, 2010		0.53
Slavin & Karweit, 1984 (see also What Works Clearninghouse, 2012)		0.01

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Guskey & Piggott, 1988 (updates Guskey & Gates, 1985)	This paper presents a synthesis of findings from 46 studies on group based applications of mastery learning strategies. Meta-analytic procedures were used to combine the results of the studies and to calculate estimates of the effects of group-based applications. Results show that such applications yield consistently positive effects on both cognitive and affective student learning outcomes, as well as several teacher variables. Variation in the size of the effect across studies was found to be quite large, however, and homogeneity tests indicated that studies do not share a common effect size. Several factors were explored as possible explanations for this variation, including the subject area to which mastery learning was applied, the grade level of students involved and the duration of the study. Other possible explanations for this variation are discussed, along with implications for future directions in the research.

Kulik, Kulik & Bangert-Drowns, 1990	A meta-analysis of findings from 108 controlled evaluations showed that mastery learning programs have positive effects on the examination performance of students in colleges, high schools, and the upper grades in elementary schools. The effects appear to be stronger on the weaker students in a class, and they also vary as a function of mastery procedures used, experimental designs of studies, and course content. Mastery programs have positive effects on student attitudes toward course content and instruction but may increase student time on instructional tasks. In addition, self-paced mastery programs often reduce the completion rates in college classes.
Slavin, 1987	Several recent reviews and meta-analyses have claimed extraordinarily positive effects of mastery learning on student achievement, and Bloom (1984a, 1984b) has hypothesized that mastery-based treatments will soon be able to produce "2-sigma" (i.e., 2 standard deviation) increases in achievement. This article examines the literature on achievement effects of practical applications of group-based mastery learning in elementary and secondary schools over periods of at least 4 weeks, using a review technique, "best-evidence synthesis," which combines features of meta-analytic and traditional narrative reviews. The review found essentially no evidence to support the effectiveness of group-based mastery learning on standardized achievement measures. On experimenter-made measures, effects were generally positive but moderate in magnitude, with little evidence that effects maintained over time. These results are discussed in light of the coverage versus mastery dilemma posed by group-based mastery learning.
Waxman et al.,1985	To estimate the effects of adaptive education on cognitive, affective, and behavioral outcomes of learning, 309 effect sizes were calculated using statistical data from 38 studies that contained a combined sample of approximately 7,200 students. The substantial mean of the study weighted effect sizes is .45, suggesting that the average student in adaptive programs scores at the 67th percentile of control group distributions. The effect appeared constant across grades, socioeconomic levels, races, private and public schools, and community types. In addition, the effects were not significantly different across the categories of adaptiveness, student outcomes, social contexts and methodological rigor of the studies.
Willett et al. 1983	This article is a report of a meta-analysis on the question: “What are the effects of different instructional systems used in science teaching?” The studies utilized in this meta-analysis were identified by a process that included a systematic screening of all dissertations completed in the field of science education since 1950, an ERIC search of the literature, a systematic screening of selected research journals, and the standard procedure of identifying potentially relevant studies through examination of the bibliographies of the studies reviewed. In all, the 130 studies coded gave rise to 341 effect sizes. The mean effect size produced over all systems was 0.10 with a standard deviation of 0.41, indicating that, on average, an innovative teaching system in this sample produced one-tenth of a standard deviation better performance than traditional science teaching. Particular kinds of teaching systems, however, produced results that varied from this overall result. Mean effect sizes were also computed by year of publication, form of publication, grade level, and subject matter.

Mentoring

Low impact for moderate cost, based on moderate evidence.

+1

Mentoring in education aims to develop young people's strengths by pairing them with an older volunteer, sometimes from a similar background, who can act as a positive role model. It is often characterised as aiming to build confidence, or to develop resilience and character, rather than directly focusing on teaching or tutoring specific skills. Mentors typically build relationships with young people by meeting with them one-to-one for about an hour or so a week either at school, at the end of the school day, or at weekends.

Activities vary from programme to programme, sometimes including direct academic support with homework or other school tasks. Mentoring has increasingly been offered to young people who are hard to reach or deemed to be at risk of educational failure or exclusion.

How effective is it?

The impact of mentoring is low in terms of direct effect on academic outcomes, accelerating learning on average by only about one month's additional progress for the majority of pupils. There is some evidence that pupils from disadvantaged backgrounds can benefit by up to about two months' additional progress.

School-based mentoring programs have on average been less effective than community-based approaches, possibly because school-based mentoring can result in fewer opportunities for young people to develop more lasting and trusting relationships with adult role models. Other positive benefits have been reported in terms of attitudes to school, attendance and behaviour.

Programmes which have a clear structure and expectation, provide training and support for mentors, and use mentors from a professional background, are associated with more successful outcomes. There are risks associated with unsuccessful mentor pairings, which may have a detrimental effect on the mentee, and the negative overall impacts reported by some studies should prompt caution.

How secure is the evidence?

The evidence is moderately secure. The quality of evaluations has improved in recent years with more rigorous designs compared with earlier studies, which often relied on correlational designs. Impact estimates have been fairly consistent over the last decade.

Most of the studies come from the USA and focus on secondary school pupils, with a few studies from the UK and other European countries such as Portugal. A recent rigorous study of mentoring for reading in Northern Ireland with eight to nine year olds found small improvements of about two months' progress in fluency, but not in reading comprehension. Further rigorous evaluation in the UK is needed of varying approaches to mentoring across different age groups.

What are the costs?

Costs overall are estimated as moderate. Compared with other professionally delivered interventions and approaches, mentoring is relatively inexpensive. Costs mainly cover mentor training and support, and the organisation and administration of the programme. Community-based programmes tend to be more expensive than school-based programmes as schools tend to absorb some of the costs, such as space costs or general administration. Estimates in the USA are between $1000-$1500 per student per year or about £600-£850 per pupil per year.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. The impact of mentoring varies, but overall, it is likely to have only a small impact on attainment.
2. Positive effects tend not to be sustained once the mentoring stops, so care must be taken to ensure that benefits are not lost.
3. Community-based approaches tend to be more successful than school-based approaches.
4. Mentor drop-out can have detrimental effects on mentees. What steps have you taken to assess the reliability of mentors?
5. What training and support have you provided for mentors?

Mentoring
Toolkit references

Full references
Bernstein, L., Rappaport, C. D., Olsho, L., Hunt, D., & Levin, M. (2009). Impact Evaluation of the US Department of Education's Student Mentoring Program. Final Report. NCEE 2009-4047. Washington US Department of Education National Center for Education Evaluation and Regional Assistance. http://ies.ed.gov/ncee/pubs/20094047/
DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. American Journal of Community Psychology, 30(2), 157-197. http://www.dx.doi.org/10.1023/A:1014628810714
Eby, L. T., Allen, T. D., Evans, S. C., Ng, T., & DuBois, D. L. (2008). Does mentoring matter? A multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *Journal of Vocational Behavior*, 72(2), 254-267. http://www.dx.doi.org/10.1016/j.jvb.2007.04.005
Herrera, C., Grossman, J. B., Kauh, T. J., & McMaken, J. (2011). Mentoring in Schools: An Impact Study of Big Brothers Big Sisters School-Based Mentoring. Child Development, 82(1), 346-361. http://www.ncbi.nlm.nih.gov/pubmed/21291446
Karcher, M. J. (2008). The study of mentoring in the learning environment (SMILE): A randomized evaluation of the effectiveness of school-based mentoring. Prevention Science, 9(2), 99-113. http://www.dx.doi.org/10.1111/j.1467-8624.2010.01559.x
Maxwell, B., Connolly, P., Demack, S., O'Hare, L., Stevens, A. & Clague, L. (2014) *TextNow Transition Programme Evaluation Report and Executive Summary October 2014*. London: Education Endowment Foundation http://educationendowmentfoundation.org.uk/project...
McQuillin, S., Smith, B., & Strait, G. (2011). Randomized evaluation of a single semester transitional mentoring program for first year middle school students: a cautionary result for brief, school-based mentoring programs. *Journal of Community Psychology*, 39(7), 844-859. http://www.dx.doi.org/10.1002/jcop.20475
Núñez, J. C., Rosário, P., Vallejo, G., & González-Pienda, J. A. (2013). A longitudinal assessment of the effectiveness of a school-based mentoring program in middle school. *Contemporary Educational Psychology,* 38(1), 11-21. http://www.dx.doi.org/10.1016/j.cedpsych.2012.10.002
Miller, S., Connolly, P., & Maguire, L. K. (2011). The effects of a volunteer mentoring programme on reading outcomes among eight-to nine-year-old children: A follow up randomized controlled trial. *Journal of Early Childhood Research*, http://www.dx.doi.org/10.1177/1476718X11407989

Mentoring
Toolkit references

Summary of effects		
Study	***Effect size on FSM***	***Overall effect size***
Bernstein et al., 2009	0.00	0.05 (maths) -0.04 (reading) -0.03 (science)
DuBois et.al., 2002	0.11	0.11 (academic)
Eby et al. 2008		0.16 (performance)
Wheeler, Keller & DuBois. 2010		-0.02 (maths) -0.01 (reading)
Wood & Mayo-Wilson, 2012		-0.01 (academic performance)
Recent studies		
McQuillin et al. 2011		-0.44 (reading) -0.12 (English & language) -0.37 (maths) 0.11 (science)
Miller et al. 2012		0.14 (reading fluency)
Nunez et al. 2013 Nunez et al. 2013		-0.03 (first language) 0.31 (maths)
Maxwell et al. 2014 (EEF TextNow Transition Programme)		0.06
Indicative effect size	**0.05**	**0.02**

For more information about the effect sizes in the Toolkit, click here.

Meta analyses and abstracts	
Study	***Abstract***
Bernstein et.al. (2009).	This report summarizes the findings from a national evaluation of mentoring programs funded under the U.S. Department of Education's (ED) Student Mentoring Program. The Office of Management and Budget (OMB) requested that the Institute of Education Sciences (IES) within ED oversee an independent evaluation of the Student Mentoring Program. In 2005, ED contracted with Abt Associates and its team of subcontractors, Branch Associates, Moore and Associates, and the Center for Resource Management, to conduct the Impact Evaluation of Student Mentoring Programs. The impact evaluation used an experimental design in which students were randomly assigned to a treatment or control group. Thirty-two purposively selected School Mentoring Programs and 2,573 students took part in the evaluation, which estimated the impact of the programs over one school year on a range of student outcomes. The evaluation also describes the characteristics of the program and the mentors, and provides information about program delivery.
DuBois et.al. (2002).	We used meta-analysis to review 55 evaluations of the effects of mentoring programs on youth. Overall, findings provide evidence of only a modest or small benefit of program participation for the average youth. Program effects are enhanced significantly; however, when greater numbers of both theory-based and empirically based "best practices" are utilized and when strong relationships are formed between mentors and youth. Youth from backgrounds of environmental risk and disadvantage appear most likely to benefit from participation in mentoring programs. Outcomes for youth at-risk due to personal vulnerabilities have varied substantially in relation to program characteristics, with a noteworthy potential evident for poorly implemented programs to actually have an adverse effect on such youth. Recommendations include greater adherence to guidelines for the design and implementation of effective mentoring programs as well as more in-depth assessment of relationship and contextual factors in the evaluation of programs.

Mentoring
Toolkit references

Eby et al. 2008	The study of mentoring has generally been conducted within disciplinary silos with a specific type of mentoring relationship as a focus. The purpose of this article is to quantitatively review the three major areas of mentoring research (youth, academic, workplace) to determine the overall effect size associated with mentoring outcomes for protégés. We also explored whether the relationship between mentoring and protégé outcomes varied by the type of mentoring relationship (youth, academic, workplace). Results demonstrate that mentoring is associated with a wide range of favourable behavioural, attitudinal, health-related, relational, motivational, and career outcomes, although the effect size is generally small. Some differences were also found across type of mentoring. Generally, larger effect sizes were detected for academic and workplace mentoring compared to youth mentoring. Implications for future research, theory, and applied practice are provided.
Wheeler et.al. (2010).	Between 2007 and 2009, reports were released on the results of three separate large-scale random assignment studies of the effectiveness of school-based mentoring programs for youth. The studies evaluated programs implemented by Big Brothers Big Sisters of America (BBBSA) affiliates (Herrera et al., 2007), Communities In Schools of San Antonio, Texas (Karcher,2008), and grantees of the U.S. Department of Education's Student Mentoring Program (Bernstein et al., 2009). Differences in the findings and conclusions of the studies have led to varying responses by those in practice and policy roles. The results of the BBBSA trial led the organization to undertake an initiative to pilot and evaluate an enhanced school-based mentoring model. Findings of the Student Mentoring Program evaluation were cited as a reason for eliminating support for the program in the FY 2010 federal budget (Office of Management and Budget, 2009). In this report, we present a comparative analysis of the three studies. We identify important differences across the studies in several areas, including agency inclusion criteria, program models, implementation fidelity and support, and criteria utilized in tests of statistical significance. When aggregating results across the studies using meta-analytic techniques, we find evidence that school-based mentoring can be modestly effective for improving selected outcomes (i.e., support from non-familial adults, peer support, perceptions of scholastic efficacy, school-related misconduct, absenteeism, and truancy). Program effects are not apparent, however, for academic achievement or other outcomes. Our analysis underscores that evidence-based decision-making as applied to youth interventions should take into account multiple programmatic and methodological influences on findings and endeavour to take stock of results from the full landscape of available studies.
Wood & Mayo-Wilson (2012).	Objectives: To evaluate the impact of school-based mentoring for adolescents (11–18 years) on academic performance, attendance, attitudes, behaviour, and self-esteem. Method: A systematic review and meta-analysis. The authors searched 12 databases from 1980 to 2011. Eight studies with 6,072 participants were included, 6 were included in meta-analysis. Studies were assessed using the Cochrane Collaboration Risk of Bias Tool. Results: Across outcomes, effect sizes were very small (random effects), and most were not significant. The magnitude of the largest effect (for self-esteem) was close to zero, g = 0.09, [0.03, 0.14]. Conclusions: The mentoring programs included in this review did not reliably improve any of the included outcomes. Well-designed programs implemented over a longer time might achieve positive results.

Meta-cognition and self-regulation

High impact for very low cost, based on extensive evidence.

+8

Meta-cognition and self-regulation approaches (sometimes known as 'learning to learn' approaches) aim to help learners think about their own learning more explicitly. This is usually by teaching pupils specific strategies to set goals, and monitor and evaluate their own academic development. Self-regulation means managing one's own motivation towards learning. The intention is often to give pupils a repertoire of strategies to choose from during learning activities.

How effective is it?

Meta-cognition and self-regulation approaches have consistently high levels of impact, with pupils making an average of eight months' additional progress. The evidence indicates that teaching these strategies can be particularly effective for low achieving and older pupils.

These strategies are usually more effective when taught in collaborative groups so learners can support each other and make their thinking explicit through discussion.

The potential impact of these approaches is very high, but can be difficult to achieve as they require pupils to take greater responsibility for their learning and develop their understanding of what is required to succeed. There is no simple method or trick for this. It is possible to support pupils' work too much, so that they do not learn to monitor and manage their own learning but come to rely on the prompts and support from the teacher. "Scaffolding" provides a useful metaphor: a teacher would provide support when first introducing a pupil to a concept, then reduce the support to ensure that the pupil continues to manage their learning autonomously.

How secure is the evidence?

A number of systematic reviews and meta-analyses have consistently found similar levels of impact for strategies related to meta-cognition and self-regulation. Most studies have looked at the impact on English or mathematics, though there is some evidence from other subject areas like science, suggesting that the approach is likely to be widely applicable.

In the UK, four recent studies indicate that programmes that seek to improve learning to learn skills can effectively improve academic outcomes. A 2014 study, Improving Writing Quality, used a structured programme of writing development based on a self-regulation strategy. The evaluation found gains, on average, of an additional nine months' progress, suggesting that the high average impact of self-regulation strategies can be achieved in English schools. In 2015, evaluations of an intervention based on "Growth Mindsets" research, Philosophy for Children, and a programme called Thinking, Doing, Talking Science found gains of between two and five additional months' progress. In three projects there were indications that the programmes were particularly beneficial for pupils from low income families.

What are the costs?

Overall, costs are estimated as very low. Many studies report the benefits of professional development or an inquiry approach for teachers, where they actively evaluate strategies as they learn to use them. Most projects are estimated as costing under £80 per pupil.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Teaching approaches which encourage learners to plan, monitor and evaluate their learning have very high potential, but require careful implementation.
2. Have you taught pupils explicit strategies on how to plan, monitor and evaluate specific aspects of their learning? Have you given them opportunities to use them with support and then independently?
3. Teaching how to plan: Have you asked pupils to identify the different ways that they could plan (general strategies) and then how best to approach a particular task (specific technique)?
4. Teaching how to monitor: Have you asked pupils to consider where the task might go wrong? Have you asked the pupils to identify the key steps for keeping the task on track?
5. Teaching how to evaluate: Have you asked pupils to consider how they would improve their approach to the task if they completed it again?

Technical Appendix:
Meta-cognition and self-regulation

Definition

Meta-cognition (sometimes known as 'learning to learn') and self-regulation approaches aim to help learners think about their own learning more explicitly so as to take increased responsibility for achievement. Meta-cognition involves consciously thinking about planning, monitoring and evaluating your own learning and is often considered to have two dimensions, knowledge and skillfulness or the extent to which a learner is aware of meta-cognitive strategies to manage learning and the individual's capability at putting these strategies into practice. Approaches usually focus on teaching pupils specific strategies to set goals, and monitor and evaluate their own academic development in relation to particular learning tasks and activities. Self-regulation relates to meta-cognitive skillfulness but also involves managing one's own motivation towards learning and the development of dispositions such as resilience and perseverance. In practical terms, the intention is often to provide pupils with a repertoire of strategies to choose from during learning activities, this often involves Feedback on use of different strategies. Approaches also frequently involve Collaborative learning activities and techniques.

Search terms:

Meta-cognition*, executive function, self-regulation*

Evidence rating

Extensive: There are seven meta-analyses with five undertaken in the last 10 years. These are mainly from experimental studies which were often undertaken in schools and which evaluated impact on pupil attainment data as well as more general cognitive outcomes with some exploration of the causes of any identified heterogeneity. The underlying studies, however, vary in quality. Most of the estimates of impact are high with pooled effect falling in a narrow range (0.62 to 0.71). Recent single studies have not consistently achieved these gains.

Additional cost information

The main financial cost of implementing a metacognition and self-regulation approach will be the cost of professional development. The average cost of professional development in EEF-funded programmes is well under £80 per pupil.

The cost of the Using Self-Regulation to Improve Writing programme, which aims to improve pupils' writing by promoting self-regulation, cost £52 per pupil.

References

Full references
*Abrami, P.C., Bernard, R.M., Borokhovski, E., Wade, A., Surkes, M.A., Tamim, R., & Zhang, D. (2008). Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage 1 Meta-Analysis. *Review of Educational Research* 78.4 pp 1102-1134. http://www.dx.doi.org/10.3102/00346543083260
*Chiu, C.W.T. (1998). *Synthesizing Metacognitive Interventions: What Training Characteristics Can Improve Reading Performance?* Paper presented at the Annual Meeting of the American Educational Research Association San Diego, CA, April 13-17, 1998. http://files.eric.ed.gov/fulltext/ED420844.pdf

Crawford, C. & Skipp, A. (2014) *LIT Programme Evaluation Report and Executive Summary* October 2014. London: EEF. http://educationendowmentfoundation.org.uk/uploads/pdf/Final_Copy__EEF_Evaluation_Report_-_The_Learning_Trust.pdf

*Dignath, C., Buettner, G. & Langfeldt, H. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review* 3.2 pp. 101-129. http://www.dx.doi.org/10.1016/j.edurev.2008.02.003

*Donker, A. S., De Boer, H., Kostons, D., Dignath van Ewijk, C. C., & Van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, 11, 1-26. http://www.dx.doi.org/10.1016/j.edurev.2013.11.002

Gorard et al. (2015) *Philosophy for Children (P4C) Evaluation Report* London: EEF

*Haller, E.P., Child, D.A. & Walberg, H.J. (1988). Can Comprehension be taught? A Quantitative Synthesis of "Metacognitive Studies." *Educational Researcher,* 17.9 pp 5-8. http://www.dx.doi.org/10.3102/0013189X017009005

Hanley, Slavin & Elliott (2015) *Thinking Doing Talking Science Evaluation Report* London: EEF

*Higgins, S., Hall, E., Baumfield, V., & Moseley, D. (2005). *A meta-analysis of the impact of the implementation of thinking skills approaches on pupils.* In: Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=339

Jacob, R., & Parkinson, J. (2015). The Potential for School-Based Interventions That Target Executive Function to Improve Academic Achievement A Review. *Review of Educational Research* http://dx.doi.org/10.3102/0034654314561338

*Klauer, K.J. & Phye, G.D. (2008). Inductive Reasoning: A Training Approach. *Review of Educational Research*, 78.1 pp 85-123. http://www.dx.doi.org/10.3102/0034654307313402

NIESR (2015) Changing Mindsets Evaluation Report London: EEF

Perry, V., Albeg, L., & Tung, C. (2012). Meta-Analysis of Single-Case Design Research on Self-Regulatory Interventions for Academic Performance. *Journal of Behavioral Education*, 21(3), 217-229. http://www.dx.doi.org/10.1007/s10864-012-9156-y

Schunk, D.H. (2008). Metacognition, self-regulation, and self-regulated learning: Research recommendations. *Educational Psychology Review*, 20.4 pp 463-467. http://www.dx.doi.org/10.1007/s10648-008-9086-3

Torgerson, D., Torgerson, C. Ainsworth, H. Buckley, H. M Heaps, C. Hewitt, C. & Mitchell, N. (2014) *Improving Writing Quality: Evaluation Report and Executive Summary May 2014* London: EEF. http://educationendowmentfoundation.org.uk/uploads/pdf/EEF_Evaluation_Report_-_Improving_Writing_Quality_-_May_2014_v2.pdf

Tracy, B., Reid, R., & Graham, S. (2009). Teaching young students strategies for planning and drafting stories: The impact of self-regulated strategy development. *The Journal of Educational Research*, *102*(5), 323-332. http://dx.doi.org/10.3200/JOER.102.5.323-332

Summary of effects		
Meta-analyses	**FSM effect size**	**Overall effect size**
Abrami et al. 2008		0.34
Chiu 1998		0.67
Dignath et al. 2008		0.62
Donker et al. 2014	0.72	0.66
Haller et al. 1988		0.71
Higgins et al. 2005		0.62
Klauer & Phye 2008		0.69
Recent studies		

Crawford et al. 2014 (EEF- LIT Programme)		0.09
Hanley, Slavin & Elliott 2015 (EEF Thinking Doing Talking Science)	0.38	0.22
Gorard et al. 2015 (EEF P4C) Reading Maths		0.14 0.13
NIESR 2015 (EEF Changing Mindsets) Pupil intervention – English Pupil intervention – Maths	 0.17 0.11	 0.18 0.10
Torgerson et al. 2014 (EEF- Using Self-Regulation to Improve Writing)	1.60	0.74
Tracy, Reid & Graham, 2009 (writing overall)		0.47
Weighted mean effect size		**0.62**

Meta-analyses and abstracts	
Study	**Abstract**
Abrami et al. 2008	Critical thinking (CT), or the ability to engage in purposeful, self-regulatory judgment, is widely recognized as an important, even essential, skill. This article describes an on-going meta-analysis that summarizes the available empirical evidence on the impact of instruction on the development and enhancement of critical thinking skills and dispositions. We found 117 studies based on 20,698 participants, which yielded 161 effects with an average effect size (g+) of 0.341 and a standard deviation of 0.610. The distribution was highly heterogeneous (QT = 1,767.86, p < .001). There was, however, little variation due to research design, so we neither separated studies according to their methodological quality nor used any statistical adjustment for the corresponding effect sizes. Type of CT intervention and pedagogical grounding were substantially related to fluctuations in CT effects sizes, together accounting for 32% of the variance. These findings make it clear that improvement in students' CT skills and dispositions cannot be a matter of implicit expectation. As important as the development of CT skills is considered to be, educators must take steps to make CT objectives explicit in courses and also to include them in both pre-service and in-service training and faculty development.
Chiu 1998	In this paper, meta-analysis is used to identify components that are associated with effective metacognitive training programs in reading research. Forty-three studies, with an average of 81 students per study, were synthesized. It was found that metacognitive training could be more effectively implemented by using small-group instruction, as opposed to large-group instruction or one-to-one instruction. Less intensive programs were more effective than intensive programs. Program intensity was defined as the average number of days in a week that instruction was provided to students. Students in higher grades were more receptive to the intervention. Measurement artifacts, namely teaching to the test and use of non-standardized tests and the quality of the studies synthesized played a significant role in the evaluation of the effectiveness of the metacognitive reading intervention.

Dignath et al. 2008	Recently, research has increasingly focused on fostering self-regulated learning amongst young children. To consider this trend, this article presents the results of a differentiated meta-analysis of 48 treatment comparisons resulting from 30 articles on enhancing self-regulated learning amongst primary school students. Based on recent models of self-regulated learning, which consider motivational, as well as cognitive, and metacognitive aspects [Boekaerts, M. (1999). Self-regulated learning: Where we are today. International Journal of Educational research, 31(6), 445–457], the effects of self-regulated learning on academic achievement, on cognitive and metacognitive strategy application, as well as on motivation were analyzed. As the results show, self-regulated learning training programmes proved to be effective, even at primary school level. Subsequent analysis tested for the effects of several moderator variables, which consisted of study features and training characteristics. Regarding factors that concern the content of the treatment, the impact of the theoretical background that underlies the intervention was tested, as well as the type of cognitive, metacognitive, or motivational strategy which were instructed, and if group work was used as instruction method. Training context related factors, which were included in the analyses, consisted of students' grade level, the length of the training, if teachers or researchers directed the intervention, as well as the school subject in which context the training took place. Following the results of these analyses, a list with the most effective training characteristics was provided.
Donker et al. 2014	In this meta-analysis the results of studies on learning strategy instruction focused on improving self-regulated learning were brought together to determine which specific analysis included 58 studies in primary and secondary education on interventions aimed at improving cognitive, metacognitive, and management strategy skills, as well as motivational aspects and metacognitive knowledge. A total of 95 interventions and 180 effect sizes demonstrated substantial effects in the domains of writing (Hedges' g = 1.25), science (.73), mathematics (.66) and comprehensive reading (.36). These domains differed in terms of which strategies were the most effective in improving academic performance. However, metacognitive knowledge instruction appeared to be valuable in all of them. Furthermore, it was found that the effects were higher when self-developed tests were used than in the case of intervention-independent tests. Finally, no differential effects were observed for students with different ability levels. To conclude, the authors have listed some implications of their analysis for the educational practice and made some suggestions for further research.
Haller et al. 1988	To assess the effect of "metacognitive" instruction on reading comprehension, 20 studies, with a total student population of 1,553, were compiled and quantitatively synthesized. For 115 effect sizes, or contrasts of experimental and control groups' performance, the mean effect size was .71, which indicates a substantial effect. In this compilation of studies, metacognitive instruction was found particularly effective for junior high students (seventh and eighth grades). Among the metacognitive skills, awareness of textual inconsistency and the use of self-questioning as both a monitoring and a regulating strategy were most effective. Reinforcement was the most effective teaching strategy.
Higgins et al. 2005	Executive Summary Methods: Relevant studies in the area of thinking skills were obtained by systematically searching a number of online databases of educational research literature, by identifying references in reviews and other relevant books and reports, and from contacts with expertise in this

	area. Twenty-six of the studies identified for this review were obtained from the database which resulted from the first thinking skills review (Higgins et al., 2004); a further three resulted from updating the original search and applying the more stringent criteria required for a quantitative synthesis. Studies were selected for the meta-analysis if they had sufficient quantitative data to calculate an effect size (relative to a control or comparison group of pupils) and if the number of research subjects was greater than 10. Effect sizes were calculated from the reported data and combined statistically using quantitative synthesis. Results: twenty-nine studies were identified which contained quantitative data on pupils' attainment and attitudes suitable for meta-analysis. The studies come from a range of countries around the world with half set in the US or UK. The studies broadly cover the ages of compulsory schooling (5–16) and include studies set in both primary and secondary schools. A number of named thinking skills interventions are included, such as Feuerstein's instrumental enrichment (FIE) and cognitive acceleration through science education (CASE) as well as studies which report a more general thinking skills approach (such as the development of metacognitive strategies). The quantitative synthesis indicates that thinking skills programmes and approaches are effective in improving the performance on tests of cognitive measures (such as Raven's progressive matrices) with an overall effect size of 0.62. (This effect would move a class ranked at 50th place in a league table of 100 similar classes to 26th or a percentile gain of 24 points.) However, these approaches also have a considerable impact on curricular outcomes with the same effect size of 0.62. The overall effect size (including cognitive, curricular and affective measures) was 0.74. Conclusions: Overall, the quantitative synthesis indicates that, when thinking skills programmes and approaches are used in schools, they are effective in improving pupils' performance on a range of tested outcomes (relative to those who did not receive thinking skills interventions). The magnitude of the gains found appears to be important when compared with the reported effect sizes of other educational interventions. This review found an overall mean effect of 0.62 for the main (cognitive) effect of each of the included studies, larger than the mean of Hattie's vast database of meta-analyses at 0.4 (Hattie, 1999) but very similar to the overall figure reported by Marzano (1998, p 76) of 0.65 for interventions across the knowledge, cognitive, metacognitive and self-system domains. In particular, our study identified metacognitive interventions as having relatively greater impact, similar to Marzano's study. Looking at a smaller part of our review, Feuerstein's instrumental enrichment is one of the most extensively researched thinking skills programme. Our results broadly concur with those of Romney and Samuels (2001), whose meta-analysis found moderate overall effects and an effect size of 0.43 on reasoning ability (p 28). Our findings were of the same order, with an overall effect size of 0.58 (one main effect from each of seven studies included) and an effect size of 0.52 on tests of reasoning (one main effect from four studies). There is some indication that the impact of thinking skills programmes and approaches may vary according to subject. In our analysis there was relatively greater impact on tests of mathematics (0.89) and science (0.78), compared with reading (0.4).
Klauer & Phye 2008	Researchers have examined inductive reasoning to identify different cognitive processes when participants deal with inductive problems. This article presents a prescriptive theory of inductive reasoning that identifies cognitive processing using a procedural strategy for making comparisons. It is hypothesized that training in the use of the procedural inductive

	reasoning strategy will improve cognitive functioning in terms of (a) increased fluid intelligence performance and (b) better academic learning of classroom subject matter. The review and meta-analysis summarizes the results of 74 training experiments with nearly 3,600 children. Both hypotheses are confirmed. Further, two moderating effects were observed: training effects on intelligence test performance increased over time, and positive problem solving transfer to academic learning is greater than transfer to intelligence test performance. The results cannot be explained by placebo or test-coaching effects. It is concluded that the proposed strategy is theoretically and educationally promising and that children of a broad age range and intellectual capacity benefit with such training.

One to one tuition

Moderate impact for high cost, based on extensive evidence.

+5

One to one tuition is where a teacher, teaching assistant or other adult gives a pupil intensive individual support. It may be undertaken outside of normal lessons as additional teaching, for example as part of extending school time or summer schools, or as a replacement for other lessons by withdrawing the pupil for extra teaching.

How effective is it?

Evidence indicates that one to one tuition can be effective, on average accelerating learning by approximately five additional months' progress.

Short, regular sessions (about 30 minutes, 3-5 times a week) over a set period of time (6-12 weeks) appear to result in optimum impact. Evidence also suggests tuition should be additional to, but explicitly linked with, normal teaching, and that teachers should monitor progress to ensure the tutoring is beneficial. Studies comparing one to one with small group tuition show mixed results. In some cases one to one tuition has led to greater improvement, while in others tuition in groups of two or three has been equally or even more effective compared to one to one. The variability in findings may suggest that the quality of teaching in one to one tuition or small groups is more important than the group size, emphasising the value of professional development for teachers.

Programmes involving teaching assistants or volunteers also have a valuable impact, but tend to be less effective than those using experienced and specifically trained teachers, which have nearly twice the effect on average. Where tuition is delivered by volunteers or teaching assistants there is evidence that training and the use of a structured programme is beneficial.

How secure is the evidence?

Overall, the evidence is consistent and strong, particularly for younger learners who are behind their peers in primary schools, and for subjects like reading and mathematics. There are fewer studies at secondary level or for other subjects.

In the UK, three recent evaluations of one to one tuition interventions (see Catch Up Numeracy, Catch Up Literacy and Switch-on Reading) found average impacts of between three and five months' additional progress, suggesting that positive impacts can be successfully replicated in English schools. In addition, an intensive coaching programme that involved one to one and small group tuition had an average impact of five additional months' progress.

What are the costs?

Overall, costs are estimated as high. A single pupil receiving 30 minutes tuition, five times a week for 12 weeks requires about four full days of a teacher's time, which is estimated to cost approximately £700 per pupil. Costs could be reduced by trialing groups of two or three (see Small group tuition).

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. One to one tuition is very effective in helping learners catch up, but is relatively expensive. Before you commit to one to one tuition, have you considered trialing intensive support groups of two or three and evaluating the impact?
2. Tuition is more likely to make an impact if it is additional to and explicitly linked with normal lessons. Have you considered how you will support pupils and regular class teachers to ensure the impact is sustained once they return to normal classes?
3. Training is likely to be particularly beneficial when tuition is delivered by experienced and well-trained teaching assistants. What training and support have you provided?
4. Have any programmes you are adopting been evaluated?

Technical Appendix: One to one tuition

Definition

One to one tuition is where a teacher, teaching assistant or other adult gives a pupil intensive tuition. This is often as catch up or remedial support for learners who are falling behind their peers with important skills or concepts, though it may also be offered to other learners such as high attainers or in subjects like music for instrumental teaching.

This may be during normal lessons (withdrawal) or it may also be undertaken outside of the pupil's normal lessons, for example as part of after school programmes or summer schools.

Such tuition is usually undertaken by trained teachers or teaching assistants or other adults, such as volunteers, and not by fellow students (see peer tutoring).

It is distinguished from mentoring which is often undertaken by volunteers who focus on building confidence, or developing resilience and character, rather than directly or only focusing on teaching or tutoring specific academic skills.

Evidence rating

There are seven meta-analyses of one-to-one tuition, mainly from well-controlled experiments or trials which were undertaken in schools using pupil attainment data. The causes of variation were explored in these studies and important influences were identified as the experience and training of tutors and the structure and intensity of the tutoring. Overall the evidence is rated as extensive.

Additional cost information

The average salary of a full-time qualified teacher is £34,600 a year (source: https://getintoteaching.education.gov.uk/competitive-salary-and-great-benefits). There are 195 days in the school year. This means that the average cost of four days of teachers' time is £709.74.

References

Full references
Allor, J., & McCathren, R. (2004). The efficacy of an early literacy tutoring program implemented by college students. *Learning Disabilities Research & Practice*, 19(2), 116-129. http://dx.doi.org/10.1111/j.1540-5826.2004.00095.x
Chappell, S., Nunnery, J., Pribesh, S., & Hager, J. (2010). *Supplemental educational services (SES) provision of no child left behind: A synthesis of provider effects* (Research Brief). Norfolk, VA: The Center for Educational Partnerships at Old Dominion University. http://eric.ed.gov/?id=ED530860 http://dx.doi.org/10.3102/00028312019002237
*D'Agostino, J. V., & Murphy, J. A. (2004). A meta-analysis of Reading Recovery in United States schools. *Educational Evaluation and Policy Analysis*, 26(1), 23-28 http://dx.doi.org/10.3102/01623737026001023
*Elbaum, B., Vaughn, S.M., Hughes, M.T. & Moody, S.M. (2000). How Effective Are One-to-One Tutoring Programs in Reading for Elementary Students at Risk for Reading Failure? A Meta-Analysis of the Intervention Research. *Journal of Educational Psychology* 92.4 (2000): 605-619. http://dx.doi.org/10.1037/0022-0663.92.4.605
Gorard, S., See, B.H. & Siddiqui, N. (2014) *Switch-on Reading Evaluation Report and Executive Summary* February 2014. London: EEF. http://educationendowmentfoundation.org.uk/uploads/pdf/FINAL_EEF_Evaluation_Report_-

Switch-on-_February_2014.pdf
*Jun, S.W., Ramirez, G., & Cumming, A. (2010). Tutoring Adolescents in Literacy: A Meta-Analysis. *Journal of Education*, 45 (2), 219-238. http://mje.mcgill.ca/article/view/4770
NfER (2014) *Catch Up® Numeracy: Evaluation Report and Executive Summary* February 2014. London: EEF https://educationendowmentfoundation.org.uk/uploads/pdf/FINAL_EEF_Evaluation_-_Catch_Up_Numeracy_-_February_2014.pdf
NfER (2014) *Catch Up® Literacy: Evaluation Report and Executive Summary* February 2014. London: EEF https://educationendowmentfoundation.org.uk/uploads/pdf/Catch_Up_Literacy_(Final).pdf http://educationendowmentfoundation.org.uk/uploads/pdf/FINAL_EEF_Evaluation_Report_-_Switch-on_-_February_2014.pdf
*Ritter, G.W., Barnett, J.H., Genny, C.S., & Albin, G.R. (2009). The Effectiveness of Volunteer Tutoring Programs for Elementary and Middle School Students: A Meta-Analysis. *Review of Educational Research*, 79 (3), 3-38. http://dx.doi.org/10.3102/0034654308325690
*Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1-26. http://dx.doi.org/10.1016/j.edurev.2010.07.002
Torgerson, C.J., Wiggins, A., Torgerson, D.J., Ainsworth, H., Barmby, P., Hewitt, C., Jones, K., Hendry, V., Askew, M., Bland, M. Coe, R., Higgins, S., Hodgen, J., Hulme, C. & Tymms, P. (2011). *Every Child Counts: The Independent evaluation. Executive Summary.* London: DfE. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/193101/DFE-RBX-10-07.pdf
Washington State Institute for Public Policy (2014a) Tutoring: By adults, one-on-one, structured: Seattle Wa. http://www.wsipp.wa.gov/BenefitCost/ProgramPdf/370/Tutoring-By-adults-one-on-one-structured
Washington State Institute for Public Policy (2014b) Tutoring: By adults, one-on-one, non-structured
Wasik, B. A., & Slavin, R. E. (1993). Preventing Early Reading Failure with One-to-One Tutoring: A Review of Five Programs. *Reading Research Quarterly*, 28(2), 179–200. http://dx.doi.org/10.2307/747888

Summary of effects		
Meta-analyses	**FSM effect size**	***Effect size***
D'Agostino & Murphy 2004	0.32	0.32
Elbaum et al. 2000		0.41
Jun, Ramirez & Cumming 2010 (by adults)		0.70
Ritter et al. 2009		0.30
Slavin et al. 2011 (One-to-one phonics tutoring)		0.62
Washington State Institute 2014(a) structured tutoring		0.53
Washington State Institute 2014(b) non-structured tutoring		0.05
Recent studies		
NfER, 2014 (EEF Catch up Numeracy) (intervention vs control) (time equivalent one to one vs control		0.21 0.27
NFER, 2015 a (EEF Catch up literacy	0.00	0.12
Gorard et al. 2014 (EEF Switch on Reading)		0.24
NFER, 2015b (EEF One to one coaching in literacy)	0.40	0.36
Allor et al., 2004		0.53
Weighted mean effect size		**0.40**

For more information about the effect sizes in the Toolkit, click here.

Study	Abstract
D'Agostino & Murphy 2004	We conducted a meta-analysis of 36 studies of Reading Recovery (RR), an intensive tutorial intervention designed to develop the literacy skills of low-performing first grade students. Few individual studies of the program have yielded conclusive evidence regarding the program's effectiveness due to various methodological limitations. We relied on specific meta-analytic strategies to combine as much available evidence as possible to study overall program effects. We also analysed the results from the few more rigorously designed studies separately. In general, we found positive program effects for both discontinued and not discontinued students on outcomes tailored to the program and standardized achievement measures. RR effects were most pronounced, however, for discontinued students on measures designed for the program. Contrary to conventional belief, we found no evidence suggesting that prior observed effects could be explained completely by factors resulting from methodological flaws (e.g., regression artifacts).
Elbaum et al. 2000	A meta-analysis of supplemental, adult-instructed one-to-one reading interventions for elementary students at risk for reading failure was conducted. Reading outcomes for 42 samples of students (N = 1,539) investigated in 29 studies reported between 1975 and 1998 had a mean weighted effect size of 0.41 when compared with controls. Interventions that used trained volunteers or college students were highly effective. For Reading Recovery interventions, effects for students identified as discontinued were substantial, whereas effects for students identified as not discontinued were not significantly different from zero. Two studies comparing one-to-one with small-group supplemental instruction showed no advantage for the one-to-one programs.
Jun et al. 2010	What does research reveal about tutoring adolescents in literacy? We conducted a meta-analysis, identifying 152 published studies, of which 12 met rigorous inclusion criteria. We analysed the 12 studies for the effects of tutoring according to the type, focus, and amount of tutoring; the number, age, and language background of students; and the quality of the research. Despite variability, these studies suggest benefits, notably for cross-age tutoring, reading, and small tutoring programs of lengthy duration.
Ritter et al. 2009	This meta-analysis assesses the effectiveness of volunteer tutoring programs for improving the academic skills of students enrolled in public schools Grades K–8 in the United States and further investigates for whom and under what conditions tutoring can be effective. The authors found 21 studies (with 28 different study cohorts in those studies) reporting on randomized field trials to guide them in assessing the effectiveness of volunteer tutoring programs. Overall, the authors found volunteer tutoring has a positive effect on student achievement. With respect to particular sub-skills, students who work with volunteer tutors are likely to earn higher scores on assessments related to letters and words, oral fluency, and writing as compared to their peers who are not tutored.
Slavin et al. 2011	This article reviews research on the achievement outcomes of alternative approaches for struggling readers ages 5–10 (US grades K-5): One-to-one tutoring, small-group tutorials, classroom instructional process approaches, and computer-assisted instruction. Study inclusion criteria included use of randomized or well-matched control groups, study duration of at least 12 weeks, and use of valid measures independent of treatments. A total of 97 studies met these criteria. The review concludes that one-to-one tutoring is very effective in improving reading performance. Tutoring models that focus on phonics obtain much better outcomes than others. Teachers are more effective than paraprofessionals and volunteers as tutors. Small-group,

	phonetic tutorials can be effective, but are not as effective as one-to-one phonetically focused tutoring. Classroom instructional process programs, especially cooperative learning, can have very positive effects for struggling readers. Computer-assisted instruction had few effects on reading. Taken together, the findings support a strong focus on improving classroom instruction and then providing one-to-one, phonetic tutoring to students who continue to experience difficulties.
Washington State Institute 2014 a	The tutoring programs included in this meta-analysis are structured, systematic approaches to tutoring struggling students in specific English language arts and/or mathematics skills. The evaluated programs include a variety of specific programs and curricula such as (in no particular order) Reading Recovery, Mathematics Recovery, Edmark Reading Program, Howard Street Tutoring, and Early Intervention Program. The programs provide, on average, about 30 hours of tutoring time to an individual student each year. Tutors are typically certificated teachers or specially trained adults (e.g. instructional aides and community volunteers). Tutors receive approximately ten hours of training per year with a focus on the specific content and general tutoring strategies.
Washington State Institute 2014 b	The tutoring programs included in this analysis provide one-on-one assistance to struggling students in English language arts and/or mathematics. The evaluated programs typically allow tutors to exercise their own discretion when selecting and implementing tutoring strategies. The programs provide, on average, about 30 hours of tutoring time to an individual student each year. The tutors are non-certificated adults (e.g. instructional aides and community volunteers) who receive approximately two hours of training per year.

Oral language interventions

Moderate impact for very low cost, based on extensive evidence.

+5

Oral language interventions emphasise the importance of spoken language and verbal interaction in the classroom.

They are based on the idea that comprehension and reading skills benefit from explicit discussion of either the content or processes of learning, or both. Oral language approaches include:

- Targeted reading aloud and discussing books with young children
- Explicitly extending pupils' spoken vocabulary
- The use of structured questioning to develop reading comprehension

All of the approaches reviewed in this section support learners' articulation of ideas and spoken expression, such as Thinking Together or Philosophy for Children. Oral language interventions therefore have some similarity to approaches based on Meta-Cognition, which make talk about learning explicit in classrooms, and to Collaborative Learning approaches, which promote pupils' talk and interaction in groups.

How effective is it?

Overall, studies of oral language interventions consistently show positive benefits on learning, including oral language skills and reading comprehension. On average, pupils who participate in oral language interventions make approximately five months' additional progress over the course of a year.

All pupils appear to benefit from oral language interventions, but some studies show slightly larger effects for younger children and pupils from disadvantaged backgrounds (up to six months' benefit). Likewise, some types of oral language interventions appear, on average, to be more effective than others.

There is consistent evidence supporting reading to young children and encouraging them to answer questions and to talk about the story with a trained adult. Conversely, 'whole language' approaches, which focus on meaning and personal understanding, do not appear to be as successful as those involving more interactive and dialogic activities. A number of studies show the benefits of trained teaching assistants effectively supporting both oral language skills and reading outcomes.

For all oral language interventions certain factors are associated with higher learning gains, suggesting that careful implementation is important. Approaches which explicitly aim to develop spoken vocabulary work best when they are related to current content being studied in school and when they involve active use of any new vocabulary. Likewise, approaches that use technology are most effective when technology is used as a medium to encourage collaborative work and interaction between pupils, rather than a taking a direct teaching or tutoring role. Most studies comment on the importance of training and teacher development or support with implementation.

How secure is the evidence?

There is an extensive evidence base on the impact of oral language interventions, including a substantial number of meta-analyses and systematic reviews. The evidence is relatively consistent, suggesting that oral language interventions can be successful in a variety of environments. Although the majority of the evidence relates to younger children, there is also clear evidence that older learners, and particularly disadvantaged pupils, can benefit.

The evidence base includes a number of high quality studies in UK schools. Additional evidence about matching specific programmes or approaches to particular learners' needs either by age or by attainment would also be useful.

What are the costs?

Overall, the costs are estimated as under £80 per pupil and very low. There are few, if any, direct financial costs associated with this approach. Additional resources such as books for discussion may be required. In a recent UK evaluation the cost of these additional resources was estimated at between £10 and £20 per pupil. Professional development or training is also likely to enhance the benefits on learning.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. How can you help pupils to make their learning explicit through verbal expression?
2. How will you match the oral language activities to learners' current stage of development so that it extends their learning and connects with the curriculum?
3. What training will the adults involved have to ensure they model and develop pupil's oral language skills?
4. If you are using technology, how will you ensure that pupils talk about their learning and interact with each other effectively?

Oral languages interventions
Toolkit references

Full references
Alexander, R. (2012). Improving oracy and classroom talk in English schools: Achievements and challenges. In DfE Seminar on Oracy, the National Curriculum and Educational Standards (Vol. 20). http://www.robinalexander.org.uk/wp-content/uploads/2012/06/DfE-oracy-120220-Alexander-FINAL.pdf
*Blok, H. (1999). Reading to Young Children in Educational Settings: A Meta-Analysis of Recent Research. Language Learning, 49(2), 343-371. DOI: 10.1111/0023-8333.0009
*Bus, A. G., Van Ijzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. Review of educational research, 65(1), 1-21. doi: 10.3102/00346543065001001
*Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. Journal of Research on Educational Effectiveness, 2(1), 1-44. DOI:10.1080/19345740802539200
*Guthrie, J. T., McRae, A., & Klauda, S. L. (2007). Contributions of concept-oriented reading instruction to knowledge about interventions for motivations in reading. Educational Psychologist, 42(4), 237-250. DOI:10.1080/00461520701621087
Jalongo, M. R. (2010). Listening in early childhood: An interdisciplinary review of the literature. The International Journal of Listening, 24(1), 1-18.
*Marulis, L. M., & Neuman, S. B. (2010). The Effects of Vocabulary Intervention on Young Children's Word Learning A Meta-Analysis. Review of Educational Research, 80(3), 300-335. doi: 10.3102/0034654310377087
*Mol, S.E.., Bus. A.G. & Maria T. de Jong (2009) Interactive Book Reading in Early Education: A Tool to Stimulate Print Knowledge as Well as Oral Language Review of Educational Research June 2009, Vol. 79, No. 2, pp. 979-1007 DOI: 10.3102/0034654309332561
*Stahl, S. A., & Miller, P. D. (1989). Whole language and language experience approaches for beginning reading: A quantitative research synthesis. Review of Educational Research, 59(1), 87-116. doi: 10.3102/00346543059001087
*Strong, G. K., Torgerson, C. J., Torgerson, D., & Hulme, C. (2011). A systematic meta-analytic review of evidence for the effectiveness of the 'Fast ForWord'language intervention program. Journal of Child Psychology and Psychiatry, 52(3), 224-235. DOI: 10.1111/j.1469-7610.2010.02329.x
Styles, B. Clarkson, & R. Fowler, K (2014) *Chatterbooks: Evaluation Report and Executive Summary* London: Education Endowment Foundation
*Swanson, E., Vaughn, S., Wanzek, J., Petscher, Y., Heckert, J., Cavanaugh, C., ... & Tackett, K. (2011). A synthesis of read-aloud interventions on early reading outcomes among preschool through third graders at risk for reading difficulties. Journal of learning disabilities, 44(3), 258-275. doi: 10.1177/0022219410378444
* included meta-analyses

Oral languages interventions
Toolkit references

Summary of effects	
Study	***Effect size***
Blok, 1999	0.41
Bus, Van Ijzendoorn & Pellegrini, 1995	0.55
Elleman, Lindo, Morphy & Compton, 2009	0.10
Guthrie, McRae & Klauda, 2007	0.91
Marulis & Neuman, 2010	0.88
Mol, Bus & de Jong, 2009	0.39
Stahl & Miller, 1989	0.09
Strong, Torgerson, Torgerson & Hulme, 2011	0.08
Swanson et al. 2011	0.29
Recent studies	
Styles, Clarkson & Fowler, 2014 (Chatterbooks)	-0.14
Indicative effect size	**0.41**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Blok, 1999	This article reviews 10 studies, comprising 11 samples, of the effects of reading to young children in schools. The age of the children varied between 31 and 90 months. Dependent variables were classified in 2 domains: oral language and reading skills. The combined effect size for the oral language domain was d = .63, and for the reading domain d = .41. Although these figures look promising, caution is needed because the empirical evidence appears to be meager. Not only is the number of studies small, but a critical analysis of the design of the studies generally reveals poor quality. In 8 of the 11 samples, students came from disadvantaged families. The mean age is 5 years (average 65 months, range 31 through 90 months).
Bus, Van Ijzendoorn & Pellegrini, 1995	The current review is a quantitative meta-analysis of the available empirical evidence related to parent-pre-schooler reading and several outcome measures. In selecting the studies to be included in this meta-analysis, we focused on studies examining the frequency of book reading to pre-schoolers. The results support the hypothesis that parent-pre-schooler reading is related to outcome measures such as language growth, emergent literacy, and reading achievement. The overall effect size of d = .59 indicates that book reading explains about 8% of the variance in the outcome measures. The results support the hypothesis that book reading, in particular, affects acquisition of the written language register. The effect of parent-pre-schooler reading is not dependent on the socioeconomic status of the families or on several methodological differences between the studies. However, the effect seems to become smaller as soon as children become conventional readers and are able to read on their own.
Marulis & Neuman, 2010	This meta-analysis examines the effects of vocabulary interventions on pre-K and kindergarten children's oral language development. The authors quantitatively reviewed 67 studies and 216 effect sizes to better understand the impact of training on word learning. Results indicated an overall effect size of .88, demonstrating, on average, a gain of nearly one standard deviation on vocabulary measures. Moderator analyses reported greater effects for trained adults in providing the treatment, combined pedagogical strategies that included explicit and implicit instruction, and author-created measures compared to standardized measures. Middle- and upper-income at-risk children were significantly more likely to benefit from vocabulary intervention than those students also at risk and poor. These results indicate that although they might improve oral language skills, vocabulary interventions are not sufficiently powerful to close the gap—even in

	the preschool and kindergarten years.
Mol, Bus & de Jong, 2009	This meta-analysis examines to what extent interactive storybook reading stimulates two pillars of learning to read: vocabulary and print knowledge. The authors quantitatively reviewed 31 (quasi) experiments (n = 2,049 children) in which educators were trained to encourage children to be actively involved before, during, and after joint book reading. A moderate effect size was found for oral language skills, implying that both quality of book reading in classrooms and frequency are important. Although teaching print-related skills is not part of interactive reading programs, 7% of the variance in kindergarten children s alphabetic knowledge could be attributed to the intervention. The study also shows that findings with experimenters were simply not replicable in a natural classroom setting. Further research is needed to disentangle the processes that explain the effects of interactive reading on children sprint knowledge and the strategies that may help transfer intervention effects from researchers to children s own teachers.
Stahl & Miller, 1989	To examine the effects of whole language and language experience approaches on beginning reading achievement, a quantitative synthesis was performed on two data bases: the five projects conducted as part of the United States Office of Education (USOE) first grade studies and 46 additional studies comparing basal reading approaches to whole language or language experience approaches. The results of both analyses suggest that, overall, whole language/language experience approaches and basal reader approaches are approximately equal in their effects, with several exceptions. First, whole language/language experience approaches may be more effective in kindergarten than in first grade. Second, they may produce stronger effects on measures of word recognition than on measures of reading comprehension. Third, more recent studies show a trend toward stronger effects for the basal reading program relative to whole language/language experience methods. Fourth, whole language/language experience approaches produce weaker effects with populations labelled specifically as disadvantaged than they do with those not specifically labelled. Finally, studies with higher rated quality tend to produce lower effect sizes and the lowest effect sizes were found in studies that evaluated existing programs, as opposed to newly implemented experimental programs. These results are discussed within a stage model of reading that suggests that whole language/language experience approaches might be most effective for teaching functional aspects of reading, such as print concepts and expectations about reading, whereas more direct approaches might be better at helping students master word recognition skills prerequisite to effective comprehension.
Strong, Torgerson, Torgerson & Hulme, 2011	Background: Fast ForWord is a suite of computer-based language intervention programs designed to improve children's reading and oral language skills. The programs are based on the hypothesis that oral language difficulties often arise from a rapid auditory temporal processing deficit that compromises the development of phonological representations. Methods: A systematic review was designed, undertaken and reported using items from the PRISMA statement. A literature search was conducted using the terms 'Fast ForWord''Fast For Word''Fastforword' with no restriction on dates of publication. Following screening of (a) titles and abstracts and (b) full papers, using pre-established inclusion and exclusion criteria, six papers were identified as meeting the criteria for inclusion (randomised controlled trial (RCT) or matched group comparison studies with baseline equivalence published in refereed journals). Data extraction and analyses were carried out on reading and language outcome measures comparing the Fast ForWord intervention groups to both active and untreated control groups. Results: Meta-analyses indicated that there was no significant effect of Fast ForWord on any outcome measure in comparison to active or untreated control groups.

	Conclusions: There is no evidence from the analysis carried out that Fast ForWord is effective as a treatment for children's oral language or reading difficulties.
Swanson et al. 2011	A synthesis and meta-analysis of the extant research on the effects of storybook read-aloud interventions for children at risk for reading difficulties ages 3 to 8 is provided. A total of 29 studies met criteria for the synthesis, with 18 studies providing sufficient data for inclusion in the meta-analysis. Read-aloud instruction has been examined using dialogic reading; repeated reading of stories; story reading with limited questioning before, during, and/or after reading; computer-assisted story reading; and story reading with extended vocabulary activities. Significant, positive effects on children's language, phonological awareness, print concepts, comprehension, and vocabulary outcomes were found. Despite the positive effects for read-aloud interventions, only a small amount of outcome variance was accounted for by intervention type.
Guthrie, McRae & Klauda, 2007	We present a theoretical and empirical explication of the intervention of Concept-Oriented Reading Instruction (CORI) that is designed to increase students' reading comprehension and motivation for reading. The framework specifies a set of five motivational constructs that represent goals for the instructional intervention. Necessary cognitive goals in reading are also presented. For this intervention, the five instructional practices of relevance, choice, success, collaboration, and thematic unit that are prominent in CORI are portrayed as components that are aligned with motivational constructs. The impact of CORI on some of the motivational processes, cognitive competencies, and reading comprehension are presented in the form of a meta-analysis of 11 CORI studies with 75 effect sizes on 20 outcome variables. The CORI motivational intervention is compared to laboratory treatments and other field studies.
Elleman, Lindo, Morphy & Compton, 2009	A meta-analysis of vocabulary interventions in grades pre-K to 12 was conducted with 37 studies to better understand the impact of vocabulary on comprehension. Vocabulary instruction was found to be effective at increasing students' ability to comprehend text with custom measures (d = 0.50), but was less effective for standardized measures (d = 0.10). When considering only custom measures, and controlling for method variables, students with reading difficulties (d = 1.23) benefited more than three times as much as students without reading problems (d = 0.39) on comprehension measures. Gains on vocabulary measures, however, were comparable across reading ability. In addition, the correlation of vocabulary and comprehension effects from studies reporting both outcomes was modest (r = .43).

Outdoor adventure learning

Moderate impact for moderate cost, based on limited evidence.

+3

Outdoor adventure learning typically involves outdoor experiences, such as climbing or mountaineering, survival, ropes or assault courses, or outdoor sports, such as orienteering, sailing and canoeing. These can be organised as intensive block experiences or shorter courses run in schools or local outdoor centres.

Adventure education usually involves collaborative learning experiences with a high level of physical (and often emotional) challenge. Practical problem-solving, explicit reflection and discussion of thinking (see also Meta-cognition and self-regulation) may also be involved.

Adventure learning interventions typically do not include a formal academic component. This summary does therefore not include approaches to outdoor learning, such as Forest Schools or field trips.

How effective is it?

Overall, studies of adventure learning interventions consistently show positive benefits on academic learning, and wider outcomes such as self-confidence. On average, pupils who participate in adventure learning interventions appear to make approximately three additional months' progress.

The evidence suggests that the impact is greater for longer courses (more than a week), and those in a 'wilderness' setting, though other types of intervention still show some positive impacts.

Understanding why adventure learning interventions appear to improve academic outcomes is not straightforward. One assumption might be that non-cognitive skills such as perseverance and resilience are developed through adventure learning and that these skills have a knock-on impact on academic outcomes. However, it should be noted that the wider evidence base on the relationship between these types of non-cognitive skills is underdeveloped.

If adventure learning interventions are effective because of their impact on non-cognitive skills, then explicitly encouraging students to actively apply these skills in the classroom is likely to increase effectiveness.

How secure is the evidence?

The existing base on adventure learning interventions is limited and relatively inconsistent. The most recent studies, which use more robust methodologies, show smaller effects than older studies. Our overall assessment of potential progress is weighted towards more recent studies. However, on average both older and more recent studies do show a positive impact on academic attainment.

The existing qualitative evidence is more consistent than the quantitative findings, showing that in most cases young people perceive adventure learning interventions to have had a positive impact on their lives and attitudes.

What are the costs?

Costs vary with a 10 day adventure sailing experience costing about £900 and an 8 day Outward Bound course about £500. An adventure ropes course costs about £30 for a day. Costs are estimated at £500 per pupil per year and are therefore moderate.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. A wide range of adventure activities are linked with increased academic achievement.
2. Experiences of over a week tend to have greater impact and tend to produce effects of a longer duration.
3. It is important to work with well-trained and well-qualified staff as adventure experiences can pose very different physical and emotional risks to those in schools.
4. Effects are evident in self-confidence, self-efficacy and motivation. Have you made all teachers aware of the intervention and how improvements in these characteristics may be supported in the classroom?
5. How you will ensure the benefits of outdoor adventure learning are transferred into the classroom?

Outdoor adventure learning Toolkit references

Full references
Cason, D., & Gillis, H. L. (1994). A meta-analysis of outdoor adventure programming with adolescents. Journal of Experiential Education, 17(1), 40-47.
Gillis, L. H., & Speelman, E. (2008). Are challenge (ropes) courses an effective tool? A meta-analysis. Journal of Experiential Education, 31(2), 111-135.
Hans, T. A. (2000). A meta-analysis of the effects of adventure programming on locus of control. Journal of Contemporary Psychotherapy, 30(1), 33-60.
Hattie, J., Marsh, H. W., Neill, J. T., & Richards, G. E. (1997). Adventure education and Outward Bound: Out-of-class experiences that make a lasting difference. Review of Educational Research, 67(1), 43-87.
Laidlaw, J. S. (2000). A meta-analysis of outdoor education programs. University of Northern Colorado. ProQuest Dissertations and Theses: 304612041, 96 p. Retrieved from http://search.proquest.com/docview/304612041.
McKenzie, M. D. (2000). How are adventure education program outcomes achieved?: A review of the literature. Australian Journal of Outdoor Education-Vol, 5(1), 19-28.
Wilson, S. J., & Lipsey, M. W. (2000). Wilderness challenge programs for delinquent youth: A meta-analysis of outcome evaluations. Evaluation and Program Planning, 23(1), 1-12.

Summary of effects	
Study	***Effect size***
Cason & Gillis, 1994	0.31 (all effects) 0.61 (on school grades)
Gillis & Speelman, 2008	0.43 (overall) 0.26 (academic outcomes)
Hattie et. al, 1997	0.34 0.45 (academic outcomes)
Hattie et. al, 1997	0.34
Laidlaw, 2000	0.17
Indicative effect size	**0.23**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	**Abstract**
Cason & Gillis, (1994).	Adventure practitioners asked to justify their work with adolescent populations have no one study to point to that statistically sums up major findings in the field. Whether it be a school board, treatment facility, or funding agency, one study is needed which can combine statistics from many studies into a format to show overall effectiveness of adventure programming? This study used the statistical technique of meta-analysis to demonstrate that adolescents who attend adventure programming are 62% better off than those who do not. While combining various populations and outcomes resulted in an overall effect that could be considered small by some accounts, the study did point to major problems with current research and offers some direction for future researchers to explore.
Gillis & Speelman, (2008).	This study reports the results of a meta-analysis of 44 studies that examined the impacts of participation in challenge (ropes) course activities. Overall, a medium standardized mean difference effect size was found (d = 0.43). Effect sizes were calculated for various study characteristics, including demographics and outcome. Higher effects were found for adult groups (d = 0.80) and for studies measuring family functioning (d = 0.67). Studies with therapeutic (d = 0.53) or developmental foci (d = 0.47) had higher effect sizes than those with educational foci (d = 0.17). Higher effect sizes for group effectiveness (d = 0.62) affirmed the use of challenge course experiences for team-building purposes. Implications for further research include the importance of recording detailed program design information, selecting

	appropriate instrumentation, and including follow-up data.
Hattie et.al. (1997).	The purpose of this meta-analysis is to examine the effects of adventure programs on a diverse array of outcomes such as self-concept, locus of control, and leadership. The meta-analysis was based on 1,728 effect sizes drawn from 151 unique samples from 96 studies, and the average effect size at the end of the programs was .34. In a remarkable contrast to most educational research, these short-term or immediate gains were followed by substantial additional gains between the end of the program and follow-up assessments (ES = .17). The effect sizes varied substantially according the particular program and outcome and improved as the length of the program and the ages of participants increased. Too little is known, however, about why adventure programs work most effectively.
Laidlaw, (2000).	The purpose of this meta-analysis was to examine research in the field of outdoor education to determine if features of studies, outcomes, and programs are significantly related to variation among the estimated effects of outdoor education programs. The primary findings of this dissertation were that study design and the degree to which outcomes were proximal to the intent of the program explained a significant part of the variance in effect estimates. Specifically, studies using poorly controlled designs had the highest mean effect size estimates (effect size = .6), in contrast to those that used controlled, experimental designs (effect size = .17). In this aspect, the findings of this study support the results of Cason and Gillis. In addition, the findings of this meta-analysis indicated that studies which evaluated outcomes proximally related to program goals had significantly higher effect sizes (effect size = .77) than those studies which evaluated distally related outcomes (effect size = .40). In a notable contrast to both prior meta-analyses in the field, after controlling for the influence of potentially confounding variables, and after controlling for a problematic issue of meta-analysis, that of independence of effect sizes, no other feature of outcomes or programs were significantly related to effect sizes. The results of this dissertation imply that the relationship between outcomes and program goals are important considerations, and that relationship between other substantive features of programs (such as length) and their subsequent outcomes (self-concept) cannot be determined from the existing literature given its inherent problem.

Parental involvement

Moderate impact for moderate cost, based on moderate evidence.

+3

Parental Involvement covers the active engagement of parents in supporting their children's learning at school. This includes programmes focused on parents and their skills (such as improving literacy or IT skills), general approaches to encourage parents to support their children to read or do mathematics, and more intensive programmes for families in crisis.

How effective is it?

Although parental involvement is consistently associated with pupils' success at school, the evidence about how to increase involvement to improve attainment is mixed and much less conclusive. This is particularly the case for disadvantaged families. There is some evidence that supporting parents with their first child will have benefits for siblings. However, there are also examples where combining parental engagement strategies with other interventions, such as extended early years provision, has not been associated with any additional educational benefit. This suggests that developing effective parental involvement to improve their children's attainment is challenging and will need careful monitoring and evaluation.

The impact of parents' aspirations is also important, though there is insufficient evidence to show that changing parents' aspirations will raise their children's aspirations and achievement over the longer term. Two recent meta-analyses from the USA suggested that increasing parental involvement in primary and secondary schools had on average 2-3 months positive impact.

How secure is the evidence?

Although there is a long history of research into parental involvement programmes, there is surprisingly little robust evidence of the impact of programmes that have tried to increase involvement to improve learning. The association between parental involvement and a child's academic success is well established, but rigorous evaluation of approaches to improve learning through parental involvement is more sparse.

The evidence is predominantly from primary level and the early years, though there are studies which have looked at secondary schools. Impact studies tend to focus on reading and mathematics attainment.

What are the costs?

The costs of different approaches vary enormously, from running parent workshops (about £80 per session) and improving communications, which are cheap, to intensive family support programmes with specially trained staff. The cost of a specialist community or home/school liaison teacher is about £35,000, or about 37 secondary-level Pupil Premium allocations. Costs per pupil are therefore estimated as moderate.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Involvement is often easier to achieve with parents of very young children.
2. What approaches will you take to support parents in working with their children?
3. Have you provided a flexible approach to allow parental involvement to fit around their schedule?
4. Parents of older children may appreciate short sessions at flexible times to involve them.
5. How will you make your school welcoming for parents whose own experience of school may not have been positive?
6. Have you provided some simple, practical ways that parents can support their children in ways that do not require a high level of ability (e.g. by ensuring that students have an environment where they can work at home)?

Parental involvement Toolkit references

Full references
Bus, A. G., Van Ijzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. Review of Educational Research, 65(1), 1-21.
Cummings, C., Laing, K., Law, J., McLaughlin, J., Papps, I., Todd, L. & Woolner, P. (2012) Can Changing Aspirations And Attitudes Impact On Educational Attainment? A Review Of Interventions York: Joseph Rowntree Foundation.
Gorard, S., See B.H. & Davies, P. (2012) The impact of attitudes and aspirations on educational attainment and participation York: Joseph Rowntree Foundation.
Jeynes, W.H. (2005). A Meta-Analysis of the Relation of Parental Involvement to Urban Elementary School Student Academic Achievement. Urban Education, 40.3 pp 237-269.
Jeynes, W.H. (2007). The Relationship Between Parental Involvement and Urban Secondary School Student Academic Achievement: A Meta-Analysis. Urban Education, 42.1 pp 82-110.
Layzer, J. I., Goodson, B. D., Bernstein, L., & Price, C. (2001). National Evaluation of Family Support Programs. Final Report Volume A: The Meta-Analysis.
Mattingly, D.J., Prislin, R., McKenzie, T.L., Rodriguez, J.L., Kayzar, B. (2002). Evaluating Evaluations: The Case of Parent Involvement Programs. Review of Educational Research, 72.4 pp 549-576.
Nye, C., Schwartz, J., & Turner, H. (2006). Approaches to Parent Involvement for Improving the Academic Performance of Elementary SchoolAge Children: A Systematic Review. Campbell Systematic Reviews, 2(4).
Pomerantz, E.M. & Moorman, E.A., Litwack, S.D. (2007). The How, Whom, and Why of Parents' Involvement in Children's Academic Lives: More Is Not Always Better. Review of Educational Research 77. 3 pp. 373–410.
Seitz V. & Apfel, N.H. (1994). Parent-Focused Intervention: Diffusion Effects on Siblings, Child Development 65.2 pp 677-683.
Van-Steensel, R., McElvany, N. Kurvers J. & Herppich S. (2011). How Effective Are Family Literacy Programs? : Results of a Meta-Analysis. Review Of Educational Research 81.1 pp 69-96.

Summary of effects	
Study	***Effect size***
Bus et al. 1995 (joint book reading)	0.59
Jeynes, 2005	0.27
Jeynes, 2007	0.25
Layzer et al, (2001)	0.29
Nye et al (2006)	0.43
Van-Steensel et.al., 2011 (family literacy)	0.18
Indicative effect size	**0.26**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Bus et.al. (1995).	The current review is a quantitative meta-analysis of the available empirical evidence related to parent-pre-schooler reading and several outcome measures. In selecting the studies to be included in this meta-analysis, we focused on studies examining the frequency of book reading to pre-schoolers. The results support the hypothesis that parent-pre-schooler

	reading is related to outcome measures such as language growth, emergent literacy, and reading achievement. The overall effect size of d= .59 indicates that book reading explains about 8% of the variance in the outcome measures. The results support the hypothesis that book reading, in particular, affects acquisition of the written language register. The effect of parent-pre-schooler reading is not dependent on the socioeconomic status of the families or on several methodological differences between the studies. However, the effect seems to become smaller as soon as children become conventional readers and are able to read on their own.
Jeynes, (2005).	This meta-analysis of 41 studies examines the relationship between parental involvement and the academic achievement of urban elementary school children. Analyses determined the effect sizes for parental involvement overall and subcategories of involvement. Results indicate a significant relationship between parental involvement overall and academic achievement. Parental involvement, as a whole, was associated with all the academic variables by about 0.7 to 0.75 of a standard deviation unit. This relationship held for White and minority children and also for boys and girls. The significance of these results is discussed.
Jeynes, (2007).	A meta-analysis is undertaken, including 52 studies, to determine the influence of parental involvement on the educational outcomes of urban secondary school children. Statistical analyses are done to determine the overall impact of parental involvement as well as specific components of parental involvement. Four different measures of educational outcomes are used. These measures include an overall measure of all components of academic achievement combined, grades, standardized tests, and other measures that generally included teacher rating scales and indices of academic attitudes and behaviours. The possible differing effects of parental involvement by race and socioeconomic status are also examined. The results indicate that the influence of parental involvement overall is significant for secondary school children. Parental involvement as a whole affects all the academic variables under study by about .5 to .55 of a standard deviation unit. The positive effects of parental involvement hold for both White and minority children.
Layzer et.al. (2001).	This volume is part of the final report of the National Evaluation of Family Support Programs and details findings from a meta-analysis of extant research on programs providing family support services. Chapter A1 of this volume

	provides a rationale for using meta-analysis. Chapter A2 describes the steps of preparation for the meta-analysis. Chapter A3 describes the 260 programs or interventions represented in the meta-analysis examines their representativeness by comparing them with 167 family support programs that were not evaluated, describes characteristics of the studies included in the analysis, and compares them with excluded studies. Chapter A4 describes the analytic approach to answering the central research questions regarding the impact of family support services on selected child and adult outcomes and the program or treatment characteristics related to impacts. Chapter A5 details the findings of the meta-analysis. The analysis revealed that programs providing family support services had small but statistically significant average short-term effects on child cognitive development and school performance, child social and emotional development, child health, child safety, parent attitudes and knowledge, parenting behaviour, family functioning, parental mental health and health risk behaviours, and economic well-being. Associated with stronger child outcomes were programs that targeted special needs children. Associated with less strong child outcomes were programs that used home visiting as their primary method of working with parents. Programs with the largest parent effects focused on developing parents' skills as effective adults.
Nye et al (2006)	The impact of parental involvement in a child's growth and development is generally accepted (Sheldon, 2003). However, educators, parent groups, and policy makers continue to debate the issue of whether or not parental involvement has a beneficial effect on the academic achievement of children (Epstein, 2001). A key element in these debates is how parental involvement is defined (Epstein, Sanders, Simon, Salinas, Jansorn, & Voorhis, 2002). Therefore, it is important to understand how parent involvement is defined before conclusions are drawn on the impact of parental involvement. In addition, it is important to understand what aspects of parent involvement have the greatest impact and whether the impact is consistent across children characteristics such as socioeconomic status, race, and the child's grade level, age and gender. During the past several decades, the have been numerous primary studies investigating various aspects of parent involvement and the effect in has on children's learning. The No Child Left Behind Act (NCLB) has served to focus attention on the need and importance of parent

	involvement in their child's education. This systematic review synthesizes findings from this research. For this review, parent involvement is defined as the active engagement of a parent with their child outside of the school day in an activity which centers on enhancing academic performance. This review is intended to provide evidence to policymakers about the level of investment in parental involvement, to educators that can guide the development of parent involvement programs for their school improvement plans, and to researchers in designing studies to rigorously investigate the effectiveness of parent involvement for improving elementary school children's academic performance in schools.
Van-Steensel et.al. (2011).	This meta-analysis examines the effects of family literacy programs on children's literacy development. It analyses the results of 30 recent effect studies (1990–2010); covering 47 samples, and distinguishes between effects in two domains: comprehension-related skills and code-related skills. A small but significant mean effect emerged (d = 0.18). There was only a minor difference between comprehension- and code-related effect measures (d = 0.22 vs. d =0.17). Moderator analyses revealed no statistically significant effects of the program, sample, and study characteristics inferred from the reviewed publications. The results highlight the need for further research into how programs are carried out by parents and children, how program activities are incorporated into existing family literacy practices, and how program contents are transferred to parents.
Senechal & Young, (2008).	This review focuses on intervention studies that tested whether parent–child reading activities would enhance children's reading acquisition. The combined results for the 16 intervention studies, representing 1,340 families, were clear: Parent involvement has a positive effect on children's reading acquisition. Further analyses revealed that interventions in which parents tutored their children using specific literacy activities produced larger effects than those in which parents listened to their children read books. The three studies in which parents read to their children did not result in significant reading gains. When deciding which type of intervention to implement, educators will have to weigh a variety of factors such as the differences in effectiveness across the different types of intervention, the amount of resources needed to implement the interventions, and the reading level of the children.

Parental involvement
Toolkit references

Education
Endowment
Foundation

Peer tutoring

Moderate impact for very low cost, based on extensive evidence.

+5

Peer tutoring includes a range of approaches in which learners work in pairs or small groups to provide each other with explicit teaching support. In cross-age tutoring, an older learner takes the tutoring role and is paired with a younger tutee or tutees. Peer assisted learning is a structured approach for mathematics and reading with sessions of 25-35 minutes two or three times a week. In reciprocal peer tutoring, learners alternate between the role of tutor and tutee. The common characteristic is that learners take on responsibility for aspects of teaching and for evaluating their success. Peer assessment involves the peer tutor providing feedback to children relating to their performance and can have different forms such as reinforcing or correcting aspects of learning.

How effective is it?

Overall, the introduction of peer tutoring approaches appears to have a positive impact on learning, with an average positive effect of approximately five additional months' progress. Studies have identified benefits for both tutors and tutees, and for a wide range of age groups. Though all types of pupils appear to benefit from peer tutoring, there is some evidence that children from disadvantaged backgrounds and low attaining pupils make the biggest gains.

Peer tutoring appears to be particularly effective when pupils are provided with support to ensure that the quality of peer interaction is high, for example by providing questioning frames. In cross-age peer tutoring some studies have found that a two year age gap is effective and that intensive blocks of tutoring are more effective, relative to longer programmes.

Peer tutoring appears to be less effective when the approach replaces normal teaching, rather than supplementing or enhancing it, suggesting that peer tutoring is most effectively used to consolidate learning, rather than to introduce new material.

How secure is the evidence?

Peer tutoring has been extensively studied and a majority of studies show moderate to high average effects. High-quality reviews have explored the impact of peer tutoring at both primary and secondary level, and in a variety of subjects.

Though overall the evidence base related to peer tutoring is relatively consistent, some recent studies of peer tutoring have found lower average effects, suggesting that monitoring the implementation and impact of peer tutoring is valuable.

Two randomised controlled trials conducted in English schools and published in 2015 found that the introduction of new peer tutoring programmes did not lead to any improvement in attainment. These findings may reinforce the finding that factors such as the amount or type of support provided to tutors are important, or indicate that some forms of peer tutoring that have been effective elsewhere are less effective in English schools. It is also possible that the introduction of new peer tutoring programmes will have less of an impact in schools where peer tutoring or collaborative learning is already commonplace. However, it would be valuable to assess this claim through further research.

What are the costs?

The direct costs of running peer tutoring in schools are very low, as few additional materials are required (£10-20 per pupil per year). Professional development and additional support for staff is recommended, particularly in the early stages of setting up a programme. Estimates are less than £80 per pupil, indicating very low overall costs.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Are the activities sufficiently challenging for the tutee to benefit from the tutor's support?
2. What support will the tutor receive to ensure that the quality of peer interaction is high?
3. Training for staff and tutors are essential ingredients for success. How will you organise sufficient time to train both staff and tutors, and to identify improvements as the programme progresses?
4. Is peer tutoring being used to review or consolidate learning, or to introduce new material?
5. Four to ten week intensive blocks appear to provide maximum impact for both tutors and tutees. Can you arrange for your peer tutoring to follow this structure?

Technical Appendix: Peer tutoring

Definition

Peer tutoring includes a range of approaches in which learners work in pairs or small groups to provide explicit teaching support. There are two main types of peer tutoring: either same age or cross age. In cross-age peer tutoring, an older learner takes the tutoring role and is paired with a younger tutee or tutees. There are a number of same-age approaches such as Peer-Assisted Learning. In Reciprocal Peer Tutoring, learners alternate between the role of tutor and tutee. In peer-assisted learning and reciprocal peer tutoring the common characteristic of these approaches is that learners take on responsibility for aspects of teaching and for evaluating the success of their peer or peers. In most peer tutoring approaches learners are instructed in how to undertake their roles effectively, often with specific and structured aspects of the interaction (such as learning the question types in reciprocal peer tutoring, or using specific prompts and questions in cross age peer tutoring).

Peer assessment involves learners of the same or different ages providing feedback relating to aspects of their academic performance and can have different forms such as reinforcing or correcting aspects of learning. Where this includes a teaching role to support the learner being assessed to act on such feedback, studies are to be included as peer tutoring. If peer assessment is undertaken purely as marking, particularly if the aim is to develop the assessor's understanding of the marking criteria, without support to improve, it would not be included.

Peer tutoring is related to a number of other Toolkit themes. It is sometimes thought of as a form of collaborative learning in a broader range of strategies sometimes described as peer-mediated learning. However in the Toolkit we consider collaborative activities (and co-operative learning) as activities where the learners have a common aim or goal. This might be a co-operative task where group members do different aspects of the task but contribute to a common overall outcome (such as taking responsibility for different aspects of a presentation), or a shared task where group members work together throughout the activity (such as writing a joint article or report). In peer tutoring, by contrast, there is an explicit teaching and evaluation role. The tutor is often thought to improve due to the development of their meta-cognitive awareness and improvement in their capability to self-regulate their own learning. This is usually implicit however and can therefore be distinguished from the teaching and application of meta-cognitive and self-regulation strategies. In Peer tutoring, tutees are provided with direct feedback about what they are learning from their tutor, so this is an integral component of the approach.

Search terms:

Peer tutoring; peer assisted learning, peer teaching, peer instruction, peer help, peer buddy, peer involvement, reciprocal tutoring, reciprocal teaching

Evidence rating

There are eight meta-analyses included in the summary, with five of these conducted in last 10 years. The pooled effects from these syntheses range from 0.30 to 1.05 so do not provide a consistent estimate of effect. This variation is not explained by moderator analyses. A number of the meta-analyses include experimental and quasi-experimental studies which are not well controlled. In addition not all included studies have been adjusted for clustering. Recent studies have not replicated these effects. Overall the evidence is rated as extensive.

Technical Appendix: Peer tutoring

Additional cost information

The main financial cost of implementing a peer tutoring approach will be the cost of professional development. The average cost of professional development in EEF-funded programmes is well under £80 per pupil.

References

References
*Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, 42(1), 39-55. http://www.nasponline.org/publications/spr/abstract.aspx?ID=3706
*Cohen, P.A., Kulik, J.A., Kulik, C.C. (1982). Educational Outcomes of Tutoring: A Meta-Analysis of Findings. *American Educational Research Journal*, 19(2), 237-248. http://www.dx.doi.org/10.3102/00028312019002237
*Cook, S. B., Scruggs, T. E., Mastropieri, M. A., & Casto, G. C. (1985). Handicapped students as tutors. *The Journal of Special Education*, 19(4), 483-492.
*Ginsburg-Block, M.D., & Rohrbeck, C.A., (2006). A Meta-Analytic Review of Social, Self-Concept and Behavioral Outcomes of Peer-Assisted Learning. *Journal of Educational Psychology*, 98(4) 732-749. http://dx.doi.org/10.1037/0022-0663.98.4.732
*Jun, S.W., Ramirez, G., & Cumming, A. (2010). Tutoring Adolescents in Literacy: A Meta-Analysis. Journal of Education, 45(2) 219-238. http://mje.mcgill.ca/article/view/4770/6491
*Leung, K. C. (2014). Preliminary Empirical Model of Crucial Determinants of Best Practice for Peer Tutoring on Academic Achievement. *Journal of Educational Psychology* 107(2), 558-579. http://dx.doi.org/10.1037/a0037698 http://dx.doi.org/10.1037/a0037698
Lloyd, C., Morris, S., Edovald, T., Skipp, A., Kiss, Z., & Haywood, S. (2015a) *Durham Shared Maths Project Evaluation Report and Executive Summary.* London: EEF.
Lloyd, C., Edovald, T., Kiss, Z., Skipp, A., Morris, S., & Ahmed, H. (2015) *Paired Reading Evaluation Report and Executive Summary* London: EEF
*Rohrbeck, C., Ginsburg-Block, M.D., Fantuzzo, J. W. & Miller, T.R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, 95(2), 240-257. http://dx.doi.org/10.1037/0022-0663.95.2.240
Topping, K.J. (2005). Trends in Peer Learning *Educational Psychology* 25.6.pp 631-645. http://dx.doi.org/10.1080/01443410500345172
Topping, K., Miller, D., Thurston, A., McGavock, K. & Conlin, N. (2011). Peer tutoring in reading in Scotland: thinking big. *Literacy* 45.1 pp 3-9. http://dx.dp
Topping, K. J., Miller, D., Murray, P., Henderson, S., Fortuna, C., & Conlin, N. (2011). Outcomes in a randomised controlled trial of mathematics tutoring. *Educational Research*, 53(1), 51-63.
Tymms, P., Merrell, C., Thurston, A., Andor, J., Topping, K., & Miller, D. (2011). Improving attainment across a whole district: school reform through peer tutoring in a randomized controlled trial. *School Effectiveness and School Improvement*, 22(3), 265-289.
*Washington State Institute for Public Policy (2014) Benefit-Cost Results Tutoring: By peers: http://www.wsipp.wa.gov/BenefitCost/ProgramPdf/107/Tutoring-By-peers
What Works Clearinghouse, U.S. Department of Education, Institute of Education Sciences (2013). Elementary School Mathematics intervention report: Peer-Assisted Learning Strategies. Retrieved from http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=619

Technical Appendix: Peer tutoring

Summary of Effects		
Study	**FSM Effect**	**Effect size**
Meta-analyses		
Bowman-Perrot et al., 2013		(0.75)[1]
Cohen, Kulik and Kulik, 1982 (on tutees)		0.40
Cohen, Kulik and Kulik, 1982 (on tutors)		0.33
Cook et al. 1985 (on tutees)		0.59
Cook et al. 1985 (on tutors)		0.65
Ginsburg-Block et al., 2006	0.38	0.35
Jun, Ramirez & Cumming, 2010 (cross age peer tutoring)		1.05
Leung, 2014	0.35	0.39
Rohrbeck et.al., 2003		0.59
Washington State Institute (2014)		0.43
Indicative effect size (median)		**0.43**
Recent studies[2]		
Lloyd et al. 2015a (cross age on tutees Y7)	-0.04	-0.02
Lloyd et al. 2015a (cross age on tutors Y9)	-0.06	-0.06
Lloyd et al. 2015b (cross age on tutees Y3)	-0.04	0.01
Lloyd et al. 2015b (cross age on tutors Y5)	0.05	0.02
What Works, 2013		0.06

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses and abstracts	
Meta-analysis	**Abstract**
Bowman-Perrott et al. 2013	Peer tutoring is an instructional strategy that involves students helping each other learn content through repetition of key concepts. This meta-analysis examined effects of peer tutoring across 26 single-case research experiments for 938 students in Grades 1-12. The TauU effect size for 195 phase contrasts was 0.75 with a confidence interval of CI₉₅ = 0.71 to 0.78, indicating that moderate to large academic benefits can be attributed to peer tutoring. Five potential moderators of these effects were examined: dosage, grade level, reward, disability status, and content area. This is the first peer tutoring meta-analysis in nearly 30 years to examine outcomes for elementary and secondary students, and extends previous peer tutoring meta-analyses by examining disability as a potential moderator. Findings suggest that peer tutoring is an effective intervention regardless of dosage, grade level, or disability status. Among students with disabilities, those with emotional and behavioral disorders benefitted most. Implications are discussed.
Cohen et al. 1982	A meta-analysis of findings from 65 independent evaluations of school

* included meta-analyses

[1] Single case meta-analysis using TauU (CI 0.71-0.78), a non-parametric effect size indicating moderate to large effects.

[2] Not included in any of the meta-analyses

	tutoring programs showed that these programs have positive effects on the academic performance and attitudes of those who receive tutoring. Tutored students outperformed control students on examinations, and they also developed positive attitudes toward the subject matter covered in the tutorial programs. The meta-analysis also showed that tutoring programs have positive effects on children who serve as tutors. Like the children they helped, the tutors gained a better understanding of and developed more positive attitudes toward the subject matter covered in the tutorial program. Participation in tutoring programs had little or no effect, however, on the self-esteem of tutors and tutees.
Cook et al. 1985	A meta-analysis was conducted on available research documenting effectiveness of handicapped students as tutors of other students. Nineteen articles yielding 74 effect sizes were located. Results indicated that (a) tutoring programs were generally effective; (b) tutees generally gained more than tutors; and (c) tutor and tutee gains on self-concept and sociometric ratings were small, while gains on attitude measures were larger. Implications for instruction and further research are given.
Ginsburg-Block et al. 2006	Meta-analysis was used to examine social, self-concept, and behavioral effects of peer-assisted learning (PAL) interventions with elementary school students. An electronic search of PsycINFO and ERIC databases resulted in 36 relevant PAL studies. Overall, effect sizes were small to moderate across the 3 outcome variable domains. Both social and self-concept outcomes were positively correlated with academic outcomes. Specific PAL components—student autonomy, individualized evaluation, structured student roles, interdependent group rewards, and same-gender grouping—were related to effect sizes. PAL interventions were more effective for low-income versus higher income, urban versus suburban–rural, minority versus nonminority, and Grades 1–3 students versus Grades 4–6 students. Results suggest that PAL interventions that focus on academics can also improve social and self-concept outcomes.
Jun et al. 2010	What does research reveal about tutoring adolescents in literacy? We conducted a meta-analysis, identifying 152 published studies, of which 12 met rigorous inclusion criteria. We analyzed the 12 studies for the effects of tutoring according to the type, focus, and amount of tutoring; the number, age, and language background of students; and the quality of the research. Despite variability, these studies suggest benefits, notably for cross-age tutoring, reading, and small tutoring programs of lengthy duration.
Leung 2014 (see also Leung et al. 2005)	Previous meta-analyses of the effects of peer tutoring on academic achievement have been plagued with theoretical and methodological flaws. Specifically, these studies have not adopted both fixed and mixed effects models for analyzing the effect size; they have not evaluated the moderating effect of some commonly used parameters, such as comparing same-age reciprocal peer tutoring, same-age nonreciprocal, or cross-age peer tutoring; considered the educational level of tutee or tutor; or properly addressed publication bias. Most studies are confined to specific populations and particular subjects (mainly mathematics and reading), and some studies are confounded by other types of intervention (such as cooperative learning or adult-led tutoring). Hence, there is a compelling need for an updated, comprehensive meta-analysis evaluating the effect of peer tutoring on academic achievement that incorporates advances in methodology, is not confounded by other modes of peer learning, and engages a wide range of participants and various subjects. The present study demonstrates that peer tutoring has a positive impact on academic achievement. The moderators and crucial determinants of the

	effectiveness of peer tutoring are identified and compared. Moreover, program parameters based on the concepts of role theory and interdependent group contingencies are evaluated. Finally, a preliminary empirical model of the crucial determinants of best practices for peer tutoring on academic achievement is proposed.
Rohrbeck et al. 2003	A meta-analytic review of group comparison design studies evaluating peer-assisted learning (PAL) interventions with elementary school students produced positive effect sizes (ESs) indicating increases in achievement (un-weighted mean ES _ 0.59, *SD* _ 0.90; weighted ES, *d* _ 0.33, *p* _ .0001, 95% confidence interval _ 0.29–0.37). PAL interventions were most effective with younger, urban, low income, and minority students. Interventions that used interdependent reward contingencies, ipsative evaluation procedures, and provided students with more autonomy had higher ESs. Adequate descriptive information was missing in many studies. Researchers are encouraged to develop PAL interventions in collaboration with practitioners to maximize those interventions' use and effectiveness and to include more detailed information about students, schools, and intervention components in their reports.
Washington State Institute 2014	Generally, peer tutoring is an instructional strategy that uses students to provide academic assistance to struggling peers. Peer tutoring may use students from the same classrooms or pair older students with younger struggling students. Tutoring assistance can occur through one-on-one interactions or in small groups and in some instances students alternate between the role of tutor and tutee. The specific types of peer tutoring that have been evaluated and are included in this meta-analysis are (in no particular order): ClassWide Peer Tutoring, Peer-Assisted Learning Strategies, and Reciprocal Peer Tutoring. The evaluated tutoring programs in this analysis provide, on average, about 30 hours of peer tutoring time each year and about 6 hours of training time for teachers and students to learn program procedures.

Performance pay

Very low or no impact for low cost, based on very limited evidence.

0

Performance pay schemes create a direct link between a teacher's wages or bonus, and the performance of their class. A distinction can be drawn between awards, where improved performance leads to a higher permanent salary, and payment by results, where teachers get a bonus for higher test scores. A second key issue is how performance is measured and how closely this is linked to outcomes for learners. In some schemes, students' test outcomes are the sole factor used to determine performance pay awards. In others, performance judgements can also include information from lesson observations or feedback from pupils, or be left to the discretion of the headteacher.

How effective is it?

Estimates based on cross-national comparisons suggest that performance pay could lead to positive impacts of around three months. However, the results of more rigorous evaluations, such as those with experimental trials or with well-controlled groups, suggest that the actual average impact has been close to zero. In India, there is evidence of the benefit of performance pay in the private sector but not the state sector, but it is not clear how this evidence applies to other systems.

As the evaluations of a number of performance pay schemes in the USA, where the approach is also known as 'merit pay', have been unable to find a clear link with student learning outcomes, investing in performance pay would not appear to be a good investment without further study. There are a number of examples of unintended consequences of performance pay from the US and elsewhere, which suggests that designing effective performance pay schemes is difficult.

Evaluations of the English threshold assessment introduced in 2000 offer a cautious endorsement of approaches which seek to reward teachers in order to benefit disadvantaged students by recognising teachers' professional skills and expertise. However, approaches which simply assume that incentives will make teachers work harder do not appear to be well supported by existing evidence.

How secure is the evidence?

The evidence is not conclusive. Although there has been extensive research into performance pay, most of this is either from correlational studies linking national pay levels with general national attainment or from naturally occurring experiments. In the latter it is hard to measure other variables which may influence the impact of pay increases, such as teaching to the test or other forms of "gaming". Overall, it is hard to make causal claims about the efficacy of performance pay on the basis of existing evidence.

What are the costs?

Increases are usually of the order of £2,500 per teacher or £100 per pupil across a class of 25. Overall cost estimates are therefore low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. It is clearly important to recruit the most effective teachers possible, and any additional resource may be better targeted at identifying and appointing the best teachers for a school.
2. Performance pay has been tried on a number of occasions, however the evidence of impact on student learning does not support the approach.
3. Spending on professional development linked to evaluation of better learning by pupils may also offer an alternative to performance pay.
4. Performance pay may lead to a narrower focus on test performance and restrict other aspects of learning.

Performance pay
Toolkit references

Full references
Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H. & Wilson, D. (2009). Evaluating the Impact of Performance-related Pay for Teachers in England. Labour Economics 16:3, pp 251-261.
Burgess, S., Croxson, B., Gregg, P. & Propper, C. (2001). The Intricacies of the Relationship Between Pay and Performance for Teachers: Do teachers respond to Performance Related Pay schemes? CMPO Working Paper Series No. 01/35.
Dolton, P. & Marcenaro-Gutierrez, O. (2011). If You Pay Peanuts, Do You Get Monkeys? A Cross-country Analysis of Teacher Pay and Pupil Performance'. Economic Policy 26(65): 5-55,
Education Commission of the States (2010). Teacher Merit Pay: What do we Know?. The Progress of Education Reform, 11(3).
Fryer, R.G. (2011). Teacher Incentives and Student Achievement: Evidence from New York City Public Schools. National Bureau of Economic Research, Working Paper 16850.
Glazerman, S. & Seifullah, A. (2010). An Evaluation of the Teacher Advancement Program (TAP) in Chicago: Year Two Impact Report. Mathematica Policy Research, Inc.
Lavy, V. (2002). Evaluating the Effects of Teachers' Group Performance Incentives on Pupil Achievement. Journal of Political Economy, 110(6), 1286-1317.
Martins, P.S. (2009). Individual Teacher Incentives, Student Achievement and Grade Inflation. Centre for Globalization Research: Working Paper 29. Queen Mary, University of London.
Podgursky, M.J. & Springer, M.G. (2006). Teacher Performance Pay: A Review. National Center on Performance Incentives.
Springer, M.G., Ballou, D., Hamilton, L., Le, V.N., Lockwood, J.R., McCaffrey, D.F., Pepper, M. & Stecher, B.M. (2010a). Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching. Project on Incentives in Teaching, National Center on Performance Incentives.
Springer, M.G., Lewis, J.L., Ehlert, M.W., Podgursky, M.J., Crader, G.D., Taylor, L.L., Gronberg, T.J., Jansen, D.W., Lopez, O.S. & Stuit, D.A. (2010b). District Awards for Teacher Excellence (D.A.T.E.) Program: Final Evaluation Report. National Center on Performance Incentives.
Woessmann, L. (2010). Cross-Country Evidence on Teacher Performance Pay. CESifo Working Paper No. 3151 Category 5: Economics Of Education Munich: CESifo.
Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H. & Wilson, D. (2009). Evaluating the Impact of Performance-related Pay for Teachers in England. Labour Economics 16:3, pp 251-261.
Burgess, S., Croxson, B., Gregg, P. & Propper, C. (2001). The Intricacies of the Relationship Between Pay and Performance for Teachers: Do teachers respond to Performance Related Pay schemes? CMPO Working Paper Series No. 01/35.
Dolton, P. & Marcenaro-Gutierrez, O. (2011). If You Pay Peanuts, Do You Get Monkeys? A Cross-country Analysis of Teacher Pay and Pupil Performance'. Economic Policy 26(65): 5-55,

Summary of effects	
Study	***Effect size***
Martins, 2009 (single study)	-0.09
Woessman, 2010 (correlational)	0.25
Indicative effect size	**0.00**

For more information about the effect sizes in the Toolkit, click here.

Performance pay
Toolkit references

Meta-analyses abstracts	
Study	***Abstract***
Martins (2009).	How do teacher incentives affect student achievement? Here we examine the effects of the recent introduction of teacher performance-related pay and tournaments in Portugal's public schools. Specifically, we conduct a difference-in-differences analysis based on population matched student-school panel data and two complementary control groups: public schools in autonomous regions that were exposed to lighter versions of the reform; and private schools, which are subject to the same national exams but whose teachers were not affected by the reform. We found that the focus on individual teacher performance decreased student achievement, particularly in terms of national exams, and increased grade inflation.
Woessmann (2010).	The general-equilibrium effects of performance-related teacher pay include long-term incentive and teacher-sorting mechanisms that usually elude experimental studies but are captured in cross-country comparisons. Combining country-level performance-pay measures with rich PISA-2003 international achievement micro data; this paper estimates student-level international education production functions. The use of teacher salary adjustments for outstanding performance is significantly associated with math, science, and reading achievement across countries. Scores in countries with performance-related pay are about one quarter standard deviations higher. Results avoid bias from within-country selection and are robust to continental fixed effects and to controlling for non-performance-based forms of teacher salary adjustments.

Phonics

Moderate impact for very low cost, based on very extensive evidence.

+4

Phonics is an approach to teaching reading, and some aspects of writing, by developing learners' phonemic awareness. This involves the skills of hearing, identifying and using phonemes or sound patterns in English. The aim is to systematically teach learners the relationship between these sounds and the written spelling patterns, or graphemes, which represent them. Phonics emphasises the skills of decoding new words by sounding them out and combining or 'blending' the sound-spelling patterns.

How effective is it?

Phonics approaches have been consistently found to be effective in supporting younger readers to master the basics of reading, with an average impact of an additional four months' progress. Research suggests that phonics is particularly beneficial for younger learners (4-7 year olds) as they begin to read. Teaching phonics is more effective on average than other approaches to early reading (such as whole language or alphabetic approaches), though it should be emphasised that effective phonics techniques are usually embedded in a rich literacy environment for early readers and are only one part of a successful literacy strategy.

For older readers who are still struggling to develop reading skills, phonics approaches may be less successful than other approaches such as Reading comprehension strategies and Meta-cognition and self-regulation. The difference may indicate that children aged 10 or above who have not succeeded using phonics approaches previously require a different approach, or that these students have other difficulties related to vocabulary and comprehension which phonics does not target.

Qualified teachers tend to get better results when delivering phonics interventions (up to twice the effectiveness of other staff), indicating that expertise is a key component of successful teaching of early reading.

How secure is the evidence?

Overall, the evidence base related to phonics is very secure. There have been a number of studies, reviews and meta-analyses that have consistently found that the systematic teaching of phonics is beneficial. There is some evidence that approaches informed by synthetic phonics (where the emphasis is on sounding out letters and blending sounds to form words) may be more beneficial than analytic approaches (where the sound/symbol relationship is inferred from identifying patterns and similarities by comparing several words). However, the evidence here is less secure and it is probably more important to match the teaching to children's particular needs and systematically teach the sound patterns with which they are not yet confident.

Several robust studies of phonics programmes in English have been published in recent years. The findings show that phonics programmes can be effective in English schools, but also underline the importance of high quality implementation. Recent evaluations of Switch-on Reading, a programme involving phonics components delivered by teaching assistants, and Fresh Start, showed that both had an average impact of three additional months' progress. However two other programmes, both targeting struggling, older readers, did not improve reading outcomes.

What are the costs?

Overall, the costs are estimated as very low. The costs associated with teaching phonics arise from the need for specific resources and professional training. Evidence suggests that the effectiveness of phonics is related to the pupil's stage of reading development, so it is also important that teachers have professional development in effective assessment as well as in the use of particular phonic techniques and materials.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Phonics can be an important component in the development of early reading skills, particularly for children from disadvantaged backgrounds. However, it is also important that children are successful in making progress in all aspects of reading including vocabulary development, comprehension and spelling, which should be taught separately and explicitly.
2. The teaching of phonics should be explicit and systematic to support children in making connections between the sound patterns they hear in words and the way that these words are written.
3. The teaching of phonics should be matched to children's current level of skill in terms of their phonemic awareness and their knowledge of letter sounds and patterns (graphemes).
4. Phonics improves the accuracy of the child's reading but not the comprehension. How are you planning on developing wider literacy skills such as comprehension?

Technical Appendix: Phonics

Definition

Phonics is an approach to teaching reading, and some aspects of writing and spelling, by developing learners' phonemic awareness. In linguistics, a phoneme is the smallest unit of speech that can be used to make one word different from another word. Phonics approaches therefore involve the skills of hearing, identifying and using sound patterns or phonemes in English. The aim is to teach learners the relationship between these sounds and the written spelling patterns, or graphemes, which represent them. Phonics emphasises the skills of decoding new words by sounding them out and combining or 'blending' the sound-spelling patterns. There are two main approaches to teaching phonics: analytic and synthetic phonics. In both approaches the learner needs to have some phonological awareness (the ability to hear and discriminate sounds in spoken words). Synthetic phonics focuses on the development of phonemic awareness as a key skill. To learn to decode written text into sounds, a reader is taught up to 44 phonemes (the smallest units of sound) and their related graphemes (the written symbols for these phonemes). Analytic phonics, also sometimes known as the "whole word" approach, involves analysis of whole words to detect phonetic or orthographic (spelling) patterns, then splitting them into smaller parts and sounding these out to help with the decoding process.

Search terms:

Phonics, analytic* phonics, synthetic phonics, phonemic awareness.

Evidence rating

There are seven meta-analyses and one best-evidence synthesis with quantitative estimates of impact on attainment (effect sizes). Five of the meta-analyses have been conducted in the last ten years. There is high quality evidence in these syntheses where the majority of the included studies have ecological validity and where the outcome measures include curriculum measures or standardised tests in school subject areas. The pooled effect size estimates range from 0.24 to 0.62, with some of the variation explained by intensity (particularly one-to-one and small group) and outcome measures (higher effects for word level measures and lower for comprehension). Overall the evidence is rated as very extensive.

Additional cost information

The main financial cost of implementing a phonics approach will be the cost of professional development. The average cost of professional development in EEF-funded programmes is well under £80 per pupil. Overall the cost is estimated as very low.

References

Full references
Berkeley, S., Scruggs, S.T. & Mastropieri, M.A. (2010). Reading Comprehension Instruction for Students With Learning Disabilities, 1995--2006: A Meta-Analysis. *Remedial and Special Education* 31, 423-436. http://dx.doi.org/10.1177/0741932509355988
*Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2008). Meta-Analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, 112.3: pp. 579–620. http://dx.doi.org/10.14507/epaa.v11n15.2003
Connor, C.M., Morrison, F.J., Fishman, B.J., Schatschneider, C. & Underwood, P. (2007). Algorithm-guided individualized reading instruction. *Science*, 315, 464–465. http://dx.doi.org/10.1126/science.1134513
*Ehri, C.L., Nunes, S.R., Stahl, S.A., & Willows, D.M. (2001). Systematic Phonics Instruction Helps Students Learn to Read: Evidence from the National Reading Panel's Meta-Analysis. *Review of Educational Research,* 71, (3) 393-447. http://dx.doi.org/10.3102/00346543071003393
*Galuschka K, Ise E, Krick K, Schulte-Körne G (2014) *Effectiveness of Treatment Approaches for Children and Adolescents with Reading Disabilities: A Meta-Analysis of Randomized Controlled Trials.* PLoS ONE 9(2): e89900. http://dx.doi.org/10.1371/journal.pone.0089900
Gorard, S., See, B. H., & Siddiqui, N. (2014). *Switch-on Reading: Evaluation Report and Executive Summary*. London: Education Endowment Foundation http://educationendowmentfoundation.org.uk/uploads/pdf/FINAL_EEF_Evaluation_Report_-_Switch-on_-_February_2014.pdf
Gorard, S., Siddiqui, N. & See, B.H. (2015) *Fresh Start Evaluation Report and Executive Summary* London: EEF https://educationendowmentfoundation.org.uk/uploads/pdf/Fresh_Start_(Final).pdf
*Jeynes, W.H. (2008). A Meta-Analysis of the Relationship between Phonics Instruction and Minority Elementary School Student Academic Achievement. *Education and Urban Society*. 40 (2), 151-166. http://dx.doi.org/10.1177/0013124507304128
King, B. & Kasim, A. (2015) Rapid Phonics Evaluation Report and Executive Summary. London: EEF https://educationendowmentfoundation.org.uk/uploads/pdf/Rapid_Phonics_(Final).pdf
*McArthur G, Eve PM, Jones K, Banales E, Kohnen S, Anandakumar T, Larsen L, Marinus E, Wang HC, Castles A. (2012) *Phonics training for English-speaking poor readers.* Cochrane Database of Systematic Reviews, Issue 12. Art. No.: CD009115. http://dx.doi.org/10.1002/14651858.CD009115.pub2
Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychological bulletin*, *138*(2), 322. http://dx.doi.org/10.1037/a0026744
Merrell, C. & Kasim, A. (2015) Butterf*ly Phonics Evaluation Report and Executive Summary.* London: EEF https://educationendowmentfoundation.org.uk/uploads/pdf/Butterfly_Phonics_(Final).pdf
Savage, R., Burgos, G., Wood, E., & Piquette, N. (2015). The Simple View of Reading as a framework for national literacy initiatives: a hierarchical model of pupil-level and classroom-level factors. *British Educational Research Journal.* http://dx.doi.org/10.1002/berj.3177
Sheard, M., Chambers, B. & Elliott, L.2015) *Units of Sound Evaluation Report and Executive Summary.* London: EEF https://educationendowmentfoundation.org.uk/uploads/pdf/Units_of_sound_(Final).pdf
* Sherman, K. H. (2007). *A meta-analysis of interventions for phonemic awareness and phonics instruction for delayed older readers.* Doctoral Thesis University of Oregon. UMI No: 3285626 ProQuest Dissertations and Theses: http://search.proquest.com/docview/304825094
* Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1-26. http://dx.doi.org/10.1016/j.edurev.2010.07.002
Suggate, S. P. (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental Psychology*, 46(6), 1556.

http://dx.doi.org/10.1037/a0020612
Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73*(4), 407-440. http://dx.doi.org/10.3102/00346543073004407
*Torgerson, C., Brooks, G., & Hall, J. (2008). A Systematic Review of the Research Literature on the Use of Phonics in the Teaching of Reading and Spelling. Department for Education and Skills.
*Studies included in the summary of effects.

Summary of Effects		
Meta-analyses	**FSM effect size**	**Overall effect size**
Camilli, Vargas & Yurecko, 2003		0.24
Ehri, Nunes, Stahl & Willows, 2001	0.66	0.41
Galuschka et al. 2014		0.32
Jeynes, 2008		0.30
McArthur et al. 2012 (word reading accuracy) (reading comprehension)		0.62 0.35
Sherman, 2007 (older readers)		0.39
Slavin et al. 2011 (one-to-one phonics tutoring) (small group phonics)		0.62 0.35
Torgerson, Brooks & Hall, 2006		0.27
Recent studies		
¶Gorard, Siddiqui & See, 2015 (EEF Fresh Start)	0.24	0.24
Gorard, See & Siddiqi, 2014 (EEF Switch-on Reading)	0.36	0.24
King & Kasim, 2015 (EEF Rapid Phonics)	-0.07	-0.07
¶Merrell & Kasim, 2015 (EEF Butterfly Phonics)	0.16	0.43
¶Sheard, Chambers, Elliott, 2015 (EEF Units of Sound)	-0.21	-0.08
Weighted mean effect size		**0.35**

¶Due to the nature of the study design and the implementation of intervention this finding has limited security.

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses and abstracts	
Camilli et al. 2003	Examined the findings of the "Teaching Children To Read" study of the National Reading Panel and the procedures of the study. Meta-analytic techniques found that the methodology and procedures were not adequate. Findings suggest that phonics, as an aspect of the complex reading process, should not be over-emphasized.
Ehri et al. 2001	A quantitative meta-analysis evaluating the effects of systematic phonics instruction compared to unsystematic or no-phonics instruction on learning to read was conducted using 66 treatment-control comparisons derived from 38 experiments. The overall effect of phonics instruction on reading was moderate, d = 0.41.
Galuschka, Ise, Krick, & Schulte-Körne 2014	Children and adolescents with reading disabilities experience a significant impairment in the acquisition of reading and spelling skills. Given the emotional and academic consequences for children with persistent reading disorders, evidence based interventions are critically needed. The present meta-analysis extracts the results of all available randomized controlled trials. The aims were to determine the effectiveness of different treatment approaches and the impact of various factors on the efficacy of interventions. The literature search for published randomized-controlled trials comprised an electronic search in the databases ERIC, PsycINFO,

	PubMed, and Cochrane, and an examination of bibliographical references. To check for unpublished trials, we searched the websites clinicaltrials.com and ProQuest, and contacted experts in the field. Twenty-two randomized controlled trials with a total of 49 comparisons of experimental and control groups could be included. The comparisons evaluated five reading fluency trainings, three phonemic awareness instructions, three reading comprehension trainings, 29 phonics instructions, three auditory trainings, two medical treatments, and four interventions with coloured overlays or lenses. One trial evaluated the effectiveness of sunflower therapy and another investigated the effectiveness of motor exercises. The results revealed that phonics instruction is not only the most frequently investigated treatment approach, but also the only approach whose efficacy on reading and spelling performance in children and adolescents with reading disabilities is statistically confirmed. The mean effect sizes of the remaining treatment approaches did not reach statistical significance. The present meta-analysis demonstrates that severe reading and spelling difficulties can be ameliorated with appropriate treatment. In order to be better able to provide evidence-based interventions to children and adolescent with reading disabilities, research should intensify the application of blinded randomized controlled trials.
Jeynes 2008	This meta-analysis of 22 studies examines the relationship between phonics and the academic achievement of urban minority elementary school children. Further analyses distinguish between those studies that are of higher quality than the others and those studies that examine all minority students and mostly minority students. Results indicate a significant relationship between phonics instruction and higher academic achievement. Phonics instruction, as a whole, is associated with academic variables by about .23 to .33 of a standard deviation unit. This relationship holds for studies that examine all minority students and those that include mostly minority students. The results also hold for higher quality studies. The significance of these results is discussed.
McArthur et al. 2012	Around 5% of English speakers have a significant problem with learning to read words. Poor word readers are often trained to use letter-sound rules to improve their reading skills. This training is commonly called phonics. Well over 100 studies have administered some form of phonics training to poor word readers. However, there are surprisingly few systematic reviews or meta-analyses of these studies. The most well-known review was done by the National Reading Panel (Ehri 2001) 12 years ago and needs updating. The most recent review (Suggate 2010) focused solely on children and did not include unpublished studies. Objectives: The primary aim of this review was to measure the effect that phonics training has on the literacy skills of English-speaking children, adolescents, and adults whose reading was at least one standard deviation (SD), one year, or one grade below the expected level, despite no reported problems that could explain their impaired ability to learn to read. A secondary objective was to explore the impact of various factors, such as length of training or training group size, that might moderate the effect of phonics training on poor word reading skills. Search methods. Selection criteria: We included studies that use randomisation, quasi-randomisation, or minimisation to allocate participants to either a phonics intervention group (phonics alone, phonics and phoneme awareness training, or phonics and irregular word reading training) or a control group (no training or alternative training, such as maths). Participants were English-speaking children, adolescents, or adults whose word reading was below the level expected for their age for no known reason (that is, they had adequate attention and no known physical,

	neurological, or psychological problems). Data collection and analysis. Two review authors independently selected studies, assessed risk of bias, and extracted data. Main results. We found 11 studies that met the criteria for this review. They involved 736 participants. We measured the effect of phonics training on eight outcomes. Authors' conclusions. Phonics training appears to be effective for improving some reading skills. Specifically, statistically significant effects were found for nonword reading accuracy (large effect), word reading accuracy (moderate effect), and letter-sound knowledge (small-to-moderate effect). For several other outcomes, there were small or moderate effect sizes that did not reach statistical significance but may be meaningful: word reading fluency, spelling, phonological output, and reading comprehension. The effect for nonword reading fluency, which was measured in only one study, was in a negative direction, but this was not statistically significant.
Sherman, 2007	The purpose of this study was to synthesize, using meta-analytical methods, the research regarding phonemic awareness and phonics (decoding) instruction with students in grades 5 through 12 who read significantly below grade level expectations. Twenty-six studies published between 1975 and 2005 met the criteria for inclusion and analysis. A total of 1358 students participated in the studies (565 in control groups, 799 in treatment groups). The effect sizes of interventions = impact on achievement were calculated on five levels of dependent variables (word identification or word attack skills of sub-syllabic or single syllable levels, and decoding multi-syllabic words; oral reading fluency and accuracy of individual words or connected text; comprehending words or vocabulary; comprehending text; decoding, fluency and comprehension). Four separate analyses were presented: (a) the full data set; (b) the data set with outliers removed; (c) the full data set without one study (Mercer, Miller, Mercer, & Lane, 2000); and (d) the data without outliers and without the Mercer study. Although many of the studies exhibited medium to high effect sizes, none of the analyses at an alpha level of 0.05 reached statistical significance. Because of the small number of studies and the variability of the population studied, the alpha level was relaxed to 0.25 to explore statistical significance of main effects or interaction effects at this level. The impact of group size and reading level on effect size was significant in many of the analyses at a 0.25 alpha level. The results were mixed for group size/intervention focus and reading level/intervention focus. Significant main effects were found for reading level (reading level*intervention focus) and the interaction between group size and intervention focus on word identification or word attack skills of sub-syllabic or single syllable levels, and decoding multi-syllabic words. The impact of reading level, group size, and intervention focus on effect size were not significant at any level. Limitations of this meta-analysis, features of interventions that show promise in accelerating the reading skills of delayed older readers, and suggestions for future research are also presented.
Slavin et al. 2011	This article reviews research on the achievement outcomes of alternative approaches for struggling readers ages 5–10 (US grades K-5): One-to-one tutoring, small-group tutorials, classroom instructional process approaches, and computer-assisted instruction. Study inclusion criteria included use of randomized or well-matched control groups, study duration of at least 12 weeks, and use of valid measures independent of treatments. A total of 97 studies met these criteria. The review concludes that one-to-one tutoring is very effective in improving reading performance. Tutoring models that focus on phonics obtain much better outcomes than others. Teachers are more

	effective than paraprofessionals and volunteers as tutors. Small-group, phonetic tutorials can be effective, but are not as effective as one-to-one phonetically focused tutoring. Classroom instructional process programs, especially cooperative learning, can have very positive effects for struggling readers. Computer-assisted instruction had few effects on reading. Taken together, the findings support a strong focus on improving classroom instruction and then providing one-to-one, phonetic tutoring to students who continue to experience difficulties.
Torgerson et al. 2008	Executive Summary The Department for Education and Skills (DfES) commissioned the Universities of York and Sheffield to conduct a systematic review of experimental research on the use of phonics instruction in the teaching of reading and spelling. This review is based on evidence from randomised controlled trials (RCTs). Key findings. The effect of phonics on reading: Systematic phonics instruction within a broad literacy curriculum was found to have a statistically significant positive effect on reading accuracy. There was no statistically significant difference between the effectiveness of systematic phonics instruction for reading accuracy for normally-developing children and for children at risk of reading failure. The weight of evidence for both these findings was moderate (there were 12 randomized controlled trials included in the analysis). Both of these findings provided some support for those of a systematic review published in the United States in 2001 (Ehri et al., 2001). An analysis of the effect of systematic phonics instruction on reading comprehension was based on weak weight of evidence (only four randomized controlled trials were found) and failed to find the statistically significant positive difference which was found in the previous review. The effect of synthetic and analytic phonics: The weight of evidence on this question was weak (only three randomized controlled trials were found). No statistically significant difference in effectiveness was found between synthetic phonics instruction and analytic phonics instruction. The effect of phonics on spelling: The weight of evidence on this question was weak (only three randomized controlled trials were found). No effect of systematic phonics instruction on spelling was found.

Physical environment

Very low or no impact for low cost based on very limited evidence.

0

Changing the physical learning environment, either by moving to a new school building; or seeking to improve the design, air quality, noise, light, or temperature of an existing building.

How effective is it?

Overall, changes to the physical environment of schools are unlikely to have a direct effect on learning beyond the extremes (i.e. once an adequate building standard has been achieved).

Moving to a new building could be an effective part of a whole school change that seeks to change behaviour and establish new norms (similar to introducing or changing School Uniform), but there is no evidence that new buildings or particular aspects of architecture directly improve learning. Where a new building is being used as a catalyst for change, there is some evidence supporting the impact of co-design, or involving potential beneficiaries in taking responsibility for learning spaces and changing their behaviours as they adapt to new settings.

Most individual factors in the physical environment show a relationship with learning only at the extremes. In terms of sound, if the noise levels are high (such as under the flight path of an airport) then there can be a measurable detrimental effect on learning. In terms of temperature, warmer and more humid conditions (particularly above 30°C conditions) can cause a loss of concentration and drowsiness. It appears that lighting in schools is usually adequate for reading and writing. The evidence on ambient music is inconclusive as it appears that people react differently to different kinds of music according to their preferences. Similarly with colour in the environment, personal preference is probably more important than any general effect.

Air quality is the one exception to the general picture on school environment. The evidence suggests low air quality does have a negative impact on attainment (reducing word recognition by 15% in one study).

How secure is the evidence?

The research on the impact of the physical environment on learning is generally weak, and is mainly based on correlational studies or drawn as inferences from wider environmental research. There are very few rigorous experimental designs, and this makes it hard to establish causal claims about the impact of physical changes.

Regarding air quality, there is evidence that some English classrooms have higher CO2 concentrations than the average recommended levels. Given that the link between air quality and academic achievement is better established than other aspects of the physical environment, this may be a worthwhile area for future study.

What are the costs?

It is very difficult to estimate the costs of physical changes as they are usually part of capital spending and a single cost, rather than a recurrent part of a school budget. A new secondary school costs about £15 million for 1,500 pupils or £10,000 per pupil. However several generations of pupils are likely to use the building. Improving air quality can be done relatively cheaply with better ventilation, filtration and the use of dehumidifiers where necessary. Overall, costs are estimated as low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Most environmental factors have an impact on classrooms only at the extremes.
2. Air quality is likely to be the most significant factor affecting learning, particularly where there is poor ventilation or high levels of dust and other pollutants
3. If you have a new learning environment, it provides an opportunity to change the expectations and behaviour of pupils, but it is unlikely to have a direct impact on learning. Have you considered how you will take advantage of the new environment to bring about improvements in these expectations?

Physical environment
Toolkit references

Full references
Baker, L., & Bernstein, H. (2012). The Impact of School Buildings on Student Health and Performance. Washington DC: The Center for Green Schools http://www.centerforgreenschools.org
Bakó-Biró, Z., Clements-Croome, D. J., Kochhar, N., Awbi, H. B., & Williams, M. J. (2012). Ventilation rates in schools and pupils' performance. Building and Environment, 48, 215-223.
Daisey, J. M., Angell, W. J., & Apte, M. G. (2003). Indoor air quality, ventilation and health symptoms in schools: an analysis of existing information. Indoor Air, 13(1), 53-64.
Tanner, C. K. (2000). The influence of school architecture on academic achievement. Journal of Educational Administration, 38(4), 309-330.
Waterhouse, L. (2006). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. Educational Psychologist, 41(4), 207-225.
Woolner, P., Hall, E., Higgins, S., McCaughey, C., & Wall, K. (2007). A sound foundation? What we know about the impact of environments on learning and the implications for Building Schools for the Future. Oxford Review of Education, 33(1), 47-70.
Woolner, P., Hall, E., Wall, K., Higgins, S., Blake, A. and McCaughey, C. (2005). School building programmes: motivations, consequences and implications. Reading: CfBT. Available at http://cdn.cfbt.com/~/media/cfbtcorporate/files/research/2005/r-school-building-programmes-2005.pdf
Woolner, P., McCarter, S., Wall, K. and Higgins, S. (2012). Changed Learning Through Changed Space. When can a Participatory Approach to the Learning Environment Challenge Preconceptions and Alter Practice? Improving Schools 15(1), 45-60.

Summary of effects	
Study	***Effect size***
There are no meta-analyses or systematic reviews with quantitative evidence of impact to allow general estimates of effect to be made. Overall effects are therefore estimated at no impact.	
Indicative effect size	**0.00**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	**Abstract**
See above.	

Reading comprehension strategies

Moderate impact for very low cost, based on extensive evidence.

+5

Reading comprehension approaches to improving reading focus on learners' understanding of the text. They teach a range of techniques that enable pupils to comprehend the meaning of what is written, such as inferring the meaning from context, summarising or identifying key points, using graphic or semantic organisers, developing questioning strategies, and monitoring their own comprehension and identifying difficulties themselves (see also Meta-cognition and self-regulation).

How effective is it?

On average, reading comprehension approaches improve learning by an additional five months' progress over the course of a school year. These approaches appear to be particularly effective for older readers (aged 8 or above) who are not making expected progress.

Successful reading comprehension approaches carefully select activities for pupils according to their reading capabilities, and ensure that texts provide an effective, but not overwhelming, challenge.

Many of the approaches can be usefully combined with phonics, collaborative and peer-learning techniques. The use of techniques such as graphic organisers and drawing pupils' attention to text structures are likely to be particularly useful when reading expository or information texts. There are also some indications that computer-based tutoring approaches can be successful in improving reading comprehension, particularly when they focus on the development of strategies and self-questioning skills, though the evidence is less robust in this area.

Comparative findings indicate that, on average, reading comprehension approaches appear to be more effective than phonics or oral language approaches for upper primary and secondary pupils, both in terms of short-term and long-term impact. However, supporting struggling readers is likely to require a concerted effort across the curriculum and a combination of different approaches. It is important to remember that no particular strategy should be seen as a panacea, and careful diagnosis of the reasons why an individual pupil is struggling is very important when exploring possible intervention strategies.

How secure is the evidence?

There is extensive evidence in this area, from a range of studies over the last 30 years. A majority of studies were conducted in the USA, and focus on pupils aged 8-18 who are falling behind their peers or have difficulties with reading.

In the UK, a recent evaluation of a programme that taught pupils to apply four reading comprehension strategies found some evidence of promise, but did not provide a robust estimate of the programme's impact.

What are the costs?

Costs for materials and professional development are estimated at £1,200 per teacher or £48 per pupil and therefore as very low. The costs associated with these approaches arise from the need for specific resources and professional training. Evidence suggests that the effectiveness of different approaches is related to the pupil's current capabilities in reading, so it is important that teachers receive professional development in effective diagnosis as well as in the use of particular techniques and materials to develop reading comprehension.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Effective diagnosis of reading difficulties is important in identifying possible solutions, particularly for older struggling readers. Are you confident that the problem(s) a pupil is facing in making expected progress is in decoding the words, understanding the structure of the language used or understanding particular vocabulary, which may be subject specific?
2. How can you focus learners' attention on developing comprehension strategies which they can apply more widely?
3. A wide range of strategies and approaches can be successful, but these need to be taught explicitly and consistently. How are you going to identify the strategies that will meet the needs of your pupils and how will these be reinforced?
4. A key issue for teachers is identifying the level of difficulty for comprehension activities that is required to extend pupils' reading capabilities. How will you ensure the texts used provide an effective challenge?

Technical Appendix:
Reading comprehension strategies

Definition

Teaching explicit reading comprehension strategies aims to improve learners' fluency and comprehension of written text by focusing on approaches and techniques a learner can use and take responsibility for themselves. It involves teaching a range of techniques that enable pupils to comprehend the meaning of what is written, such as inferring the meaning from context, summarising or identifying key points, using graphic or semantic organisers, developing questioning strategies, monitoring their own comprehension and identifying difficulties themselves (see also Meta-cognition and self-regulation).

Search terms: reading comprehension strategies; text comprehension strategies.

Evidence rating

There are seven meta-analyses which look at improving reading by developing reading comprehension strategies, five of which have been undertaken in the last 10 years. These mainly focus on helping low attaining readers aged 8 - 18 to catch up with their peers, rather than accelerating normal progress so tend to have relatively small samples. The evidence in this area dates back over last 30 years, with a majority of studies were conducted in the USA. The pooled effect sizes range between 0.10 to 0.52 (less than half of a standard deviation). Overall the evidence is rated as very extensive.

Additional cost information

Costs for materials and professional development are estimated at £1,200 per teacher or £48 per pupil and therefore as very low. The costs associated with these approaches arise from the need for specific resources and professional training. Evidence suggests that the effectiveness of different approaches is related to the pupil's current capabilities in reading, so it is important that teachers receive professional development in effective diagnosis as well as in the use of particular techniques and materials to develop reading comprehension.

References

Full references
Berkeley, S., Mastropieri, M. A., & Scruggs, T. E. (2011). Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities*, 44(1), 18-32. http://dx.doi.org/10.1177/0022219410371677
*Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading comprehension instruction for students with learning disabilities, 1995–2006: A meta-analysis. Remedial and Special Education. 31 (6): 423-436 http://dx.doi.org/10.1177/0741932509355988
Crawford, C. & Skipp, A. (2014*) LIT Programme Evaluation Report and Executive Summary October 2014.* London: EEF
*Davis, D. S. (2010). A meta-analysis of comprehension strategy instruction for upper elementary and middle school students (Doctoral dissertation, Vanderbilt University, USA). http://etd.library.vanderbilt.edu/available/etd-06162010-100830/unrestricted/Davis_dissertation.pdf
*Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. Review of Educational Research, 79(1), 262-300. http://www.dx.doi.org/10.3102/0034654308325998.

*Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. Journal of Research on Educational Effectiveness, 2(1), 1-44. http://www.tandfonline.com/doi/abs/10.1080/19345740802539200#.VMZuGv6sVK0
*Fauzan, N. (2003). The effects of metacognitive strategies on reading comprehension: a quantitative synthesis and the empirical investigation (Doctoral dissertation, University of Durham). http://etheses.dur.ac.uk/1086/
*Fukkink, R. G., & De Glopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. Review of Educational Research, 68(4), 450-469. http://www.dx.doi.org/10.3102/00346543068004450
Ness, M. (2011). Explicit reading comprehension instruction in elementary classrooms: Teacher use of reading comprehension strategies. Journal of Research in Childhood Education, 25(1), 98-117. http://www.dx.doi.org/10.1080/02568543.2010.531076
*Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in Grades 4–12: 1980–2011. *Journal of Learning Disabilities*, 48 (4): 369-390 http://www.dx.doi.org/10.1177/0022219413504995
Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). Interventions for Adolescent Struggling Readers: A Meta-Analysis with Implications for Practice. Portsmouth, NH: RMC Research Corporation: Center on Instruction. http://www.rti4success.org/sites/default/files/coi...
Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). Improving Reading Comprehension in Kindergarten through 3rd Grade: IES Practice Guide. NCEE 2010-4038. What Works Clearinghouse. http://files.eric.ed.gov/fulltext/ED512029.pdf
Stetter, M. E., & Hughes, M. T. (2010). Computer-Assisted Instruction to Enhance the Reading Comprehension of Struggling Readers: A Review of the Literature. Journal of Special Education Technology, 25(4). http://www.tamcec.org/jset/index/
Suggate, S. P. (2014). A Meta-Analysis of the Long-Term Effects of Phonemic Awareness, Phonics, Fluency, and Reading Comprehension Interventions. Journal of Learning Disabilities, http://www.dx.doi.org/10.1177/0022219414528540
*Studies included in the summary of effects.

Summary of effects		
Study	***FSM effect size***	***Effect size***
Berkeley et al. 2010		0.52
Davis 2010	0.31	0.36
Edmonds et al. 2009		0.47
Elleman et al. 2009		0.10
Fauzan 2003		0.50
Fukkink & De Glopper 1998		0.43
Scammacca et al. 2015		0.49
Recent studies		
Crawford et al. 2014 (EEF- LIT Programme)		0.09
Berkeley et al, 2011		0.71
Weighted mean effect size		**0.42**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Davis, 2010	This meta-analytic review includes intervention studies published between 1980 and 2009 in which students in grades 4-8 are taught to use two or more comprehension strategies. The collected studies were coded using a systematic data extraction scheme developed to address the central questions of the review. Information related to the characteristics of the student sample and instructional and methodological characteristics of each study were compiled in a database. Numerical effect sizes for each study for each major outcome measure were computed. The mean effect of comprehension strategy instruction on each of the targeted outcome constructs was calculated to provide an overall summary of instructional effectiveness.
Edmonds, Vaughn, Wexler, Reutebuch, Cable Tackett & Schnakenberg, 2009	This article reports a synthesis of intervention studies conducted between 1994 and 2004 with older students (Grades 6–12) with reading difficulties. Interventions addressing decoding, fluency, vocabulary, and comprehension were included if they measured the effects on reading comprehension. Twenty-nine studies were located and synthesized. Thirteen studies met criteria for a meta-analysis, yielding an effect size (ES) of 0.89 for the weighted average of the difference in comprehension outcomes between treatment and comparison students. Word-level interventions were associated with ES = 0.34 in comprehension outcomes between treatment and comparison
Elleman, Lindo, Morphy & Compton, 2009	A meta-analysis of vocabulary interventions in grades pre-K to 12 was conducted with 37 studies to better understand the impact of vocabulary on comprehension. Vocabulary instruction was found to be effective at increasing students' ability to comprehend text with custom measures (d = 0.50), but was less effective for standardized measures (d = 0.10). When considering only custom measures, and controlling for method variables, students with reading difficulties (d = 1.23) benefited more than three times as much as students without reading problems (d = 0.39) on comprehension measures. Gains on vocabulary measures, however, were comparable across reading ability. In addition, the correlation of vocabulary and comprehension effects from studies reporting both outcomes was modest (r = .43).
Fauzan, 2003	The purpose of the study was to investigate the effectiveness of metacognitive strategies on reading comprehension by means of (a) a meta-analysis and (b) an experiment designed following the meta-analysis implemented in Sarawak, Malaysia. Before the meta-analysis, the prevalent theories and issues in the reading literature such as metacognition, models of reading, measurements, motivation and previous meta-analysis were discussed to provide a better understanding of the research area in this study. A meta-analytic procedure conducted to review the primary research studies of metacognitive strategies used effect size as the measure of effectiveness. Searching for the articles and theses in the 1980s until 2001 yielded a record of 473 abstracts and articles from which there were twenty seven studies with a total number of eighty two effect sizes that could be quantitatively synthesized to compare the group performance of the experimental and control groups. The weighted effect size was 0.50 (95% CI = 0.45 to 0.56) when dependent effect sizes were synthesized, and 0.55 (95% CI=0.48 to 0.63) when the extreme 'outliers' or deviated effect sizes were excluded

	and independent effect sizes were created. Overall, the effect size was moderate indicating a positive outcome of the metacognitive strategies. The effect sizes were not homogeneous and further analyses of the qualitative and quantitative features of the studies were made to develop possible reliable estimates.
Fukkink & De Glopper, 1998	A meta-analysis of 21 instructional treatments aimed at enhancing the skill of deliberately deriving word meaning from context during reading shows a medium effect size of 0.43 standard deviation units ($p < .000$). An exploratory multilevel regression analysis shows that clue instruction appears to be more effective than other instruction types or just practice ($\beta = 0.40$). Effect size correlates negatively with class size ($\beta = .03$). Implications for instruction and future research are discussed. Future studies should investigate the effect of instruction on both the skill of deriving word meaning from context and incidental word learning to evaluate its contribution to vocabulary growth.
Scammacca, Roberts, Vaughn & Stuebing, 2013	This meta-analysis synthesizes the literature on interventions for struggling readers in Grades 4 through 12 published between 1980 and 2011. It updates Scammacca et al.'s analysis of studies published between 1980 and 2004. The combined corpus of 82 study-wise effect sizes was meta-analyzed to determine (a) the overall effectiveness of reading interventions studied over the past 30 years, (b) how the magnitude of the effect varies based on student, intervention, and research design characteristics, and (c) what differences in effectiveness exist between more recent interventions and older ones. The analysis yielded a mean effect of 0.49, considerably smaller than the 0.95 mean effect reported in 2007. The mean effect for standardized measures was 0.21, also much smaller than the 0.42 mean effect reported in 2007. The mean effects for reading comprehension measures were similarly diminished. Results indicated that the mean effects for the 1980–2004 and 2005–2011 groups of studies were different to a statistically significant degree. The decline in effect sizes over time is attributed at least in part to increased use of standardized measures, more rigorous and complex research designs, differences in participant characteristics, and improvements in the school's "business-as-usual" instruction that often serves as the comparison condition in intervention studies.
Scammacca, Roberts, Vaughn, Edmonds, Wexler, Reutebuch & Torgesen, 2007	This meta-analysis offers decision-makers research-based guidance for intervening with adolescent struggling readers. The authors outline major implications for practice: (1) Adolescence is not too late to intervene. Interventions do benefit older students; (2) Older students with reading difficulties benefit from interventions focused at both the word and the text level; (3) Older students with reading difficulties benefit from improved knowledge of word meanings and concepts; (4) Word-study interventions are appropriate for older students struggling at the word level; (5) Teachers can provide interventions that are associated with positive effects; (6) Teaching comprehension strategies to older students with reading difficulties is beneficial; (7) Older readers' average gains in reading comprehension are somewhat smaller than those in other reading and reading-related areas studied; (8) Older students with learning disabilities (LD) benefit from reading intervention when it is appropriately focused; and (9) To learn more about instructional conditions that could close the reading gap for struggling readers, individuals will need studies that provide instruction over longer periods of time and assess outcomes with measures more like those schools use to monitor reading progress of all students. This report summarizes aspects of recent research on reading instruction for adolescent struggling readers. It both synthesizes research findings to determine the

	relative effectiveness of interventions for struggling older readers and outlines the implications of these findings for practice. Its purpose is to advance the knowledge of technical assistance providers working with state departments of education and local education agencies concerning reading-related issues for students with reading difficulties and learning disabilities (LD). While the authors' methods and general findings are described, they are presented in terms of their impact on practice and policy. Specific suggestions for implementing these and other research findings are provided in an accompanying practice brief (ED521836). This report is intended primarily for technical assistance providers at Regional Comprehensive Centers for their use in crafting evidence-based guidance for states and local educational agencies.

Reducing class size

Moderate impact for very high cost, based on moderate evidence.

+3

Reducing the number of pupils in a class. As the size of a class or teaching group gets smaller it is suggested that the range of approaches a teacher can employ and the amount of attention each student will achieve will increase.

How effective is it?

Intuitively, it seems obvious that reducing the number of pupils in a class will improve the quality of teaching and learning, for example by increasing the amount of high quality feedback or one to one attention learners receive. However, overall the evidence does not show particularly large or clear effects, until class size is reduced to under 20 or even below 15.

The key issue appears to be whether the reduction is large enough to permit the teacher to change their teaching approach when working with a smaller class and whether, as a result, the pupils change their learning behaviours. If no change occurs then, perhaps unsurprisingly, learning is unlikely to improve. When a change in teaching approach does accompany a class size reduction (which appears hard to achieve until classes are smaller than about 20) then benefits on attainment can be identified, in addition to improvements on behaviour and attitudes. In some studies these benefits persist for a number of years (from early primary school through to at least the end of primary school). It appears to be very hard to achieve improvements from modest class size reductions above 20, e.g. from 30 to 25.

There is some evidence that reducing class sizes is more likely to be effective when supported with professional development to learn and develop teaching skills and approaches. Some evidence suggests slightly larger effects are documented for the lower achievers and those from the lower socio-economic status for very young pupils. Additionally teachers may potentially further develop their teaching skills and approaches in a smaller class.

How secure is the evidence?

There are a number of issues in interpreting the evidence about class size as many countries or schools already teach lower-attaining pupils in smaller groups. Overall, there is a relatively consistent picture where smaller classes are associated with slightly higher attainment (when other factors are controlled for) and when class sizes have been deliberately reduced in experimental evaluations.

The strongest evidence comes from research into primary schools in the USA where the benefits appear to be sustained for 3-4 years when classes are reduced below 18. There is some evidence that pupils in disadvantaged areas in the UK benefit from classes of fewer than 20 pupils in primary schools.

What are the costs?

The costs associated with reducing class sizes to a level where a significant benefit is likely are very high. The evidence suggests that typical classes would need to be halved to 15 pupils or even fewer. A class of 25 pupils with 50% of them receiving free school meals would be allocated an extra £8,000 under the pupil premium in 2012/13; this would not be sufficient to appoint an additional teacher. In 2013-14, a year group of 60 pupils where 50% were eligible for the Pupil Premium would increase funding by £27,000, enabling two classes of 30 to be split between three teachers with 20 pupils in each class. Costs are estimated as very high.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Small reductions (e.g. from 30 to 25 pupils) are unlikely to be cost-effective relative to other strategies.
2. Reducing class sizes for younger children may provide longer term benefits.
3. Smaller classes only impact upon learning if the reduced numbers allow teachers to teach differently. Have you considered how you will adjust your teaching strategies?
4. The gains from smaller class sizes are likely to come from the increased flexibility for organising learners and the quality and quantity of feedback the pupils receive. Have you considered how you will organise learning in smaller classes and how you will improve feedback to your pupils?
5. As an alternative to reducing class sizes, have you considered deploying staff (including teaching assistants) so that teachers can work more intensively with smaller groups?

Reducing class size
Toolkit references

Full references
Blatchford, P., Goldstein, H., Martin, C., & Browne, W. (2002). A study of class size effects in English school reception year classes. *British Educational Research Journal*, 28(2), 169–185.
Finn, J.D. & Achilles, C.M. (1999). Tennessee's class size study: findings implications, misconceptions *Educational Evaluation and Policy Analysis* 21.2 pp 97-109.
Glass, G.V. & Smith, M.L. (1978). *meta-analysis">Meta-analysis of research on the relationship of class size and achievement.* San Francisco: Far West Laboratory for Educational Research and Development.
Goldstein, H., Yang, M., Omar, R., Turner, R., & Thompson, S. (2000). meta-analysis">Meta-analysis using multilevel models with an application to the study of class size effects. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* 49(3), 399-412.
Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research* 43 (2005) pp 387–425.
McGiverin, J., Gilman, D., & Tillitski, C. (1989). A Meta-Analysis of the Relation between Class Size and Achievement. *The Elementary School Journal*, 90(1), 47.
Nye, B., Hedges, L.V., Konstantopoulos, S. (2004). Do Minorities Experience Larger Lasting Benefits from Small Classes? *Journal of Educational Research*, 98. 2 pp. 94-100.
Slavin, R.E. (1989). Class Size and Student Achievement: Small Effects of Small Classes. *Educational Psychologist*, 24. Pp 25-77.

Summary of effects	
Study	***Effect size***
Goldstein, Yang, Omar, Turner & Thompson, 2000 (correlational study)	0.20
Glass & Smith, 1978	0.01
McGiverin, Gilman & Tillitski, 1989	0.34
Slavin, 1989	0.17
Indicative effect size	**0.20**

For more information about the effect sizes in the Toolkit, click here.

Reducing class size
Toolkit references

Meta-analyses abstracts	
Study	***Abstract***
Goldstein et.al. (2000).	Meta-analysis is formulated as a special case of a multilevel (hierarchical data) model in which the highest level is that of the study and the lowest level that of an observation on an individual respondent. Studies can be combined within a single model where the responses occur at different levels of the data hierarchy and efficient estimates are obtained. An example is given from studies of class sizes and achievement in schools, where study data are available at the aggregate level in terms of overall mean values for classes of different sizes, and also at the student level.
Glass & Smith (1978).	Not available.
McGiverin et.al. (1989).	The purpose of this investigation was to examine the effects of Indiana's project Prime Time on reading and math achievement test scores of second graders who had completed 2 years of a state supported reduced-class-size program. PRIME TIME reduced class sizes in grades K-3. The results of 10 studies yielding a total of 24 comparisons (3,967 scores) of PRIME TIME (small) and pre-Prime Time large) classes were combined using Fisher's inverse chi-square procedure. Large classes averaged 26.4 students and small classes averaged 19.1 students. The results of this meta-analysis were significant at the .001 level. 10 comparisons (1,148 scores) were combined in a second meta-analysis for a control group in which class size was not reduced, and these results were not significant. The effect size for the PRIME TIME group was .34 standard deviations. This suggests that Prime Time students had higher achievement in basic skills after 2 years than did their cohorts in larger classes and indicates that primary children learn more effectively in smaller classes.
Slavin (1989).	Based on reviews by Glass, Cahen, Smith, and Filby (1982) and the Educational Research Service (1978), Cooper (this issue) concludes that substantial reductions in class size can have important effects on low-achieving students in the early grades. This article critiques these reviews and summarizes the findings of experimental studies that compared the achievement levels of elementary school students in larger classes to classes with no more than 20 students. Even in studies that made such substantial reductions, achievement differences were slight, averaging only 13% of a standard deviation. Not until class size approaches one is there evidence of meaningful effects. Based on this and other evidence, it is suggested that Chapter 1 programs provide one-to-one tutoring in reading rather than providing small-group pull-outs or reducing overall class size.

Repeating a year

Negative impact for very high cost based on extensive evidence.

-4

Pupils who do not reach a given standard of learning at the end of a year are required to repeat the year by joining a class of younger students the following academic year. Also known as “grade retention”, “non-promotion” or “failing a grade”. For students at secondary school level, repeating a year is usually limited to the particular subject or classes that a student has not passed.

Repeating a year is relatively common in the USA where the No Child Left Behind Act (2002) recommended that students be required to demonstrate a set standard of achievement before progressing to the next grade level. Students can also be required to repeat a year in some countries in Europe including Spain, France and Germany. In Finland, pupils can repeat a year in exceptional circumstances, but this decision is made collectively by teachers, parents and the student rather than on the basis of end of year testing.

How effective is it?

Evidence suggests that in the majority of cases repeating a year is harmful to a student’s chances of academic success. In addition, studies consistently show greater negative effects for students from disadvantaged backgrounds who repeat a year, suggesting that the practice of repeating a year is likely to increase educational inequality. Repeating a year is also likely to lead to greater negative effects when used in the early years of primary school and for students from ethnic minorities.

On average, students who repeat a year fall behind peers of a similar level of attainment who move on. After one year, students who repeat a year are four months behind those who move on in terms of academic achievement. In addition, studies suggest that students who repeat a year are unlikely to catch up with peers of a similar level who move on, even after completing an additional year’s schooling. Studies also suggest that students who repeat a year are more likely to drop out of school prior to completion.

After one year, students who repeat a year are four months behind those who move on in terms of academic achievement. In addition, studies suggest that students who repeat a year are unlikely to catch up with peers of a similar level who move on, even after completing an additional year’s schooling. Studies also suggest that students who repeat a year are more likely to drop out of school prior to completion.

Although the overall average impact is negative, some studies suggest that in individual circumstances students can benefit, particularly in the short term. However, it does not appear to be easy to identify which students will benefit, suggesting that repeating a year is a significant risk.

How secure is the evidence?

Overall, negative effects have been found consistently over the last fifty years in studies from Europe and North America, where much of the research has been conducted.

What are the costs?

The costs are for an additional year of schooling. In the US this was estimated at $8,916 per pupil in 2006. Annual costs of schooling vary widely in England with secondary school costs tending to fall between £4,000 and £9,000, and primary school costs between £3,000 and £8,000. Costs are therefore estimated at £6,000 per pupil per year.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Negative effects are rare for educational interventions, so the extent to which pupils who repeat a year go backwards is striking.
2. Negative effects are disproportionately greater for disadvantaged pupils, for pupils from ethnic minorities and for summer-born pupils.
3. Have you considered alternative interventions such as intensive tuition or one to one support? They are considerably cheaper and may make repeating a school year unnecessary.
4. Negative effects tend to increase with time and repeating more than one year significantly increases the risk of students dropping out of school.

Repeating a year
Toolkit references

Education Endowment Foundation

Full references
Allen, C. S., Chen, Q., Willson, V. L., & Hughes, J. N. (2009). Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. Educational Evaluation and Policy Analysis, 31(4), 480-499.
Bright, A. D. (2011). A meta-analysis">Meta-analysis of the Effects of Grade Retention of K-6 Students on Student Achievement, 1990-2010 (Doctoral dissertation, Tuscaloosa, AL: The University of Alabama.
Ehmke, T., Drechsel, B., & Carstensen, C. H. (2010). Effects of grade retention on achievement and self-concept in science and mathematics. Studies in Educational Evaluation, 36(1), 27-35.
Holmes, C. T., & Matthews, K. M. (1984). The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis. Review of Educational Research, 54(2), 225-236.
Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. Educational Evaluation and Policy Analysis, 27(3), 205-224.
Jacob, B. and Lefgren, L. (2007). The Effect of Grade Retention on High School Completion. NBER Working Paper Series.
Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. School Psychology Review, 30(3), 420-437.
Silberglitt, B., Appleton, J. J., Burns, M. K., & Jimerson, S. R. (2006). Examining the effects of grade retention on student reading performance: A longitudinal study. *Journal of School Psychology*, *44*(4), 255–270. doi:10.1016/j.jsp.2006.05.004
Uysal, S. D. (2010) : The Effect of Grade Retention on School Outcomes: An Application of Doubly Robust Estimation Methods, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Evaluation Econometrics, No. A6-V3 http://www.econstor.eu/bitstream/10419/37510/2/VfS_2010_pid_885.pdf
Warren, J. R. (2012). First-Through Eighth-Grade Retention Rates for All 50 States A New Method and Initial Results. Educational Researcher, 41(8), 320-329.
Yoshida, S. A. S. (1989). A meta-analysis of the Effects of Grade Retention on the Achievement of Elementary School Children (Doctoral dissertation, Fordham University).

Repeating a year
Toolkit references

Summary of effects	
Study	***Effect size***
Allen et al. 2009	-0.30 (low quality studies) 0.04 (medium and high quality studies)
Bright, 2011	-0.50
Jimerson, 2001	-0.31
Holmes & Matthews, 1984	-0.34
Yoshiba, 1989	-0.60
Indicative effect size	**-0.32**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Allen et.al. (2009).	The present meta-analysis examined the effect of grade retention on academic outcomes and investigated systemic sources of variability in effect sizes. Using multi-level modeling, we investigated characteristics of 207 effect sizes across 22 studies published between 1990 and 2007 at two levels: the study (between) and individual (within) levels. Design quality was a study-level variable. Individual level variables were median grade retained and median number of years post retention. Quality of design was associated with less negative effects. Studies employing middle to high methodological designs yielded effect sizes not statistically significantly different from zero and 0.34 higher (more positive) than studies with low design quality. Years post retention was negatively associated with retention effects, and this effect was stronger for studies using grade comparisons versus age comparisons. Results challenge the widely held view that retention has a negative impact on achievement. Suggestions for future research are discussed.
Bright, (2011).	This dissertation investigates the relationship between grade retention and students' academic achievement for K-6 students. A meta-analysis was conducted from studies published between 1990 and 2010 that reported data on the effects of elementary grade retention and students' academic achievement. The primary hypothesis for this dissertation was that there was a positive relationship between grade retention and students' academic performance. An extensive systematic review of the literature was conducted using bibliographic databases and other sources, resulting in the review of hundreds of abstracts and articles. Initially, this review resulted in the identification of approximately 120 articles, from which, 68 were identified as potential studies for inclusion in this meta-analysis. As data were abstracted from each potential study and evaluated, 43 studies remained for inclusion in this meta-analysis. Of these 43 studies, 31 either reported effect sizes in the results section or included sufficient data to calculate the effect sizes. After testing each study for statistical significance and eliminating insignificant studies, 26 studies remained. Effect sizes for these 26 studies were averaged and an effect size of medium strength was revealed (*ES* = -0.50). This effect size indicated that retained students scored 0.50 standard deviations lower than promoted students on academic outcome measures. Sixteen studies included in this meta-analysis had never been included in any prior meta-analysis, thus adding to the existing literature. This study found that there was not a POSITIVE relationship between grade retention and students' academic performance. Results support the findings of most prior studies on grade retention concluding that grade retention is not an effective intervention. Major findings are iii provided for the four research questions examined in this study. In addition, implications for practitioners and implications for researchers are

	included, as well as, suggestions for future research.
Holmes & Matthews (1984).	In this study data from all studies identified as meeting the selection criteria were mathematically integrated to determine the effect of grade-level retention on elementary and/or junior high school pupils. When each effect size calculated was treated equally, a grand mean effect size of -.37 was obtained indicating that, on the average, promoted children scored .37 standard deviation units higher than retained children on the various outcome measures. When the effect sizes within each study were first averaged so that each study could be given equal weight, a grand mean of -.34 was obtained. By using the effect sizes from only those studies in which the promoted and non-promoted pupils had been matched, a grand mean of -.38 was calculated. The high degree of consistency in these measures lends credibility to the validity of these findings. In addition to the grand means, effects sizes were calculated on various dependent variable measures, including academic achievement (further subdivided into various areas), personal adjustment (which included self-concept, social adjustment, and emotional adjustment), and attitude toward school, behaviour, and attendance. In all cases, the outcomes for promoted pupils were more positive than for retained pupils.
Jimerson (2001).	Retaining a child at grade level has become increasingly popular, consistent with the emphasis on accountability and standards in elementary education. This article provides a comprehensive review of the research examining the academic and socio-emotional outcomes associated with grade retention. Following a brief historical overview of previously published literature reviews, a summary of studies published between 1990 and 1999 is provided. A systematic review and meta-analysis of 20 recent studies includes: outcome variables (i.e., achievement and socio-emotional adjustment), age or grade of retained population, matched or controlled for variables in analyses with comparison groups, and the overall conclusion regarding the efficacy of grade retention. Results of recent studies and this meta-analysis are consistent with past literature reviews from the 1970s and 1980s. In addition to a summary of the results, the discussion addresses the disparity between educational practice and converging research regarding grade retention and suggests directions for practice. This review encourages researchers, educational professionals, and legislators to abandon the debate regarding social promotion and grade retention in favour of a more productive course of action in the new millennium.

School uniform

Very low or no impact for very low cost, based on very limited evidence.

0

Schools identify clothing considered appropriate for pupils to wear in school, usually including style and colour.

How effective is it?

There is a general belief in many countries that school uniform supports the development of a whole school ethos and therefore is supportive of discipline and motivation. However, there is no robust evidence that introducing a school uniform will, by itself, improve academic performance, behaviour or attendance. There are studies that have information about these outcomes linked to the introduction of a school uniform policy, but uniform was usually one factor amongst other improvement measures, such as changes in behaviour policy or other teaching and learning developments.

How secure is the evidence?

One of the problems in interpreting the evidence is that schools in challenging circumstances often choose a school uniform policy as part of a broader range of improvement measures. There are no systematic reviews of well-controlled interventions of a school uniform policy. The evidence rests mainly on correlational studies that have compared the performance of schools with uniforms to those without, or that have examined a school's trajectory of improvement after the introduction of school uniform. The most rigorous reviews and analyses have so far been unable to establish a causal link, but speculate that adoption of a uniform policy may provide a symbolic and public commitment to school improvement.

There are cultural issues about how a school uniform is perceived which play an important role in determining the acceptability and success (in terms of compliance). There is some evidence that in areas of very high poverty free school uniforms improve attendance, however this does not appear to be true in all areas. In other cultures school uniforms are associated with regulation and the loss of individuality, so care must be taken in generalising from studies from abroad.

What are the costs?

The costs associated with introducing a school uniform are very low and mainly depend on parents buying the clothes instead of others the child would wear.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Wearing a uniform is not, on its own, going to improve learning, but where it is combined with the development of a school ethos and the improvement of behaviour and discipline, it can be successfully included as part of this process.
2. Improved behaviour, on its own, does not necessarily lead to better learning, though it may be an important precondition (see Behaviour).
3. The commitment of staff to uphold and enforce a uniform policy is crucial to its success. Have you considered how you will implement a uniform policy as part of a wider school policy?

School uniform
Toolkit references

Full references
Brunsma, D.L. & Rockquemore, K. (1998). Examining the effects of student uniforms on attendance, substance abuse, disciplinary behavior problems, academic achievement. Journal of Educational Research 92 pp 53-62.
Brunsma, D.L. & Rockquemore, K. (2003). Statistics, sound bites and school uniforms: a reply to Bodine. Journal of Educational Research 97.2 pp 72-77.
Reynolds, B.A. (2004). An analysis of the evidence produced by quantitative research on the effects of school uniforms on student discipline and academic achievement PhD thesis submitted to the faculty of Brigham Young University, Salt Lake City, Utah (ProQuest Dissertations).
Samuels, K.S. (2003). The relationship of school uniforms to students' achievement, attendance, discipline referrals and perceptions: an analysis of one school district. PhD dissertation University of Alabama at Birmingham (ProQuest Dissertations and Theses).
Sowell, R. E. (2012). The relationship of school uniforms to student attendance, achievement, and discipline (Doctoral dissertation, Liberty University: ProQuest Dissertations and Theses).

Summary of effects	
Study	***Effect size***
Samuels, 2003 (language arts)	0.03
Samuels, 2003 (mathematics)	-0.06
Sowell, 2012 (single study)	0.02
Indicative effect size	**0.00**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Samuels, (2003).	The purpose of this study was to examine school uniforms and how they affect high school students' achievement, attendance, discipline referrals, and perceptions based on grade level and gender regarding the mandatory school uniform policy in the Birmingham, Alabama; City Schools (BCS). BCS students (Grades 9-12) comprised the population for this study. Instrumentation and materials' source of data wire archival records of selected students secured from the central office's testing centre. Three out of five hypotheses were tested using the following source of data: (a) Stanford Achievement Test-8/9 (SAT-8/9) Normal Curve Equivalent (NCE) scores relative to Hypothesis 1, (b) the number of suspensions and expulsions relative to Hypothesis 2, and (c) students' average daily attendance relative to Hypothesis 3. Data for Hypotheses 4 and 5 were obtained using a researcher-developed survey relative to perceptions of students based on grade level and gender regarding the school uniform policy in BCS. The participants were asked to respond to the three-section survey instrument that included Section I, information provided by respondent about grade level, gender, and name of school. Section II, which contained 13 Likert-type scale items that provided data regarding the usefulness of the policy in curbing violence and improving students' behaviour; and the effectiveness of the policy in helping students to be better students. Section III, the final section, solicited respondents' general comments about BCS mandatory uniform policy. The results of this study revealed significant change in the high school students' achievement during the selected years (1995-1998). Discipline referrals during selected years (1994-1999) decreased, and the average daily attendance during selected years

	(1994-1998) increased. There was no significant difference between the high school students' perceptions based on grade level and gender regarding BCS mandatory uniform policy.
Sowell, (2012).	This causal-comparative study examined the relationship of school uniforms to attendance, academic achievement, and discipline referral rates, using data collected from two high schools in rural southwest Georgia county school systems, one with a uniforms program and one without a uniforms program. After accounting for race and students with disabilities status, School A (with uniforms) had significantly better attendance and somewhat fewer minor behaviour infractions, but trended lower in standardized math scores and more intermediate and major behavioural infractions than School B (without uniforms). These findings failed to demonstrate an unambiguous advantage of school uniforms, consistent with the mixed results across reports in the published literature. Implications and suggestions for further research are detailed.

Setting or streaming

Negative impact for very low or no cost, based on moderate evidence.

-1

Pupils with similar levels of current attainment are grouped together either for specific lessons on a regular basis (setting or regrouping), or as a whole class (streaming or tracking). The assumption is that it will be possible to teach more effectively or more efficiently with a narrower range of attainment in a class.

How effective is it?

Overall, setting or streaming appears to benefit higher attaining pupils and be detrimental to the learning of mid-range and lower attaining learners. On average, setting or streaming does not appear to be an effective strategy for raising the attainment of disadvantaged pupils, who are more likely to be assigned to lower groups.

On average, studies show that higher attaining learners make between one and two additional months' progress when set or streamed compared to when taught in mixed ability groups. Studies of targeted interventions for pupils identified as "gifted and talented" are consistent with this finding. They show that high attaining pupils benefit from a range of different kinds of grouping, including pull-out classes, accelerated classes and promotion (where high attaining pupils move up a year). The effects of these programmes potentially provide an advantage for these pupils of 3 to 12 months' additional learning. However, research into gifted and talented schemes rarely records the impact of the schemes on the students not identified as gifted and talented, who are more likely to be from disadvantaged backgrounds. Low attaining learners fall behind by one or two months a year, on average, when compared with the progress of similar students in classes with mixed ability groups. It appears likely that routine setting or streaming arrangements undermine low attainers' confidence and discourage the belief that attainment can be improved through effort. Research also suggests that ability grouping can have a longer term negative effect on the attitudes and engagement of low attaining pupils. It should be noted that there are some exceptions to this average, where setting or streaming has benefitted all learners. Further study could be undertaken to understand what happened differently in these examples.

Though the average impact of setting or streaming on low attaining pupils is negative, evidence suggests that certain types of grouping are more effective than others. Some studies have shown that reducing the size of the lowest attaining groups and assigning high-performing teachers to these groups can be effective, as can providing additional targeted catch up support.

How secure is the evidence?

The evidence on setting and streaming is fairly consistent and has accumulated over at least 30 years of research. Although there is some variation depending on methods and research design, conclusions on the impact of ability grouping are relatively consistent.

The majority of the evidence comes from the USA, and there are few rigorous UK studies. In a 2014 evaluation Year 6 and 7 pupils made an additional three months' progress from a structured programme of small group tuition. A 2011 evaluation of Every Child Counts found that the programme had a positive impact when delivered on a one to one basis, or with groups of two or three, with all group sizes making similar amounts of progress. These findings suggest that small group tuition can be effective in English schools. More research could be undertaken to understand the impact of different group sizes, and to explore whether these effects vary by subject or age of pupils.

What are the costs?

Setting or streaming are organisational strategies that have few, if any, associated financial costs. Additional resources may be needed to support different groups. Overall the costs are estimated as very low.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Flexible within-class grouping is preferable to tracking or streaming for low attaining pupils.
2. It is important to recognise that a measure of current attainment, such as a test, is not the same as a measure of potential.
3. Have you considered how the differences in grouping will enable more effective teaching for all pupils, including lower attaining pupils?
4. How will you monitor the impact of ability grouping on pupils' attitudes to learning and their engagement?

Setting or streaming
Toolkit references

Full references
Attfield., R. (2009). Developing a Gifted and Talented Strategy: Lessons from the UK experience Reading: CfBT.
Boaler, J. (2008). Promoting 'relational equity' and high mathematics achievement through an innovative mixed-ability approach. *British Educational Research Journal* 34.2 pp 167-194.
Collins, C. A., & Gan, L. (2013). Does Sorting Students Improve Scores? An Analysis of Class Composition (No. w18848). Cambridge, MA: National Bureau of Economic Research. http://www.nber.org/papers/w18848.
Dunne, M., Humphreys, S., Dyson, A., Sebba, J., Gallannaugh, F., & Muijs, D. (2011). The teaching and learning of pupils in low-attainment sets. Curriculum Journal, 22(4), 485-513.
Duflo, E., Dupas, P., Kremer, M. (2011). "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya". *American Economic Review* 101 (5): pp 1739-1774.
Hallam, S., & Ireson, J. (2007). Secondary school pupils' satisfaction with their ability grouping placements. British Educational Research Journal, 33(1), 27-45.
Hanushek, E. A. & Woessmann, L. (2005) Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries, CESifo working papers, No. 1415, http://hdl.handle.net/10419/18779.
Ireson, J., Hallam, S. & Plewis, I. (2001). Ability grouping in secondary schools: Effects on pupils' self-concepts *British Journal of Educational Psychology* 71. 2, pp 315-326.
Ireson, J., Hallam, S., Mortimore, P., Hack, S., Clark, H. & Plewis, I. (1999). Ability grouping in the secondary school: the effects on academic achievement and pupils' self-esteem Paper presented at the *British Educational Research Association Annual Conference*, University of Sussex at Brighton, September 2-5 1999.
Kulik C-L.C & Kulik J.A. (1982). Effects of Ability Grouping on Secondary School Students: A Meta-Analysis of Evaluation Findings, *American Educational Research Journal, 19* (3), 415-428.
Kulik C-L.C & Kulik J.A. (1984). Effects of Ability Grouping on Elementary School Pupils: A Meta-Analysis. *Annual Meeting of the American Psychological Association.*
Kulik, J.A., & Kulik, C.L.C. (1987). Effects of ability grouping on student achievement. Equity and Excellence in Education, 23(1-2), 22-30.
Kulik, J.A. (1992). *An Analysis of the Research on Ability Grouping: Historical and Contemporary Perspectives* The National Research Center On The Gifted And Talented.
Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research*, *66*(4), 423-458..
Marks, R (2013) 'The Blue Table Means You Don't Have a Clue': the persistence of fixed-ability thinking and practices in primary mathematics in English schools, FORUM, 55(1), 31-44. http://dx.doi.org/10.2304/forum.2013.55.1.31

Summary of effects	
Study	***Effect size***
Gutierez & Slavin, 1992 (mixed age attainment/non-graded classes)	0.34
Kulik & Kulik 1982 (on secondary pupils)	0.1
Kulik & Kulik, 1984 (on elementary/primary pupils)	0.1
Lou et al., 1996 (on low attainers)	-0.12
Puzio & Colby, 2010	0.22
Slavin, 1990 (on low attainers)	-0.6
Weighted mean effect size	**-0.09 (on low attainers)**

For more information about the effect sizes in the Toolkit, click here.

Setting or streaming
Toolkit references

Meta-analyses abstracts	
Study	***Abstract***
Gutierez & Slavin, 1992	This article reviews research on the effect of ability grouping on the achievement of secondary students. Six randomized experiments, 9 matched experiments and 14 correlational studies compared ability grouping to heterogeneous plans over periods of from one semester to 5 years. Overall, achievement effects were found to be essentially zero at all grade levels although there is much more evidence regarding Grades 7-9 and 10-12. Results were similar for all subjects except social studies, for which there was a trend favouring heterogeneous placement. Results were close to zero for students of all levels of prior performance. This finding contrasts with those of studies comparing the achievement of students in different tracks, which generally find positive effects of ability grouping for high achievers and negative effects for low achievers, and these contrasting findings are reconciled. A nongraded elementary program is one in which children are flexibly grouped according to performance level, not age, and proceed through the elementary school at their own rates. Popular in the 1950s, 1960s, and early 1970s, the nongraded plan is returning today. This article reviews research on the achievement effects of nongraded organization. Results indicated consistent positive achievement effects of simple forms of nongrading generally developed early: cross-grade grouping for one subject (me-dian ES = + .46) and cross-grade grouping for many subjects (median ES = + .34). Forms of nongrading making extensive use of individualization were less consistently successful (median ES = +.02). Studies of Individually Guided Education (IGE), which used nongrading and individualization, also produced inconsistent effects (median ES = + .11). The article concludes that nongraded organization can have a positive impact on student achievement if cross-age grouping is used to allow teachers to provide more direct instruction to students but not if it is used as a framework for individualized instruction.
Sowell, (2012).	This causal-comparative study examined the relationship of school uniforms to attendance, academic achievement, and discipline referral rates, using data collected from two high schools in rural southwest Georgia county school systems, one with a uniforms program and one without a uniforms program. After accounting for race and students with disabilities status, School A (with uniforms) had significantly better attendance and somewhat fewer minor behaviour infractions, but trended lower in standardized math scores and more intermediate and major behavioural infractions than School B (without uniforms). These findings failed to demonstrate an unambiguous advantage of school uniforms, consistent with the mixed results across reports in the published literature. Implications and suggestions for further research are detailed.
Kulik & Kulik, 1982	This article reports results from a meta-analysis of findings from 52 studies of ability grouping carried out in secondary schools. In the typical study the benefits of ability grouping were small but significant on achievement examinations-an average increase of one tenth standard deviations on examination scores, or an increase from the 50th to the 54th percentile for the typical student in a grouped class. The size of achievement effect differed in different types of studies of grouping however. Studies in which high ability students received enriched instruction in honours classes produced especially clear effects, for example, while studies of average and below average students produced near-zero effects. The benefits of grouping were also clear in the area of student attitudes towards the subjects they were studying than did students in ungrouped classes.
Kulik & Kulik, 1984	A meta-analysis of finding from 31 separate studies showed that ability grouping has significant positive effects on the academic performance of elementary school children. The benefits of grouping tended to be small in the typical study of achievement-an increase from the 50th to the 58th percentile for the typical student in a grouped class. One subgroup of studies however produced especially clear effects. In this type of study students of high ability or gifted students were

	put into special classes in which they received enriched instruction. Studies of this type usually reported significant results and usually reported effects on achievement were moderate in size. meta-analysis">Meta-analysis also showed that ability grouping has trivially small effects on the self-concepts of elementary school pupils.
Kulik & Kulik, 1987	In this article, we analyse results from a wider variety and greater number of studies than were covered in our earlier meta-analyses. Among the major questions that we try to answer are the following: What are the effects of ability grouping in the typical study? Does grouping have different effects on different types of students—e.g., students of high, average, and low aptitude? Do different approaches to grouping produce different outcomes? Why have different meta-analysts reached different conclusions about grouping? The average effect size was 0.15 standard deviations in the elementary studies and 0.07 standard deviations in the secondary school studies. Special within-class grouping programs designed especially for the benefit of talented students raised their achievement scores by 0.62 standard deviations. Special between-class grouping programs for these students raised their scores by 0.33 standard deviations. Separating talented students into homogeneous groups apparently enabled teachers to provide learning opportunities for the students that were unavailable in more heterogeneous groups. Programs that were designed for all students in a grade—not solely for the benefit of talented learners—had significantly lower effects. Comprehensive between-class grouping raised overall achievement levels by only 0.06 standard deviations, a very small amount. Comprehensive within-class grouping raised overall achievement levels by 0.17 standard deviations.
Kulik & Kulik, 1992	Meta-analytic reviews have focused on five distinct instructional programs that separate students by ability: multi-level classes, cross-grade programs, within-class grouping, enriched classes for the gifted and talented and accelerated classes. The review shows that effects are a function of program type. Multilevel classes which entail only minor adjustments of course content for ability groups, usually have little or no effect on student achievement. Programs that entail more substantial adjustment of curriculum to ability such as cross-grade and within-class programs produce clear positive effects. Programs of enrichment and acceleration which usually involve the greatest amount of curricular adjustment have the largest effect on student learning. These results do not support recent claims that no one benefits from grouping or that students in the lower groups are harmed academically and emotionally from grouping.
Lou et.al., 1990	The effects of within-class grouping on student achievement and other outcomes were quantitatively integrated using two sets of study findings. The first set included 145 effect sizes and explored the effects of grouping versus no grouping on several outcomes. Overall, the average achievement effect size was +0.17, favoring small-group learning. The second set included 20 effect sizes which directly compared the achievement effects of homogeneous versus heterogeneous ability grouping. Overall, the results favored homogeneous grouping; the average effect size was +0.12. The variability in both sets of study findings was heterogeneous, and the effects were explored further. To be maximally effective, within-class grouping practices require the adaptation of instruction methods and materials for small-group learning.
Puzio & Colby, 2010	Although some literacy researchers consider grouping students for reading instruction to be a proven educational practice, the support for this belief is lacking from a research synthesis perspective. With this idea in mind, Slavin comments in the middle of his review on the effects of grouping: "there is not enough research on within-class ability grouping in reading to permit any conclusions" (Slavin, 1987, p. 320). Because of this, the question of whether "and" how to group students is often cast and answered ideologically rather than empirically. This review attempts to see if the empirical research available can answer either or both of these

	questions. It will contribute to the literacy field by focusing specifically on reading outcomes for classroom teachers, who instruct a wide variety of students. Informed by previous research on within-class grouping, the following three research questions guide the present study: (1) To what extent does within-class grouping impact student achievement in reading?; (2) For which grade(s) or which students is within-class grouping most or least beneficial?; and (3) Do any moderators, especially those identified by previous research (measurement source, teacher development, and grouping type), help explain this effect?
Rogers, 2007	This article discusses five reconsiderations (lessons) the research on the education of the gifted and talented suggests. Although several of the considerations derive from traditional practice in the field, some reconsideration is warranted because of more currently researched differences in how the gifted learner intellectually functions. It is argued that thinking of the gifted learner as idiosyncratic, not necessarily one of many classified as “the gifted,” requires a reconceptualization of how to appropriately and fully serve this unique learner. The research synthesized here covers the period from 1861 to present and represents the entire body of published research studies and representative literature (theory, program descriptions, and persuasive essays). Implications for service development and implementation are also discussed.
Slavin, 1990	This article reviews research on the effect of ability grouping on the achievement of secondary students. Six randomized experiments, 9 matched experiments and 14 correlational studies compared ability grouping to heterogeneous plans over periods of from one semester to 5 years. Overall, achievement effects were found to be essentially zero at all grade levels although there is much more evidence regarding Grades 7-9 and 10-12. Results were similar for all subjects except social studies, for which there was a trend favouring heterogeneous placement. Results were close to zero for students of all levels of prior performance. This finding contrasts with those of studies comparing the achievement of students in different tracks, which generally find positive effects of ability grouping for high achievers and negative effects for low achievers, and these contrasting findings are reconciled.
Steenbergen-Hu & Moon, 2011	Current empirical research about the effects of acceleration on high-ability learners’ academic achievement and social– emotional development were synthesized using meta-analytic techniques. A total of 38 primary studies conducted between 1984 and 2008 were included. The results were broken down by developmental level (P-12 and postsecondary) and comparison group (whether the accelerants were compared with same-age, older, or mixed-age peers). The findings are consistent with the conclusions from previous meta-analytic studies, suggesting that acceleration had a positive impact on high-ability learners’ academic achievement (g = 0.180, 95% CI = -.072, .431, under a random-effects model). In addition, the social–emotional development effects appeared to be slightly positive (g = 0.076, 95% CI = -.025, .176, under a random effects model), although not as strong as for academic achievement. No strong evidence regarding the moderators of the effects was found.
Vaughn et.al. 1991	The purpose of this research was to evaluate the effectiveness of pull-out programs in gifted education. Nine experimental studies were located that dealt with pull-out programs for gifted students. The variables of self-concept, achievement, critical thinking, and creativity were quantified via meta-analysis. The results indicate that pull-out models in gifted education have significant positive effects for the variables of achievement, critical thinking, and creativity. However, gifted students' self- concepts were not affected by the pull-out programs

Small group tuition

Moderate impact for moderate cost, based on limited evidence.

+4

Small group tuition is defined as one teacher or professional educator working with two, three, four, or five pupils. This arrangement enables the teacher to focus exclusively on a small number of learners, usually on their own in a separate classroom or working area. Intensive tuition in small groups is often provided to support lower attaining learners or those who are falling behind, but it can also be used as a more general strategy to ensure effective progress, or to teach challenging topics or skills.

How effective is it?

Overall, the pattern is that small group tuition is effective and, as a rule of thumb, the smaller the group the better, e.g. groups of two have slightly higher impact than groups of three, but slightly lower impact compared to one to one tuition. Some studies suggest that greater feedback from the teacher, more sustained engagement in smaller groups, or work which is more closely matched to learners' needs explains this impact. Once group size increases above six or seven there is a noticeable reduction in effectiveness.

However, although the above pattern is usually consistent, there is some variability in impact within the existing evidence. For example, in reading, small group teaching can sometimes be more effective than either one to one or paired tuition. It may be that in these cases reading practice can be efficiently organised so that all the group stay fully engaged as each take their turn, such as in Guided Reading. The variability in findings suggests two things. First, the quality of the teaching in small groups may be as or more important than the group size, and there is evidence of the benefits of professional development on pupil outcomes. Second, it is important to evaluate the effectiveness of different arrangements as the specific subject matter being taught and composition of the groups may influence outcomes.

Given this uncertainty and the lower cost of small group tuition, it may be sensible to trial small group tuition as an initial option, before moving to one to one tuition if small group tuition is ineffective.

How secure is the evidence?

The evidence is limited and mainly relates to low attaining pupils receiving additional support to catch up with their peers. More research has been undertaken into paired tuition than other kinds of small group tuition, so the evidence for small group teaching across varying sizes of groups and at different levels of intensity is not conclusive and mainly comes from single studies. There are very few studies where group size has been varied systematically to explore the effects beyond one-to-two and one-to-three, so more research would be useful in this area.

To date the majority of the evidence comes from the USA. However, in recent years a growing number of rigorous UK studies have been conducted. In a 2014 evaluation Year 6 and 7 pupils made an additional three months' progress from Switch-on Reading, a structured programme involving small group tuition. In addition, an intensive coaching programme that involved one to one and small group tuition has an average impact of five additional months' progress.

A 2011 evaluation of Every Child Counts also found that the programme had a positive impact when delivered on a one to one basis, or with groups of two or three, with all group sizes making similar amounts of progress.

What are the costs?

Overall, costs are estimated as moderate. Costs decrease with group size, as the majority of the costs are for teaching time. The cost of paired tuition is approximated as £350 per pupil per term (based on two pupils receiving 30 minutes tuition, five times a week for 12 weeks) plus any resource or equipment costs, with one to three cheaper still (£233 per pupil).

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Small group tuition is most likely to be effective if it is targeted at pupils' specific needs. How will you assess pupils' needs accurately before adopting a new approach?
2. One to one tuition and small group tuition are both effective interventions. However, the cost effectiveness of one to two and one to three indicates that greater use of these approaches may be worthwhile. Have you considered trying one to two or one to three as an initial option?
3. Training and support are likely to increase the effectiveness of small group tuition. Have those leading the small group tuition been trained in the programme they are delivering?

Technical Appendix: Small group tuition

Definition

Small group tuition is where one teacher or professional educator works directly with a small group (defined as two to five pupils). This arrangement enables the teacher to focus exclusively on a small number of learners, usually in a separate classroom or working area. Intensive tuition in small groups is often provided to support lower attaining learners or to help those who are falling behind to catch up, but it can also be used as a more general strategy to ensure effective progress, or for teaching challenging topics or skills. It is distinguished from One to One Tuition by group size (i.e. more than one and less than six in the group of learners) and from group work in Collaborative learning where there the teacher has responsibility for managing several groups in a whole class setting.

Search terms:

Small group teaching/ tuition/ instruction; dyad/paired teaching

Evidence rating

Limited: Although there are three meta-analyses and one best evidence synthesis, two of these reviews consider group work in relation to other strategies (teaching pupils with reading difficulties and the use of technology) and one calculates effect sizes for cost/benefit estimates. Only one meta-analysis was conducted in the last ten years. None of the reviews have the impact of group work on academic attainment as the main focus of the review. Overall the evidence is therefore rated as limited.

Additional cost information

30 minutes tuition, five times a week for 12 weeks is equivalent to four days of a teacher's time. The average salary of a a full-time qualified teacher is £34,600 a year (source: https://getintoteaching.education.gov.uk/competitive-salary-and-great-benefits). There are 195 days in the school year. This means that the average cost of four days' of teachers' time is approximately £700. If this is split between two pupils then the per-pupil cost is £350.

References

Full references
Buckingham, J., Wheldall, K., & Beaman, R. (2012). A randomised control trial of a Tier-2 small-group intervention ('MiniLit') for young struggling readers 1. Australian Journal of Learning Difficulties, 17(2), 79-99.
*Elbaum, B., Vaughn, S., Hughes, M. T., Moody, S. W., & Schumm, J. S. (2000). *How reading outcomes of students with disabilities are related to instructional grouping formats: A meta-analytic review*. In R. Gersten, E.P. Schiller & S. Vaughn (Eds) Contemporary special education research: Syntheses of the knowledge base on critical instructional issues, 105-135. Mahwah, NJ: Lawrence Erlbaum Associates.
Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice*, 16(4), 203-212. http://www.dx.doi.org/10.1111/0938-8982.00020
*Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, 71(3), 449-521. http://www.dx.doi.org/10.3102/00346543071003449

Schwartz, R. M., Schmitt, M. C., & Lose, M. K. (2012). Effects of Teacher-Student Ratio in Response to Intervention Approaches. *Elementary School Journal*, 112(4), 547-567. http://www.jstor.org/stable/10.1086/664490
*Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1-26. http://www.dx.doi.org/10.1016/j.edurev.2010.07.002
Torgerson. D., Torgerson, C., Mitchell, N., Buckley, H., Ainsworth, H., Heaps, C. & Jefferson, L. (2014) *Grammar for Writing Evaluation Report and Executive Summary February 2014*. London: EEF https://educationendowmentfoundation.org.uk/uploads/pdf/FINAL_EEF_Evaluation_Report_-_Grammar_for_Writing_-_February_2014.pdf
Torgerson, C. J., Wiggins, A., Torgerson, D. J., Ainsworth, H., Barmby, H., Hewitt, C., Jones, K., Hendry, V., Askew, M., Bland, M., Coe, R., Higgins, S., Hodgen, J., Hulme, C.& Tymms, P. (2011). *Every child counts: the independent evaluation.* DfE Research Report DFE-RR091a London DfE. http://dera.ioe.ac.uk/2376/1/2376_DFE-RR091A.pdf
Vaughn, S., Linan-Thompson, S., Kouzekanani, K., Bryant, D. P., Dickson, S., & Blozis, S. A. (2003). Reading instruction grouping for students with reading difficulties. *Remedial and Special Education*, 24(5), 301-315. http://www.dx.doi.org/10.1177/07419325030240050501
Vaughn, S., Wanzek, J., Wexler, J., Barth, A., Cirino, P. T., Fletcher, J. & Francis, D. (2010). The relative effects of group size on reading progress of older students with reading difficulties. *Reading and Writing*, 23(8), 931-956. http://www.dx.doi.org/10.1007/s11145-009-9183-9
*Washington State Institute for Public Policy (2014) Tutoring: By certificated teachers, small-group, structured Seattle, Wa: WSIPP http://www.wsipp.wa.gov/BenefitCost/ProgramPdf/351/Tutoring-By-certificated-teachers-small-group-structured

Summary of effects		
Meta-analyses	***FSM effect size***	***Effect size***
Elbaum et al. 2000		0.40 (pairs) 1.61 (small group – NB only one study)
Lou et al. 2001 (with digital technology)		0.16 (individual) (CI 0.12 to 0.20) 0.31 (small group) (CI 0.20 to 0.43) 0.08 (pairs compared with groups of 3-5)
Slavin et al. 2011		0.31
Washington State Institute, 2014		0.27
Recent Studies		
Buckingham et al. 2012		0.36 (Word reading)
Fuchs et al. 2008		1.18 (validated instruction in groups) 1.13 (group instruction non-validated)
NFER, 2015a EEF Primary Tutor Trust	-0.05 -0.02	-0.08 (English) -0.04 (maths)
NFER, 2015b EEF Secondary Tutor Trust	0.09	0.05 (maths)
Torgerson et al. 2014 (EEF Grammar for Writing)		0.24 small group intervention vs. control 0.21 small group vs. whole group (0.06 class level intervention vs. control *without* small group)

Technical Appendix: Small group tuition

Weighted mean effect size		**0.31**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses and abstracts	
Elbaum et al. 2000	This meta-analysis examines the link between reading outcomes and grouping format. The following grouping formats were investigated: Cooperative Learning: mixed ability groups work together on class assignments; Student Pairing: students work together in groups of 2; Peer-Tutoring: Originally peer-tutoring usually meant an older or higher ability student tutoring a younger, disabled, or low-achieving student, but studies have shown that children with disabilities benefit from being the tutor as well; Reciprocal tutoring: where students take turns leading the group, has also shown to be effective for both regular and special education students; Small-Group Instruction: this within class grouping practice can be done either with homogeneous or heterogeneous ability groups that are either led by the students or the teacher; Multiple-Grouping Formats: many classrooms use a variety of the grouping formats listed above instead of limiting themselves to just one. Findings: Grouping children instead of teaching the whole class at once significantly improves the reading performance of students with disabilities. There are not many studies on the effect of small-group instruction for teaching reading to students with LD, but the studies that exist imply that groups of 4 or fewer are better than larger groups, and reciprocal teaching is an effective strategy to use in small-group reading instruction. Being paired with another student was beneficial for students with disabilities regardless of whether the student with a disability was in the role of the tutee or acting as reciprocal tutor. Cross-age tutoring has a positive impact on older students with disabilities who tutor younger students. However, younger students with disabilities do not benefit from being tutored by older students with disabilities. The average effect size for all grouping formats used in the reading instruction of students with LD was 0.43.
Lou et al. 2001	*This study quantitatively synthesized the empirical research on the effects of social context (i.e. small group versus individual learning) when students learn using computer technology. In total, 486 independent findings were extracted from 122 studies involving 11,317 learners. The results indicate that, on average, small group learning had significantly more positive effects than individual learning on student individual achievement (mean ES = +0.15), group task performance (mean ES = +0.31), and several process and affective outcomes. However, findings on both individual achievement and group task performance were significantly heterogeneous. Through weighted least squares univariate and multiple regression analyses, we found that variability in each of the two cognitive outcomes could be accounted for by a few technology, task, grouping, and learner characteristics in the studies.*
Slavin et al. 2003	This article reviews research on the achievement outcomes of alternative approaches for struggling readers ages 5–10 (US grades K-5): One-to-one tutoring, small-group tutorials, classroom instructional process approaches, and computer-assisted instruction. Study inclusion criteria included use of randomized or well-matched control groups, study duration of at least 12 weeks, and use of valid measures independent of treatments. A total of 97 studies met these criteria. The review concludes that one-to-one tutoring is very effective in improving reading performance. Tutoring models that focus on phonics obtain

	much better outcomes than others. Teachers are more effective than paraprofessionals and volunteers as tutors. Small-group, phonetic tutorials can be effective, but are not as effective as one-to-one phonetically focused tutoring. Classroom instructional process programs, especially cooperative learning, can have very positive effects for struggling readers. Computer-assisted instruction had few effects on reading. Taken together, the findings support a strong focus on improving classroom instruction and then providing one-to-one, phonetic tutoring to students who continue to experience difficulties.
Washington State Institute 2014	The small-group tutoring programs included in this analysis are structured, systematic approaches to tutoring struggling students in specific English language arts and/or mathematics skills. The evaluated programs include a variety of specific approaches and curricula such as (in no particular order) Read Aloud, Proactive Reading, Responsive Reading, Leveled Literacy, Spell Read, Corrective Reading, and Number Rockets. An average program provides about 40 hours of tutoring time to groups of two to six (usually three) early elementary students. Certificated teachers provide tutoring and receive about 35 hours of training with a focus on the specific content and strategies used in the programs.

Social and emotional learning

+4

Moderate impact for moderate cost, based on extensive evidence.

Interventions which target social and emotional learning (SEL) seek to improve attainment by improving the social and emotional dimensions of learning, as opposed to focusing directly on the academic or cognitive elements of learning. SEL interventions might focus on the ways in which students work with (and alongside) their peers, teachers, family or community. Three broad categories of SEL interventions can be identified: 1. Universal programmes which generally take place in the classroom; 2. More specialised programmes which are targeted at students with particular social or emotional problems; 3. School-level approaches to developing a positive school ethos which also aim to support greater engagement in learning.

In 2005, a national SEL programme was introduced to support effective learning, positive behaviour, attendance and emotional well-being, first in primary schools and then in secondary schools.

How effective is it?

On average, SEL interventions have an identifiable and significant impact on attitudes to learning, social relationships in school, and attainment itself (four months' additional progress on average).

Although SEL interventions almost always improve emotional or attitudinal outcomes, not all interventions are equally effective at raising attainment. Improvements appear more likely when approaches are embedded into routine educational practices, and supported by professional development and training for staff. In addition, the implementation of the programme and the degree to which teachers are committed to the approach appear to be important.

SEL programmes appear to benefit disadvantaged or low-attaining pupils more than other pupils, though all pupils benefit on average. Approaches have been found to be effective from nursery to secondary school.

How secure is the evidence?

There is extensive research in this area and a number of meta-analyses, though more research has been undertaken with younger children in primary, than in secondary schools, and more studies have evaluated the impact on disadvantaged or low attaining pupils.

In England, a number of studies have identified a link between SEL interventions and academic outcomes. However, evidence from the nationwide SEL programme introduced in 2005 does underline the fact that benefits on learning will not be automatically achieved and that the quality of engagement with the SEL approach is likely to affect outcomes. A quasi-experimental evaluation of the impact of the secondary programme did not find a secure impact on attainment in the SEL schools. Similarly, a 2015 randomised controlled trial of an SEL intervention, Promoting Alternative Thinking Strategies, found that it did not have a positive impact on academic attainment.

What are the costs?

Universal approaches that encourage social and emotional learning throughout a school will benefit from professional development and may require new materials and resources, but these costs are likely to be very low. Social and emotional strategies targeted at specific individuals will be much more expensive. Estimates from the US suggest targeted programs cost about £2,800 per child per year and involve professional counselling or psychological services. On average, the costs per child are estimated as moderate.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Skills should be taught purposefully and explicitly linked to direct learning in schools, encouraging pupils to apply the skills they learn.
2. Teachers and other school staff can effectively support these approaches, particularly with appropriate professional development
3. How will you ensure that staff commit to supporting the programme and consistently apply the skills more widely in school?
4. Sensitive and targeted intervention may benefit at risk or more vulnerable pupils.
5. The impact on attainment of social and emotional aspects of learning is not found consistently, so it is important to evaluate the impact of any initiative. Have you considered how you will evaluate the impact of these approaches?

Technical Appendix: Social and emotional learning

Definition

Interventions and approaches which focus on social and emotional learning (SEL) aim to improve attainment by improving pupils' interaction with others and self-management of emotions, as opposed to focusing directly on academic or cognitive skills and knowledge. It is sometimes characterised as a process for learning broader life skills, such as how to deal with oneself, others and relationships, and to be able to work in an effective manner. SEL may focus on recognizing one's own emotions and learning how to manage those feelings. In dealing with others, SEL helps with developing sympathy and empathy for others, and maintaining positive relationships. SEL interventions might focus on the ways in which students work with (and alongside) their peers, teachers, family or community. Three broad categories of interventions can be identified: 1. Universal programmes which seek to improve participation and engagement and generally take place in the classroom involving all pupils; 2. More specialised programmes which are targeted at students with either behavioural issues or behaviour and academic problems (see also Behaviour interventions); 3. School level approaches to developing a positive school ethos or improving discipline which also aim to support greater engagement in learning. Some programmes include aspects of Self-regulated learning and often include Collaborative learning approaches and techniques.

Search terms:

Social and emotional learning, SEAL/SEL interventions; social skills, skills-for-life, self-esteem, empathy, emotional intelligence

Evidence rating

There are three meta-analyses of evaluations of interventions mainly undertaken in schools using pupil attainment data to assess impact, with some exploration of causes of any identified heterogeneity. All three were conducted in the last 10 years. However the quality of the underlying studies varies considerably and there are few well-controlled trials, particularly at scale. Overall the evidence is extensive, but not very consistent or very high quality.

Additional cost information

The main financial cost of implementing a whole-school social and emotional learning approach will be the cost of professional development. The average cost of professional development in EEF-funded programmes is well under £80 per pupil.

Targeted programmes are likely to carry much higher costs

References

Full references
Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). Benefits and costs of prevention and early intervention programs for youth (No. 04-07, p. 3901). Olympia, WA: Washington State Institute for Public Policy. http://courses.washington.edu/pbaf513m/prevention%20tech%20appendix.pdf
Bavarian, N., Lewis, K. M., DuBois, D. L., Acock, A., Vuchinich, S., Silverthorn, N., ... & Flay, B. R. (2013). Using Social,Emotional and Character Development to Improve Academic Outcomes: A Matched Pair, Cluster Randomized Controlled Trial in Low Income, Urban Schools. Journal of

School Health, 83(11), 771-779. http://dx.doi.org/10.1111/josh.12093
Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432. http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x
Humphrey, N., Lendrum, A., & Wigelsworth, M. (2010). *Social and emotional aspects of learning (SEAL) programme in secondary schools: national evaluation*. Research Report DFE-RR049 London: DfE. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/181718/DFE-RR049.pdf
Manchester Institue of Education (2015) *Promoting Alternative Thinking Strategies (PATHS) Evaluation report and Executive summary* London: EEF https://educationendowmentfoundation.org.uk/uploads/pdf/PATHS.pdf
Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38(1), 30. http://dx.doi.org/10.1037/0022-0167.38.1.30
Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). The positive impact of social and emotional learning for kindergarten to eighth-grade students. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning (CASEL). http://files.eric.ed.gov/fulltext/ED505370.pdf
Rimm-Kaufman, S. E., Larsen, R. A., Baroody, A. E., Curby, T. W., Ko, M., Thomas, J. B., & DeCoster, J. (2014). Efficacy of the responsive classroom approach results from a 3-year, longitudinal randomized controlled trial. *American Educational Research Journal.* http://dx.doi.org/0002831214523821
Sklad, M., Diekstra, R., Ritter, M. D., Ben, J., & Gravesteijn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools*, *49*(9), 892-909. http://dx.doi.org/10.1002/pits.21641
Valentine, J.C., DuBois, D. & Cooper, H. (2004). The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review, *Educational Psychologist*, 39(2), 111-133. http://dx.doi.org/10.1207/s15326985ep3902_3

Meta-analyses	FSM effect size	Effect size
Durlak et al. 2011		0.27
Payton et al. 2008		0.28
Sklad et al. 2012		0.46
Recent studies		
Bavarian et al. 2013	0.42	0.22 (English) 0.38 (maths)
MioE 2015 (EEF PATHS)	-0.03 (maths) 0.01 (reading)	0.26 (maths) -0.02 (reading)
Rimm-Kaufmann et al. 2014		0.26 (maths) 0.20 (reading)
Weighted mean effect size		**0.32**

For more information about the effect sizes in the Toolkit, click here.

Technical Appendix:
Social and emotional learning

Education Endowment Foundation

Meta-analyses and abstracts	
Study	***Abstract***
Durlak et al. 2011	This article presents findings from a meta-analysis of 213 school-based, universal social and emotional learning (SEL) programs involving 270,034 kindergarten through high school students. Compared to controls, SEL participants demonstrated significantly improved social and emotional skills, attitudes, behaviour, and academic performance that reflected an 11-percentile-point gain in achievement. School teaching staff successfully conducted SEL programs. The use of 4 recommended practices for developing skills and the presence of implementation problems moderated program outcomes. The findings add to the growing empirical evidence regarding the positive impact of SEL programs. Policy makers, educators, and the public can contribute to healthy development of children by supporting the incorporation of evidence-based SEL programming into standard educational practice.
Payton et al. 2008	This report summarizes results from three large-scale reviews of research on the impact of social and emotional learning (SEL) programs on elementary and middle-school students — that is, programs that seek to promote various social and emotional skills. Collectively the three reviews included 317 studies and involved 324,303 children. SEL programs yielded multiple benefits in each review and were effective in both school and after-school settings and for students with and without behavioral and emotional problems. They were also effective across the K-8 grade range and for racially and ethnically diverse students from urban, rural, and suburban settings. SEL programs improved students' social-emotional skills, attitudes about self and others, connection to school, positive social behavior, and academic performance; they also reduced students' conduct problems and emotional distress. Comparing results from these reviews to findings obtained in reviews of interventions by other research teams suggests that SEL programs are among the most successful youth-development programs offered to school-age youth. Furthermore, school staff (e.g., teachers, student support staff) carried out SEL programs effectively, indicating that they can be incorporated into routine educational practice. In addition, SEL programming improved students' academic performance by 11 to 17 percentile points.
Sklad et al. 2012	To answer the question of whether teaching social and emotional skills to foster social–emotional development can help schools extend their role beyond the transfer of knowledge, the authors conducted a meta-analytical review of 75 recently published studies that reported the effects of universal, school-based social, emotional, and/or behavioral (SEB) programs. The analyzed interventions had a variety of intended outcomes, but the increase in social skills and decrease in antisocial behavior were most often reported. Although considerable differences in efficacy exist, the analysis demonstrated that overall beneficial effects on all seven major categories of outcomes occurred: social skills, antisocial behavior, substance abuse, positive self-image, academic achievement, mental health, and prosocial behavior. Generally, immediate effects were stronger than delayed effects, with the exception of substance abuse, which showed a sleeper effect. Limitations of the analysis and moderators of the effectiveness of SEB programs in schools are discussed in the final section of the article.

Sports participation

Low impact for moderate cost based on limited evidence.

+2

Sport participation is engaging in sports as a means to increase educational engagement and attainment. This might be through organised after school activities or as an organised programme by a local sporting club or association. Sometimes sporting activity is used to encourage young people to engage in additional learning activities, such as football training at a local football club combined with study skills, ICT, literacy or mathematics lessons.

How effective is it?

The overall impact of sports participation on academic achievement tends to be low (less than one additional month's progress), though there is recent evidence from the UK that sports and learning participation can have a more dramatic effect on, for example, mathematics learning as assessed by standardised tests when combined with a structured numeracy programme (with one study showing an impact of up to 10 months' additional progress). In this circumstance the 'participation' acted as an enticement to undertake additional instruction.

The variability in effects suggests that the quality of the programme and the emphasis on or connection with academic learning may make more difference than the specific type of approach or activities involved.

How secure is the evidence?

There have been a number of reviews linking the benefits of participation in sport with academic benefits. There is, however, considerable variation in impact, including some studies which show negative effects.

In England, a 2010 systematic review for the Department for Culture, Media and Sport explored the impact of sport on learning, but did not find any high quality studies conducted in the UK.

What are the costs?

Cost estimates are hard to identify in terms of costs of participation in specific activities (such as a football coaching club, linked with after school study), but are estimated here at up to about £400 per year excluding clothing and equipment. These costs vary according to equipment and venue. Costs are therefore estimated as moderate.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Being involved in extra-curricular sporting activities may increase attendance and retention.
2. Participation in sports does not straightforwardly transfer to academic learning.
3. Planned extra-curricular activities which include short, regular and structured teaching in literacy and mathematics (either tutoring or group teaching) as part of a sports programme, such as an after school club or summer school, are much more likely to offer academic benefits.
4. If you are considering sports participation as an approach to improve attendance, engagement and attainment, have you considered how you will evaluate the impact?

Sports participation Toolkit references

Full references
Cummings, C., Laing, K., Law, J., McLaughlin, J., Papps, I., Todd, L. & Woolner, P. (2012). Can Changing Aspirations And Attitudes Impact On Educational Attainment? A Review Of Interventions York: Joseph Rowntree Foundation.
Coalter, F. (2005). The social benefits of sport: an overview to inform the community planning process. Research Report no. 98. Edinburgh: SportScotland.
Foster, C., Hillsdon, M., Cavill, N., Allender, S., Cowburn, G. (2005). Understanding participation in sport: a systematic review. London: Sport England.
Lewis, C.P. (2004). The Relation Between Extracurricular Activities With Academic And Social Competencies In School Age Children: A Meta-Analysis PhD thesis submitted to the Office of Graduate Studies of Texas A&M University, College Station, Tx (ProQuest Dissertations).
Newman, M., Bird, K., Tripney, J., Kalra, N., Kwan, I., Bangpan, M., Vigurs, C. (2010). Understanding the impact of engagement in culture and sport: A systematic review of the learning impacts for young people In: Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Shulruf, B. (2010). Do extra-curricular activities in schools improve educational outcomes? A critical review and meta-analysis of the literature. International Review of Education, 56(5), 591-612.

Summary of effects	
Study	***Effect size***
Newman et.al., 2010 (academic outcomes)	0.19
Newman et.al., 2010 (mathematics)	0.80
Lewis, 2004	0.10
Shulruf, 2010 (on GPA)	0.15
Indicative effect size	**0.18**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	Abstract
Newman et.al. (2010).	No abstract provided!
Lewis, (2004).	There has been a growing discussion in the fields of education and psychology about the relationship between social skill proficiency and academic excellence. However, the presence of extracurricular involvement as promoting both academic and social development has not been thoroughly explored. The most recent literature syntheses and meta-analyses on extracurricular activity participation were conducted in the 1980.s. An updated review and quantitative look at the participation literature is due. The purpose of this study is to integrate participation studies from the 1990s and give summative information as to the impact of extracurricular activity participation on various educational and psycho-social characteristics. Of the 164 identified studies, 41 were included in these meta-analyses. The current analyses produced 6 different activity categories: general extracurricular activity, sports, work and vocational activities, performing arts, pro-social activities, and community-based activities. The current meta-analyses suggest student outcomes were significantly related to general extracurricular activity and pro-social activity participation. General activities and pro-social activities had the most impact on academic achievement, while performing arts and pro-social activities. Participants reported the largest effect on identity and self-esteem related outcomes.

	Sports and related activities (i.e. Cheerleading) were not as strongly linked to academic achievement indicators as anticipated and student workers had more negative outcomes than any other activity participants. In conclusion, the best outcomes for children and adolescents are brought about through well-built, developmentally appropriate structured activities. Moreover, the academic and social profits of extracurricular activities that have been examined in this study can be used to inform program planning and implementation.
Shulruf, (2010).	Secondary schools tend to sponsor a large number of extra-curricular activities (ECA) yet little is known about their contribution to students' educational outcomes. This meta-analysis aims to determine what it is about ECA participation that supports positive educational outcomes. Furthermore, this study challenges the theoretical assumptions about the benefits of participation in ECA. 29 studies (all except for one based on data collected in the United States) met the search criteria for inclusion in the analysis. Most effect sizes on academic achievements yielded from non-specific ECA, academic clubs and journalism were small, as were participation in performing arts, sports and leadership activities on a range of educational outcomes. Although the results show associations between participation in ECA and educational outcomes, causal effects could not be confirmed. It is concluded that the lack of evidence supporting the causal effects, and thus the common theoretical assumptions about the effects of ECA on educational outcomes, is due to methodology limitations in these studies.

Summer schools

Low impact for moderate cost based on extensive evidence.

+2

Summer schools are lessons or classes during the summer holidays, and are often designed as catch-up programmes. Some summer schools do not have an academic focus and concentrate on sports or other non-academic activities. Others may have a specific focus, such as pupils at the transition from primary to secondary school, or advanced classes to prepare high-attaining pupils for university.

How effective is it?

On average, evidence suggests that pupils who attend a summer school make approximately two additional months' progress, compared to similar pupils who to do not.

Greater impacts (as much as four additional months' progress) can be achieved when summer schools are intensive, well-resourced, and involve small group tuition by trained and experienced teachers. Summer schools without a clear academic component are not usually associated with learning gains. Other variables, such as whether the teacher is one of the student's usual teachers, seem to make less difference on average.

A recent evaluation for the Department for Education in the UK concluded that one of the greatest barriers to impact was achieving high levels of attendance.

How secure is the evidence?

Overall, the level of evidence related to summer schools is extensive. There are a number of meta-analyses, which consistently find small average effects. Studies include both primary and secondary aged pupils and mainly focus on reading and literacy. Some studies indicate that gains are greater for disadvantaged pupils, but this is not consistent.

Most studies have taken place in the USA. However, a meta-analysis combining findings from three recent evaluations of summer schools in England indicates that average gains in literacy of an additional two months progress are also achievable in the UK.

What are the costs?

Overall, the costs are estimated as moderate. Costs include the employment of teachers for the duration of the summer school, hiring a venue and providing resources (books, photocopying etc.). Courses typically cost in the region of £250-£300 per week per student. A recent evaluation of a summer school programme in England estimated the costs as being slightly higher for a summer programme at £1,370 per pupil over four weeks (£340 per pupil, per week).

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Summer school provision that aims to improve learning needs to have an academic component. Does your summer school include an intensive teaching component (small group or one to one)?
2. Summer schools are relatively expensive. Have you considered delivering alternative approaches during the school year, which may provide similar benefits for a lower cost?
3. Maintaining high attendance at summer schools can be a challenge. If you are running a summer school, what steps might you take to engage pupils and their families?

Summer schools
Toolkit references

Education Endowment Foundation

Full references
Day, L. Martin, K., Sharp, C., Gardner, R. & Barham, J. (2013) *Summer Schools Programme for Disadvantaged Pupils: Key findings for Schools Research Report* (DFE- RR271B). London: DfE
Cooper, H, Charlton, V., Muhlenbruck, M., Borman, G.D. (2000). Making the Most of Summer School: A Meta-Analytic and Narrative Review. Monographs of the Society for Research in Child Development, 65.1, pp. 1-127. http://www.jstor.org/discover/10.2307/3181549?sid=21105174367251&uid=2&uid=3738032&uid=4
Gorard, S., Siddiqui, N. & See, B.H. (2014) Future Foundations Evaluation Report and Executive Summary February 2014. London:EEF. http://educationendowmentfoundation.org.uk/uploads/pdf/FINAL_EEF_Evaluation_Report_-_Future_Foundations_-_February_2014.pdf
Lauer P., Akiba, M., Wilkerson, S.B., Apthorp, H.S., Snow, D., & Martin-Glenn, M.L. (2006). Out-of-School-Time Programs: A Meta-Analysis of Effects for At-Risk Students. Review of Educational Research, 76(2), 275-313. http://rer.sagepub.com/content/76/2/275
Lewis, C.P. (2004). The Relation Between Extracurricular Activities With Academic And Social Competencies In School Age Children: A Meta-Analysis PhD thesis submitted to the Office of Graduate Studies of Texas A&M University, College Station, Tx (ProQuest Dissertations).
McClanahan, W.S., Sipe, C.L., & Smith, T.J. (2004). Enriching Summer Work: An Evaluation of the Summer Career Exploration Program Philadelphia, Pa: Public Private Ventures. http://eric.ed.gov/?id=ED503188
Maxwell, B., Connolly, P., Demack, S., O'Hare, L., Stevens, L. & Clague, L. (2014) Summer Active Reading Programme Evaluation Report and Executive Summary October 2014. London: EEF http://educationendowmentfoundation.org.uk/uploads/pdf/EEF_Evaluation_Report_-_Summer_Active_Reading_Programme_-_October_2014.pdf
Torgerson, D., Torgerson, C., Jefferson, L., Buckley, H., Ainsworth, H., Heaps, C. & Mitchell, N. (2014) Discover Summer School Evaluation Report and Executive Summary May 2014. London: EEF http://educationendowmentfoundation.org.uk/uploads/pdf/EEF_Evaluation_Report_-_Discover_Summer_School_-_May_2014.pdf
Borman, G. D., & Dowling, N. M. (2006). Longitudinal achievement effects of multiyear summer school: Evidence from the Teach Baltimore randomized field trial. Educational Evaluation and Policy Analysis, 28(1), 25-48. http://www.dx.doi.org/10.3102/01623737028001025

Summary of effects		
Study	***FSM effect size***	***Effect size***
Cooper et al. 2000	0.24	0.26
Kidron & Lindsay, 2014		0.16 (literacy) 0.08 (maths)
Lauer et.al., 2006.	0.05	0.16
Lewis, 2010		0.10
Recent studies		
Gorard et al. 2014a (EEF Future Foundations)	0.16 (English) - 0.12 (maths)	0.17 (English) 0.00 (maths)
Kim, 2006		0.08 (reading)
Maxwell et al. 2014 (EEF Summer Active Reading Programme)	0.22 (reading)	0.13 (reading)
Torgerson et al. 2014 (EEF Discover Summer School)	0.17 (post-hoc)	0.21 (reading) 0.24 (writing)
Indicative effect size		**0.18**

Summer schools
Toolkit references

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	***Abstract***
Cooper et.al. (2000).	Summer schools serve multiple purposes for students, families, educators, and communities. The current need for summer programs is driven by changes in American families and by calls for an educational system that is competitive globally and embodies higher academic standards. A research synthesis is reported that used both meta-analytic and narrative procedures to integrate the results of 93 evaluations of summer school. Results revealed that summer programs focusing on remedial or accelerated learning or other goals have a positive impact on the knowledge and skills of participants. Although all students benefit from summer school, students from middle-class homes show larger positive effects than students from disadvantaged homes. Remedial programs have larger effects when the program is relatively small and when instruction is individualized. Remedial programs may have more positive effects on math than on reading. Requiring parent involvement also appears related to more effective programs. Students at all grade levels benefit from remedial summer school, but students in the earliest grades and in secondary school may benefit most. These and other findings are examined for their implications for future research, public policy, and the implementation of summer programs. Based on these results, our recommendations to policy makers are that summer programs (a) contain substantial components aimed at teaching math and reading and (b) include rigorous evaluations, but also (c) permit local control of curricula and delivery systems. Funds should be set aside to foster participation in summer programs, especially among disadvantaged youth. Program implementers should (a) begin summer program planning earlier in the year, (b) strive for continuity of staffing and programs across years, (c) use summer school in conjunction with summer staff development opportunities, and (d) begin integrating summer school experiences with those that occur during the regular school year.
Kidron & Lindsay, (2014)	Interest in increased learning time programs delivered beyond the regular school day has grown (Stonehill et al., 2011). These programs provide additional instruction in English language arts, math, and other subjects and are meant to enhance students' academic interests and success (Redd et al., 2012). The most common approaches include out-of-school programs (before- and after-school and weekend programs); summer school; schools with longer school days, weeks, or years; and year-round schools. Numerous evaluations have tested the effects of such programs on students' academic knowledge, study skills, social skills, and motivation to learn. This meta-analysis examined more than 7,000 studies, sorted them by scientific rigor, and identified 30 that used research designs capable of yielding strong evidence about the outcomes of increased learning time. In some cases the 30 studies found that increased learning time programs had a positive effect on student outcomes; in other cases the studies found no positive effect. This suggests that no single increased learning time program fits the needs of all students. The information in this report should help practitioners decide how best to select and implement an increased learning time approach. The programs were found, for example, to improve academic outcomes when instruction was led by certified teachers. Ten studies reported that literacy instruction was delivered by certified teachers and found a statistically significant positive effect on literacy achievement. Five studies reported that math instruction was conducted by certified teachers and found a statistically significant positive effect on math achievement. In both cases, however, the effects were small. Programs that used a traditional instruction style (with the teacher responsible for the progression of activities and students following directions to complete tasks) improved academic outcomes in literacy (nine studies) and math (four studies). The effects were small for both subjects. Programs that used an experiential learning instruction style (such as hands-on, inquiry-based instruction) improved student social-emotional skill

	development (for example, self-confidence and self-management; four studies). Again, the effects were small. The findings also show that increased learning time can benefit students at risk of academic failure. Increased learning time improved the literacy achievement of students performing below standards (three studies). Increased learning time also promoted the social-emotional skill development (for example, emotional well-being and externalizing behaviour) of students with attention deficit/hyperactivity disorder (three studies). Programs that targeted specific student subgroups (such as struggling readers) and used explicit instruction to teach well specified skills tended to show a positive effect on student outcomes. Practitioners who wish to use increased learning time programs might there-fore set goals and design activities based on a deep understanding of student needs and interests. Because this study examined the data one category at a time, it does not provide information on potential interactions among implementation features, such as how the effective-ness of experiential learning, might vary with teacher–student ratio or the frequency and duration of classes. As the evidence base grows, studies like this one will be able to assess the effects of increased learning time using multiple factors at the same time.
Kim & Quin (2013).	This meta-analysis reviewed research on summer reading interventions conducted in the United States and Canada from 1998 to 2011. The synthesis included 41 classroom- and home-based summer reading interventions involving children from kindergarten to Grade 8. Compared to control group children, children who participated in classroom interventions, involving teacher-directed literacy lessons, or home interventions, involving child initiated book reading activities, enjoyed significant improvement on multiple reading outcomes. The magnitude of the treatment effect was positive for summer reading interventions that employed research-based reading instruction and included a majority of low-income children. Sensitivity analyses based on within-study comparisons indicated that summer reading interventions had significantly larger benefits for children from low-income backgrounds than for children from a mix of income backgrounds. The findings highlight the potentially positive impact of classroom- and home-based summer reading interventions on the reading comprehension ability of low income children.
Lauer et.al. (2006).	Schools and districts are adopting out-of-school-time (OST) programs such as after-school programs and summer schools to supplement the education of low-achieving students. However, research has painted a mixed picture of their effectiveness. To clarify OST impacts, this synthesis examined research on OST programs for assisting at-risk students in reading and/or mathematics. Researchers analysed 35 OST studies that employed control or comparison groups and met other inclusion criteria. Meta-analyses indicated small but statistically significant positive effects of OST on both reading and mathematics student achievement and larger positive effect sizes for programs with specific characteristics such as tutoring in reading. Whether the OST program took place after school or during the summer did not make a difference in effectiveness.
Lewis, (2004).	There has been a growing discussion in the fields of education and psychology about the relationship between social skill proficiency and academic excellence. However, the presence of extracurricular involvement as promoting both academic and social development has not been thoroughly explored. The most recent literature syntheses and meta-analyses on extracurricular activity participation were conducted in the 1980.s. An updated review and quantitative look at the participation literature is due. The purpose of this study is to integrate participation studies from the 1990s and give summative information as to the impact of extracurricular activity participation on various educational and psycho-social characteristics. Of the 164 identified studies, 41 were included in these meta-analyses. The current analyses produced 6 different activity categories: general extracurricular activity, sports, work and vocational activities, performing arts, pro-social activities, and community-based activities. The current meta-analyses suggest student outcomes were significantly related to general extracurricular activity and pro-social activity participation. General activities and pro-social activities had the most impact on academic achievement, while performing arts and pro-social

	activities. Participants reported the largest effect on identity and self-esteem related outcomes. Sports and related activities (i.e. Cheerleading) were not as strongly linked to academic achievement indicators as anticipated and student workers had more negative outcomes than any other activity participants. In conclusion, the best outcomes for children and adolescents are brought about through well-built, developmentally appropriate structured activities. Moreover, the academic and social profits of extracurricular activities that have been examined in this study can be used to inform program planning and implementation.

Teaching assistants

Low impact for high cost, based on limited evidence.

+1

Teaching assistants (also known as TAs or classroom support assistants) are adults who support teachers in the classroom. Teaching assistants' duties can vary widely from school to school, ranging from providing administrative and classroom support to providing targeted academic support to individual pupils or small groups.

How effective is it?

Evidence suggests that TAs can have a positive impact on academic achievement. However, effects tend to vary widely between those studies where TAs provide general administrative or classroom support, which on average do not show a positive benefit, and those where TAs support individual pupils or small groups, which on average show moderate positive benefits. The headline figure of one additional month's progress lies between these figures.

Research that examines the impact of TAs providing general classroom support suggests that students in a class with a teaching assistant present do not, on average, outperform those in one where only a teacher is present. This average finding covers a range of impacts. In some cases teachers and TAs work together effectively, leading to increases in attainment. In other cases pupils, particularly those who are low attaining or identified as having special educational needs, can perform worse in classes with teaching assistants.

Where overall negative impacts have been recorded, it is likely that support from TAs has substituted rather than supplemented teaching from teachers. In the most positive examples, it is likely that support and training will have been provided for both teachers and TAs so that they understand how to work together effectively, e.g. by making time for discussion before and after lessons.

Research which focuses on teaching assistants who provide one to one or small group support shows a stronger positive benefit of between three and five additional months on average. Often support is based on a clearly specified approach which teaching assistants have been trained to deliver. Though comparisons with qualified teachers suggest that teaching assistants tend not to be as effective in terms of raising attainment (achieving, on average about half the gains), studies suggest that benefits are possible across subjects and at both primary and secondary level.

How secure is the evidence?

Overall, the level of evidence related to teaching assistants is limited. A number of systematic reviews of the impact of support staff in schools have been conducted in the UK and internationally. However, there are no meta-analyses specifically looking at the impact of teaching assistants on learning.

Correlational studies looking at the impact of TAs providing general classroom support have shown broadly similar effects. One of the most recent studies, conducted in England, suggests that on average low attaining pupils do less well in a class with a TA present, compared to a class where only a teacher is present. More recent intervention studies, including two randomised controlled trials conducted in England in 2013, provide a strong indication that TAs can improve learning if they are trained and deployed carefully. Given the limited amount of existing evidence, these studies made a substantial contribution to the overall evidence base, changing the overall average impact from zero to one additional months' progress.

What are the costs?

The average cost of employing a teaching assistant, including salary and on-costs, is estimated at about £18,000. Overall, costs are estimated as high.

What should I consider?

Before you implement this strategy in your learning environment, consider the following:

1. Have you identified the activities where TAs can support learning, rather than simply managing tasks?
2. Have you provided support and training for teachers and TAs so that they understand how to work together effectively?
3. How will you ensure that teachers do not reduce their support or input to the pupils supported by TAs?
4. Have you considered how you will evaluate the impact of how you deploy your TAs?

Teaching assistants
Toolkit references

Education Endowment Foundation

Full references
Alborz, A., Pearson, D., Farrell, P., Howes, A. (2009). The impact of adult support staff on pupils and mainstream schools. Technical Report. In: Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
Blatchford, P., Bassett, P., Brown, P., Koutsoubou, M., Martin, C., Russell, A. and Webster, R., with Rubie-Davies, C. (2009). The impact of support staff in schools. Results from the Deployment and Impact of Support Staff project. (Strand 2 Wave 2) (DCSF-RR148). London: Department for Children, Schools and Families.
Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., ... & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. Journal of Child Psychology and Psychiatry, 49(4), 422-432.
Brown, K. J., Morris, D., & Fields, M. (2005). Intervention after grade 1: Serving increased numbers of struggling readers effectively. Journal of Literacy Research, 37(1), 61-94.
Burgoyne, K., Duff, F. J., Clarke, P. J., Buckley, S., Snowling, M. J., & Hulme, C. (2012). Efficacy of a reading and language intervention for children with Down syndrome: a randomized controlled trial. Journal of Child Psychology and Psychiatry, 53(10), 1044-1053.
Butt, R. and Lowe, K. (2012). Teaching assistants and class teachers: differing perceptions, role confusion and the benefits of skills-based training. *International Journal of Inclusive Education*, 16(2), 207–219.
Butt, R. (2014). The Changing Role of Teaching Assistants – Where Being a 'Mum' is Not Enough. Doctoral Thesis. University of Canberra. http://www.canberra.edu.au/researchrepository/file/145cc5ef-a45e-461f-b4bf-2b41553fa9cb/1/full_text.pdf [URL].
Duff, F. J., Fieldsend, E., Bowyer-Crane, C., Hulme, C., Smith, G., Gibbs, S., & Snowling, M. J. (2008). Reading with vocabulary intervention: Evaluation of an instruction for children with poor response to reading intervention. Journal of Research in Reading, 31(3), 319-336.
Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading Rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. American Educational Research Journal, 44(2), 414-448.
Farrell, P., Alborz, A., Howes, A. & Pearson, D. (2010). 'The impact of teaching assistants on improving pupils' academic achievement in mainstream schools: a review of the literature'. Educational Review, 62.4 pp 435-448.
Gerber, S.B., Finn, J.D., Achilles, C.M. and Boyd-Zacharias, J. (2001). Teacher aides and students' academic achievement. Educational Evaluation and Policy Analysis 23.2 pp 123-143.
Gorard, S., See, B. H., Siddiqui, N. (2014) Switch-on Reading Evaluation Report and Executive Summary February 2014. London: Education Endowment Foundation.
Gray, C., McCloy, S., Dunbar, C., Dunn, J., Mitchell, D., & Ferguson, J. (2007). Added value or a familiar face? The impact of learning support assistants on young readers. Journal of Early Childhood Research, 5(3), 285-300.
Hatcher, P. J., Goetz, K., Snowling, M. J., Hulme, C., Gibbs, S., & Smith, G. (2006). Evidence for the effectiveness of the Early Literacy Support programme. British Journal of Educational Psychology, 76(2), 351-367.
Mecrow, C., Beckwith, J., & Klee, T. (2010). An exploratory trial of the effectiveness of an enhanced consultative approach to delivering speech and language intervention in schools. International Journal of Language & Communication Disorders, 45(3), 354-367.
Mercer, C. D., Campbell, K. U., Miller, M. D., Mercer, K. D., & Lane, H. B. (2000). Effects of a reading fluency intervention for middle schoolers with specific learning disabilities. Learning Disabilities Research & Practice, 15(4), 179-189.
Miller, S. D. (2003). Partners-in-reading: Using classroom assistants to provide tutorial assistance to struggling first-grade readers. Journal of Education for Students Placed At Risk, 8(3), 333-349.
Moore, W. and Hammond, L. (2011). Using education assistants to help pave the road to literacy: Supporting oral language, letter-sound knowledge and phonemic awareness in the pre-primary year. *Australian Journal of Learning Difficulties*, 16(2), 85–110.

Teaching assistants
Toolkit references

Morris, D. (2006). Using Noncertified Tutors to Work with At Risk Readers: An Evidence-Based Model. The Elementary School Journal, 106(4), 351-362.
Muijs, D. & Reynolds, D. (2003). The effectiveness of the use of learning support assistants in improving the mathematics achievement of low achieving pupils in primary school. Educational Research, 45.3 pp 219-230.
NFER (2014) A Randomised Trial of Catch Up Numeracy® Evaluation Report and Executive Summary February 2014. London: Education Endowment Foundation.
Ritter, G.W., Barnett, J.H., Genny, C.S., & Albin, G.R. (2009). The Effectiveness of Volunteer Tutoring Programs for Elementary and Middle School Students: A Meta-Analysis. Review of Educational Research, 79 (3), 3-38.
Savage, R., Carless, S., & Stuart, M. (2003). The effects of rime and phoneme-based teaching delivered by learning support assistants. Journal of Research in Reading, 26(3), 211-233. Savage, R., & Carless, S. (2008). The impact of early reading interventions delivered by classroom assistants on attainment at the end of Year 2. British Educational Research Journal, 34(3), 363-385. (Follow up of 2003 study)
Swann, W., & Loxley, A. (1998). The impact of school-based training on classroom assistants in primary schools. Research papers in education, 13(2), 141-160.
Vadasy, P. F., Sanders, E. A., & Tudor, S. (2007). Effectiveness of paraeducator-supplemented individual instruction beyond basic decoding skills. Journal of Learning Disabilities, 40(6), 508-525.
Vadasy, P. F., & Sanders, E. A. (2008). Repeated reading intervention: Outcomes and interactions with readers' skills and classroom instruction. Journal of Educational Psychology, 100(2), 272.

Teaching assistants
Toolkit references

Summary of effects	
Study (NB All single studies. No meta-analyses)	***Effect size***
Ehri et al, 2007.	0.74
Gerber et.al., 2001 (compared with regular classes)	0.00
Gerber et.al., 2001 (compared with small classes)	-0.15
Blatchford et.al., 2009 (single study)	0.00
Gerber, Finn, Achilles & Boyd-Zacharias, 2001	0.00
Muijs, & Reynolds, 2003)	0.00
Blatchford, Russell, Bassett, Brown, & Martin, 2007	0.00
Blatchford et al. 2011	-0.01
Gray et al. 2007	0.00
Hatcher et al. 2006	0.15
Mecrow et al., 2010	0.12
Savage, Carless, & Stuart, 2003	0.37
Moore, & Hammond, 2011	0.79
Brown, Morris, & Fields, 2005	1.18
Mercer et al., 2000	1.5
Morris, 2006	0.76
Vadasy & Sanders, 2007	0.22
Indicative effect size	**0.08**

For more information about the effect sizes in the Toolkit, click here.

Meta-analyses abstracts	
Study	**Abstract**
Blatchford et.al. (2009).	In recent years there has been an unprecedented increase in support staff in schools in England and Wales. There were widespread expectations that this will be of benefit to teachers and pupils but there has been little systematic research to address the impact of support staff. This study used a naturalistic longitudinal design to investigate the relationship between the amount of support (measured by teacher estimates and systematic observation) and pupils' 'Positive Approaches to Learning' (PAL) and academic progress. There were over 8000 pupils across two cohorts and seven age groups. Results on PAL were not straightforward by there was a consistent trend for those with most support to make less academic progress than similar pupils with less support, and this was not explained by characteristics of the pupils such as prior attainment or level of special educational need. In Years 1, 3, and 7 in English and mathematics and for English in Year 10 the higher the level of support, the lower the level of attainment. In English in Year 1 those pupils with a medium level of support obtained attainment scores that were almost two points less than those with a low level of support. Two points equate to one sub-level of the main National Curriculum levels (e.g. the difference between level 1B and 1C). There was a difference of roughly three units between the pupils with most and least support, or about one and a half sub-levels.
Gerber et.al. (2001).	Despite more than 600,000 teacher aides in American schools today, research provides little information about their classroom activities, their qualifications for carrying out their duties, or their impact on student achievement and behavior. This investigation asked whether the presence of a teacher aide in the classroom has any noticeable impact on pupils' learning. Three primary questions were addressed: (1) In Grades K through 3, does the presence of a full-time teacher aide in the classroom affect

	students' academic achievement? (2) If teacher aides have a positive effect on students' performance, does the effect depend on the number of years the student attends classes with a teacher aide? (3) Do some functions of aides (i.e., clerical tasks, instructional tasks, non-instructional tasks) have a greater impact on student achievement than others? This investigation showed that the teacher aide movement in the United States has created a state of affairs that requires many aides to perform tasks for which they are ill-prepared. In addition, teacher aide data were analyzed from Tennessee's Project STAR, a longitudinal experiment in which students were assigned at random to small classes, regular-size classes without an aide, or regular-size classes with a full-time teacher aide. The analyses reported here extend previous investigations, examining the functions and effects of teacher aides in depth. The results showed that teacher aides have little, if any, positive effect on students' academic achievement. The only positive effect was an improvement in readings cores for students who attended a class with a teacher aide for 2 or 3 years. These results were the only exceptions to a plethora of negative findings. The study also showed that the types of duties aides performed had no bearing on student achievement. Because teacher aides are called upon increasingly to provide instruction to pupils, policies and research must help us select and prepare aides to perform effectively.
Muijs, D., & Reynolds, D. (2003).	The use of learning support assistants in schools has become increasingly common in England, partly as a result of government support for this strategy. One suggested advantage of the deployment of learning support assistants is to provide additional support to low achieving pupils. However, so far research on the effects of this strategy is very limited. In this study the effect of using trained learning support assistants to help pupils underachieving in mathematics is examined using a quasi-experimental design whereby pupils receiving support were matched to those not receiving support on prior achievement and pupil background factors. Having matched the 180 pupils in the project schools who had received support from an NSA with 180 pupils who had not on a number of variables, including free meal eligibility, prior achievement, special needs, ethnicity and gender, we found that pupils who had received NSA support did not make more progress in mathematics than those who had not.
Blatchford, P., Russell, A., Bassett, P., Brown, P., & Martin, C. (2007).	It is widely assumed that increasing the number of teaching assistants (TAs) in the classroom will be beneficial to children, and this is one important aim of the recently implemented Workforce Agreement. But there are still significant gaps in knowledge about many aspects of their deployment and impact. The Class Size and Pupil-Adult Ratios (CSPAR) KS2 study built on earlier findings when the pupils were in reception and Key Stage (KS) 1, and investigated: (1) the deployment of TAs in classrooms and how key parties involved perceived this; and (2) the effect of TAs on interactions involving pupils and teachers in the same classrooms, and on pupil attainments. The study had a longitudinal, mixed method and multi-informant design. Methods of data collection included: (for the whole sample) questionnaires completed by TAs, teachers and head teachers, assessments of pupil attainments in mathematics, English and science, data on pupil background, and (for a sub-sample) case studies and a systematic observation study. This study found that the TA's role in KS2 is predominantly a direct one, in the sense of face-to-face interactions supporting certain pupils. There was no evidence that the presence of TAs, or any characteristic of TAs, had a measurable effect on pupil attainment. However, results were clear in showing that TAs had an indirect effect on teaching, e.g. pupils had a more active form of interaction with the teacher and there was more individualised teacher attention. This

	supported teachers' views that TAs are effective in supporting them in this way. The authors conclude that more attention needs to be paid to what they call the pedagogical role of TAs so that they can be used effectively to help teachers and pupils.
Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., ... & Hulme, C. (2008).	This study compares the efficacy of two school-based intervention programmes (Phonology with Reading (P + R) and Oral Language (OL)) for children with poor oral language at school entry. Methods: Following screening of 960 children, 152 children (mean age 4;09) were selected from 19 schools on the basis of poor vocabulary and verbal reasoning skills and randomly allocated to either the P + R programme or the OL programme. Both groups of children received 20 weeks of daily intervention alternating between small group and individual sessions, delivered by trained teaching assistants. Children in the P + R group received training in letter-sound knowledge, phonological awareness and book level reading skills. Children in the OL group received instruction in vocabulary, comprehension, inference generation and narrative skills. The children's progress was monitored at four time points: pre-, mid- and post-intervention, and after a 5-month delay, using measures of literacy, language and phonological awareness. Results: The data are clustered (children within schools) and robust confidence intervals are reported. At the end of the 20-week intervention programme, children in the P + R group showed an advantage over the OL group on literacy and phonological measures, while children in the OL group showed an advantage over the P + R group on measures of vocabulary and grammatical skills. These gains were maintained over a 5-month period. Conclusions: Intervention programmes designed to develop oral language skills can be delivered successfully by trained teaching assistants to children at school entry. Training using P + R fostered decoding ability whereas the OL programme improved vocabulary and grammatical skills that are foundations for reading comprehension. However, at the end of the intervention, more than 50% of at-risk children remain in need of literacy support.
Burgoyne, K., Duff, F. J., Clarke, P. J., Buckley, S., Snowling, M. J., & Hulme, C. (2012).	This study evaluates the effects of a language and literacy intervention for children with Down syndrome. Methods: Teaching assistants (TAs) were trained to deliver a reading and language intervention to children in individual daily 40-min sessions. We used a waiting list control design, in which half the sample received the intervention immediately, whereas the remaining children received the treatment after a 20-week delay. Fifty-seven children with Down syndrome in mainstream primary schools in two UK locations (Yorkshire and Hampshire) were randomly allocated to intervention (40 weeks of intervention) and waiting control (20 weeks of intervention) groups. Assessments were conducted at three time points: pre-intervention, after 20 weeks of intervention, and after 40 weeks of intervention. Results: After 20 weeks of intervention, the intervention group showed significantly greater progress than the waiting control group on measures of single word reading, letter-sound knowledge, phoneme blending and taught expressive vocabulary. Effects did not transfer to other skills (nonword reading, spelling, standardised expressive and receptive vocabulary, expressive information and grammar). After 40 weeks of intervention, the intervention group remained numerically ahead of the control group on most key outcome measures; but these differences were not significant. Children who were younger, attended more intervention sessions, and had better initial receptive language skills made greater progress during the course of the intervention. Conclusions: A TA-delivered intervention produced

	improvements in the reading and language skills of children with Down syndrome. Gains were largest in skills directly taught with little evidence of generalization to skills not directly taught in the intervention. Keywords: Down syndrome, early literacy, intervention, language, phonological awareness,RCT.
Gray, C., McCloy, S., Dunbar, C., Dunn, J., Mitchell, D., & Ferguson, J. (2007).	As the number of learning support assistants (also referred to as classroom assistants, teaching aides, auxiliary support and paraeducators) employed in the primary sector continues to increase, questions concerning the contribution they make to raising standards of attainment remain to be answered. This article reports evidence from a small-scale study undertaken to evaluate the impact learning support assistants have on the reading attainment of young children taught in schools using a whole class systematic phonics approach. A matched sample test/retest approach was employed to examine differences in the reading performance of pupils who did and did not receive additional reading support. Whilst overall, pupils exposed to the whole class systematic phonics approach showed a significant improvement in reading performance, no added value was noted for pupils receiving learning support assistance. On the contrary, the results suggest that learning support may have a detrimental impact on lower ability readers. Contradictions between the empirical findings and the views of teachers are discussed here.
Hatcher, P. J., Goetz, K., Snowling, M. J., Hulme, C., Gibbs, S., & Smith, G. (2006).	It is widely recognized that effective interventions for poor reading involve training in phoneme awareness and letter-sound knowledge, linked in the context of reading books. From the applied perspective, it is important to gather data on the effectiveness of different forms of implementation of literacy support within this framework. Aim. We evaluate the effectiveness of the UK Early Literacy Support (ELS) programme (Department for Education and Skills [DfES], 2001) relative to a programme of Reading Intervention based on 'sound linkage' (Hatcher, Hulme, & Ellis, 1994), both delivered by teaching assistants. Sample. The sample comprised 128 6-year-old children, from 16 primary schools in a Local Education Authority (LEA) in the north of England. Method. The children were nominated as in need of special help by their class teachers and allocated to one of the two programmes. Results. After controlling for a difference in spelling ability at the start of the intervention, it was found that both groups of children made equivalent and significant gains in reading and spelling that were maintained at follow-up. The standardized scores were in the average range. Regression analyses confirmed the importance of initial reading ability as well as letter identification, phoneme manipulation, and sound linkage in predicting progress in learning to read and to spell. Conclusion. The ELS programme provides a cost effective method of boosting 6-year-old children's reading to an average level.
Mecrow, C., Beckwith, J., & Klee, T. (2010).	Increased demand for access to specialist services for providing support to children with speech, language and communication needs prompted a local service review of how best to allocate limited resources. This study arose as a consequence of a wish to evaluate the effectiveness of an enhanced consultative approach to delivering speech and language intervention in local schools. Aims: The purpose was to evaluate an intensive speech and language intervention for children in mainstream schools delivered by specialist teaching assistants. Methods & Procedures: A within-subjects, quasi-experimental exploratory trial was conducted, with each child serving as his or her own control with respect to the primary outcome measure. Thirty-five children between the ages of 4;2 and 6;10 (years; months) received speech and/or language intervention for an average of four 1-hour

	sessions per week over 10 weeks. The primary outcome measure consisted of change between pre- and post-intervention scores on probe tasks of treated and untreated behaviours summed across the group of children, and maintenance probes of treated behaviours. Secondary outcome measures included standardized tests (Clinical Evaluation of Language Fundamentals — Preschool UK (CELF-PUK); Diagnostic Evaluation of Articulation and Phonology (DEAP)) and questionnaires completed by parents/carers and school staff before and after the intervention period. Outcome & Results: The primary outcome measure showed improvement over the intervention period, with target behaviours showing a significantly larger increase than control behaviours. The gains made on the target behaviours as a result of intervention were sustained when reassessed 3–12 months later. These findings were replicated on a second set of targets and controls. Significant gains were also observed on CELF-PreschoolUK receptive and expressive language standard scores from pre- to post-intervention. However, DEAP standard scores of speech ability did not increase over the intervention period, although improvements in raw scores were observed. Questionnaires completed before and after intervention showed some significant differences relating to how much the child's speech and language difficulties affected him/her at home and at school. (CELF-P UK receptive composite) 0.08 (CELF-P UK expressive composite) Conclusions & Implications: This exploratory study demonstrates the benefit of an intensive therapy delivered by specialist teaching assistants for remediating speech and language difficulties experienced by young children in mainstream schools. The service delivery model was perceived by professionals as offering an inclusive and effective practice and provides empirical support for using both direct and indirect intervention in the school setting.
Boyle J, McCartney E, Forbes J, O'Hare A.A. (2007).	Objectives: To compare language outcomes following direct individual therapy [speech and language therapist (SLT) working individually with a child], indirect individual therapy [speech and language therapy assistant (SLTA) working individually with a child], direct group therapy (SLT working with a small group of children) and indirect group therapy (SLTA working with a small group of children) for primary school-age children with persistent primary receptive and/or expressive language impairment relative to a comparison group receiving current models and levels of SLT service. Design: The trial had a 2 x 2 factorial design (direct/indirect versus individual/group therapy) together with a control group that received existing levels of community-based speech and language therapy and served as a comparator for the economic analysis. All postintervention language outcome measures were blind assessed. A short-run economic evaluation across the four different modes of therapy was carried out using the primary outcome measure. A comparable method was used for estimating the costs of providing services in the community for children allocated to the control group. Setting: Research intervention took place in school settings in Scotland, with some of the children randomised to group therapies transported to join a group in a different school. Participants: Children aged between 6 and 11 years, attending a mainstream school, with standard scores on the Clinical Evaluation of Language Fundamentals (CELF-3UK) of less than -1.25 SD (receptive and/or expressive) and non-verbal IQ on the Wechsler Abbreviated Scale of Intelligence (WASI) above 75, and no reported hearing loss, no moderate/severe articulation/phonology/dysfluency problems or otherwise

	requiring individual SLT work. Informed, written parental consent was required. Interventions: A therapy manual was constructed that provided a range of procedures and activities for intervention in areas identified by a search of the research and professional literature for examples of language therapies of proven effectiveness. SLTs planned activities for children seen by therapists and SLTAs, using the manual. Main outcome measures: Primary outcome measures were standardised scores on tests of expressive and receptive language. Secondary outcome measures were scores on a test of receptive vocabulary, together with questionnaire, rating scale and focus group data from parents, teachers, project SLTs and SLTAs, and an audit of therapy sessions. Results: There was no evidence that the five modes involved in the project were different at the onset in terms of primary outcome measures, although there were significant gender differences. The results from both the intention-to-treat analyses of the outcomes from the 161 children randomised who met the eligibility criteria and the protocol analyses of the outcomes from the 152 children for whom postbaseline data were available revealed that there were no significant postintervention differences between direct and indirect modes of therapy on the one hand, or between individual and group modes on the other on any of the primary language outcome measures. However, there was evidence of some benefits from direct therapy from an SLT in secondary outcome measures. Intervention delivered three times a week for 30-40 minutes over a 15-week period also yielded significant improvements in age-corrected standardised scores for expressive language, although not for receptive language, relative to those receiving community-based SLT services. Children with specific expressive language delay were more likely to show improvement than those with mixed receptive-expressive difficulties, and non-verbal IQ was not a significant moderating variable. The within-trial economic evaluation identified indirect therapy, particularly indirect group therapy, as the least costly of the modes investigated in the study, with direct individual therapy as the most costly option. Conclusions: SLTAs can act as effective surrogates for SLTs in the delivery of services within primary schools to children with primary language impairment who do not to require the specialist skills of an SLT. Generalising the central estimates of the relative cost of different therapy modes to other educational/health systems is possible, but the precise differences reported in resource use need to be qualified by the level of programme intensity and other characteristic features of education and therapy services that may differ from those observed in this trial. Further research is needed into effective interventions for receptive language problems and also investigations of the efficacy of the relationship between dose and treatment effect in both expressive and receptive language. Research is also needed into models of integrative service delivery, cluster models of delivery via integrated community schools, and the involvement of class teachers, classroom assistants and parents/carers. There is also a need for studies to identify the characteristics of children who are most likely to succeed with indirect intervention approaches, and also to evaluate alternative methods of working with those who may benefit from different modes. Finally, research to refine the therapy manual would also be helpful.

Duff, F. J., Fieldsend, E., Bowyer-Crane, C., Hulme, C., Smith, G., Gibbs, S., & Snowling, M. J. (2008).	Interventions combining phonically based reading instruction with phonological training are generally effective for children with reading (decoding) difficulties. However, a minority of children respond poorly to such interventions. This study explored the characteristics of children who showed poor response to reading intervention and aimed to improve their literacy and language skills via a new theoretically motivated intervention. Twelve 8-year-old treatment poor responders with severe and persisting reading difficulties participated. A 9-week reading intervention incorporating reading, phonological and vocabulary training was implemented. Before the intervention began the children showed almost no progress over 6 months of regular classroom education, on measures of oral language and literacy. Over the intervention period improvements were made on measures of reading, phonological awareness and language skills, which were maintained 6 months later. Although the intervention was effective, it should be noted that most children remained poor readers and require ongoing remediation.
Moore, W. & Hammond, L. (2011)	Children with weak oral language skills are at risk of experiencing difficulty with early literacy acquisition. Intensive small group intervention during the pre-primary year has the potential to improve children's success in developing emergent literacy skills. Education assistants are a potentially powerful resource for supporting students at educational risk. In this study, education assistants at four schools were trained to provide a daily half-hour emergent literacy program to pre-primary students with low oral language skills. The program focused on developing phonological awareness, letter-sound knowledge and vocabulary using both explicit and in-context (embedded) learning activities. The students undertaking the program made significant gains on early language and literacy measures. Case studies are presented that illustrate the strengths and limitations of the intervention for children and schools.
Brown, K. J., Morris, D., & Fields, M. (2005).	The present study replicated the original evaluation of the Howard Street tutoring model (Morris, Shaw, & Perney, 1990), an intervention for struggling readers in second and third grade. It also evaluated the effectiveness of supervised paraprofessionals (Title I aides) in delivering that tutorial. For an entire school year, teachers or paraprofessionals, working under the supervision of a reading specialist, tutored 40 struggling readers twice per week for 45 minutes per session. The tutored group's instruction included guided reading in levelled texts with controlled vocabulary, word study, and reading for fluency. The control group's instruction, which was provided daily in a small-group context, featured guided reading and phonics work in the classroom basal reader. Analysis of covariance was used to compare the performance of the two groups on several end-of-year reading measures. Results showed that, overall, the tutored group outperformed the control group on each of the posttest reading measures (standardized and informal). In addition, the subset of students tutored by paraprofessionals outperformed the control students. In fact, results indicated that in the structured tutoring context, paraprofessional tutors were almost as effective as certified teachers.
Mercer, C. D., Campbell, K. U., Miller, M. D., Mercer, K. D., & Lane, H. B. (2000).	A fluency-reading intervention was developed and used to supplement reading instruction of middle school students with learning disabilities (LD). The daily 5- to 6-min fluency intervention focused on phonics, sight phrases, and oral reading. Repeated readings were used in each area until the student achieved mastery on each respective task. The students were

	divided in three groups according to the length of the intervention (6-9 months, 10-18 months, and 19-25 months). Dependent t tests were used to test the effects of the intervention. A pretest and posttest curriculum-based assessment (CBA) measure was used for each group to determine progress in reading rate. In addition, progress within the intervention was measured using comparisons of beginning and ending reading levels. Significant growth in reading level and reading rate was found. Implications of these results for students with LD are discussed.
Morris, D. (2006)., Miller, S. D. (2003).	This article synthesizes results from 5 studies that used noncertified tutors to work with at-risk primary-grade readers. Each of the studies featured (1) twice-weekly tutoring lessons that included guided reading, word study, and reading for fluency; and (2) supervision of the tutoring by a knowledgeable reading teacher. Results from the 5 studies provided convergent evidence that noncertified reading tutors (community volunteers and teacher aides) can be effective with struggling readers. However, their effectiveness was, in large part, due to the amount and quality of guidance they received from the supervising reading teacher. Partners-in-Reading (Miller, 2003), SC Teacher assistants First graders (N = 54) Word recognition .78; Comprehension .76 Next Steps (Morris, 2005) Salt Lake City, UT Teacher assistants Second and third graders (N = 21) Word recognition 0.78; Passage reading 0.55; Comprehension 1.01. This study evaluated the feasibility of using classroom assistants as tutors of 1st-grade struggling readers in a school with limited financial and personnel resources. The tutoring program, Partners-in-Reading (PIR), offered assistance to 54 first graders in 2 cohorts. Classroom assistants scheduled tutoring a minimum of 4 times per week for 30 to 40 min per session: A typical session included the reading and rereading of familiar texts, an introduction of texts at or slightly above a student's instructional level, and various word recognition activities. PIR students' word recognition and development spelling scores were compared with Reading-Recovery (RR) students (n = 62) and a control group (n = 58). Although equivalent at the year's start, PIR and RR students outperformed controls on these measures at the end of 1st grade. They also scored higher than did the controls on a norm-referenced word recognition subtest and were less likely to be retained. PIR students also outperformed the controls on a norm-referenced comprehension subtest. This discussion focuses on the benefits of using classroom assistants as tutors and the related questions of when tutoring should be offered, its duration, and its evaluation.]
Vadasy, P. F., Sanders, E. A., & Tudor, S. (2007).	A total of 46 children in Grades 2 and 3 with low word-level skills were randomly assigned to 1 of 2 groups that received supplemental phonics-based reading instruction. One group received intervention October through March (21.5 hours), and one group served as a control from October through March and later received intervention March through May (17.5 hours). Paraeducators trained in a standard treatment protocol provided individual instruction for 30 min per day, 4 days per week. At the March posttest, the early treatment (ET; n = 23) group outperformed the controls (late treatment, LT; n = 20) on reading accuracy and passage fluency. Across both groups, second graders outperformed third graders on these same measures. At the 3-month follow-up, the ET group showed no evidence of decline in reading accuracy, passage fluency, or words spelled; however, 3rd-grade ET students had significantly higher spelling skills compared to 2nd graders. The LT group demonstrated significant growth during their intervention in reading

	accuracy and spelling, but not passage fluency. When we compared the ET and LT groups on their gains per instructional hour, we found that the ET group made significantly greater gains than the LT group across all 3 measures. The results support the value of paraeducator-supplemented reading instruction for students below grade level in word identification and reading fluency.
Vadasy, P. F., & Sanders, E. A. (2008).	This study examined effects of a repeated reading intervention, Quick Reads, with incidental word-level scaffolding instruction. Second- and third-grade students with passage-reading fluency performance between the 10th and 60th percentiles were randomly assigned to dyads, which were in turn randomly assigned to treatment (paired tutoring, n = 82) or control (no tutoring, n = 80) conditions. Paraeducators tutored dyads for 30 min per day, 4 days per week, for 15 weeks (November–March). At midintervention, most teachers with students in the study were formally observed during their literacy blocks. Multilevel modeling was used to test for direct treatment effects on pretest–posttest gains as well as to test for unique treatment effects after classroom oral text reading time, 2 pretests, and corresponding interactions were accounted for. Model results revealed both direct and unique treatment effects on gains in word reading and fluency. Moreover, complex interactions between group, oral text reading time, and pretests were also detected, suggesting that pretest skills should be taken into account when considering repeated reading instruction for 2nd and 3rd graders with low to average passage-reading fluency. d = .29, .31, .32, and .43, for gains in word reading accuracy, fluency rate, PRF–U, and PRF–A, respectively. results from these models revealed significant unique treatment effects on word reading efficiency (approximate d = .32), PRF–U (d = .65), PRF–A (d = .81), fluency rate (d = .58), and comprehension (d =.51).
Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007).	The Reading Rescue tutoring intervention model was investigated with 64 low–socioeconomic status, language-minority first graders with reading dif-ficulties. School staff provided tutoring in phonological awareness, systematic phonics, vocabulary, fluency, and reading comprehension. Tutored students made significantly greater gains reading words and comprehending text than controls, who received a small-group intervention (d = 0.70) or neither inter- vention (d = 0.74). The majority of tutored students reached average reading levels whereas the majority of controls did not. Paraprofessionals tutored students as effectively as reading specialists except in skills benefiting nonword decoding. Paraprofessionals required more sessions to achieve equivalent gains. Contrary to conventional wisdom, results suggest that students make greater gains when they read text at an independent level than at an instructional level.
NFER (2013) A Randomised Trial of Catch Up Numeracy®Evaluation Report and Executive Summary November 2014.	Catch Up Numeracy is a one-to-one intervention for primary school children who are struggling with numeracy. It consists of two 15-minute lessons per week for up to 30 weeks, usually delivered by teaching assistants. The intervention is aimed at primary aged children struggling with mathematics. The project ran from September 2012 to July 2013 and was delivered by Catch Up, and Professor Ann Dowker of Oxford University. The external evaluator was the National Foundation for Educational Research (NFER) One to one teaching was shown to significantly increase numeracy scores. There was no evidence that Catch Up added anything over and above these effects. Sub group analysis did not identify any differential effects for pupil gender or eligibility for free school meals. Successful implementation would benefit from; TAs having sufficient time to plan and prepare for the sessions and finding

	time to fit in two 15 minute sessions each week within the existing school timetable □ managing the negative attitude towards the intervention, for example pupils resenting being singled out as 'not very good at maths' or being taken out of lessons they enjoy such as PE in order to do Catch Up. □ finding suitable resources to run the intervention □ finding a suitable location within school to run the sessions □ senior management support and commitment to the programme.
Gorard, S., See, B. H., Siddiqui, N. (2014) Switch-on Reading Evaluation Report and Executive Summary February 2014.	Switch-on Reading is an intensive 10-week literacy intervention. It is delivered on a one to one basis by staff, most commonly teaching assistants, who have been trained in the approach. The purpose of Switch-on is to achieve functional literacy for as many pupils as possible, and so to close the reading achievement gap for vulnerable children working below age-expected levels. It is based on the longer and somewhat more expensive intervention Reading Recovery, which is teacher-led and delivered over a 12-20 week period. In this evaluation, the programme involved regular sessions for pupils who had not achieved Level 4 English at Key Stage 2. The identified pupils attended regular 20 minute reading sessions over the course of one term. The students were removed from class to attend the sessions, which aimed to improve their reading comprehension and fluency. Each session required students to read from four different books graded on the basis of their difficulty. Training and support for staff was provided by the Every Child a Reader staff of Nottinghamshire local authority. What impact did it have? The overall result was an effect size of +0.24, based on the pooled standard deviation of the post-test score for both groups, meaning that the programme made a noticeable positive impact. This effect can be envisaged as suggesting that on average pupils receiving the intervention would make approximately three additional months' progress over the course of a year compared to similar pupils who did not. The evaluation identified positive results for all groups of pupils (defined by sex, first language, ethnicity, special educational needs, free school meal eligibility and measured attainment at the outset). However, it is important to note that conclusions about specific groups of pupils are more tentative than the overall finding, because the study was primarily designed to test the average impact on all identified children. Pupils with low attainment prior to the intervention showed particularly positive results, making five additional months progress on average. Pupils eligible for free school meals and pupils identified as having special educational needs made four additional months progress on average. As such, this evaluation suggests that Switch-on can be an effective intervention for weak and disadvantaged readers at the stage of transition to secondary school. The intervention was generally well-conducted and the pupils seemed very happy with their reading sessions. Staff needed training and then some monitoring to ensure that they adhered to the protocol, which was necessary for the intervention to have the largest possible effect.