Sutton Trust - EEF
Teaching and Learning Toolkit

Summer 2013

Approaches

APPROACH	COST ESTIMATE	EVIDENCE ESTIMATE	AVERAGE IMPACT	SUMMARY
Ability grouping	£	★★★	- 1 Month	Negative impact for very low or no cost, based on moderate evidence.
After school programmes	££££	★★	+ 2 Months	Low impact for high cost, based on limited evidence.
Arts participation	££	★★★	+ 2 Months	Low impact for low cost, based on moderate evidence.
Aspiration interventions	£££	★	0 Months	Very low or no impact for moderate cost based on very limited evidence.
Behaviour interventions	£££	★★★★	+ 4 Months	Moderate impact for very high cost, based on extensive evidence.
Block scheduling	£	★★	0 Months	Very low or negative impact for very low of no cost, based on limited evidence.
Collaborative learning	£	★★★★	+ 5 Months	Moderate impact for very low cost, based on extensive evidence.
Digital technology	££££	★★★★	+ 4 Months	Moderate impact for high cost, based on extensive evidence.
Early years intervention	£££££	★★★★	+ 6 Months	High impact for very high costs, based on extensive evidence.
Extended school time	£££	★★	+ 2 Months	Low impact for moderate cost, based on limited evidence.
Feedback	££	★★★	+ 8 Months	High impact for low cost, based on moderate evidence.
Homework (Primary)	£	★★★	+ 1 Month	Low impact for very low or no cost, based on moderate evidence.
Homework (Secondary)	£	★★★	+ 5 Months	Moderate impact for very low or no cost, based on moderate evidence.
Individualised instruction	£	★★★	+ 2 Months	Low impact for low cost, based on moderate evidence.
Learning styles	£	★★★	+ 2 Months	Low impact for very low cost, based on moderate evidence.
Mastery learning	££	★★★	+ 5 Months	Moderate impact for low cost, based on moderate evidence.
Mentoring	£££	★★★	+ 1 Month	Low impact for moderate cost, based on moderate evidence.
Meta-cognition and self-regulation	££	★★★★	+ 8 Months	High impact for low cost, based on extensive evidence.
One to one tuition	££££	★★★★	+ 5 Months	Moderate impact for high cost, based on extensive evidence.
Outdoor adventure learning	£££	★★	+ 3 Months	Moderate impact for moderate cost, based on limited evidence.
Parental involvement	£££	★★★	+ 3 Months	Moderate impact for moderate cost, based on moderate evidence.
Peer tutoring	££	★★★★	+ 6 Months	High impact for low cost, based on extensive evidence.
Performance pay	££	★	0 Months	Low or no impact for moderate cost, based on very limited evidence.
Phonics	£	★★★★	+ 4 Months	Moderate impact for very low cost, based on extensive evidence.
Physical environment	££	★	0 Months	Very low or no impact for low cost based on very limited evidence.
Reducing class size	£££££	★★★	+ 3 Months	Low impact for very high cost, based on moderate evidence.
Repeating a year	£££££	★★★★	- 4 Months	Negative impact for very high cost based on extensive evidence.
School uniform	£	★	0 Months	Very low or no impact for very low cost, based on very limited evidence.
Small group tuition	£££	★★	+ 4 Months	Moderate impact for moderate cost, based on limited evidence.
Social and emotional learning	£	★★★★	+ 4 Months	Moderate impact for very low cost, based on extensive evidence.
Sports participation	£££	★★	+ 2 Months	Moderate impact for moderate cost based on moderate evidence.
Summer schools	£££	★★	+ 3 Months	Moderate impact for moderate cost based on limited evidence.
Teaching assistants	££££	★★	0 Months	Very low or no impact for high cost, based on limited evidence.

Introduction

The Sutton Trust-EEF Teaching and Learning Toolkit is an accessible summary of educational research which provides guidance for teachers and schools on how to use their resources to improve the attainment of disadvantaged pupils.

The Toolkit currently covers 33 topics, each summarised in terms of their average impact on attainment, the strength of the evidence supporting them and their cost.

The Toolkit is a live resource which will be updated on a regular basis as findings from EEF-funded projects and other high-quality research become available. In addition, we would welcome suggestions for topics to be included in future editions. If you have a topic suggestion, or any other comments or questions about the Toolkit, please contact Robbie Coleman at robbie.coleman@eefoundation.org.uk.

Why is research useful?

We know that the relationship between spending and pupil outcomes is not simple. Between 1997 and 2011 per pupil spending increased by 85% but over this period improvements in pupil outcomes were marginal on most measures. At 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 taken 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 achievement 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 2012-13 the Pupil Premium is worth £623 per child, and by 2014-15 this is expected to rise to approximately £1,200 per child. 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.

Average impact

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

For example, research summarised in the Toolkit shows that improving the quality feedback provided to pupils has an average impact of eight months. This means that pupils in a class where high quality feedback is provided will make on average eight months more progress over the course of a year compared to another class of pupils which were 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 estimations are based on ‘effect sizes’ reported in British and international comparative data (see table below). Effect sizes are quantitative measures of the impact of different approaches on learning. The Toolkit prioritises systematic reviews of research and quantitative syntheses of data such as meta-analyses of experimental studies. To be included in the analysis an approach needed to have some quantifiable evidence base for comparison. For more information about the Toolkit's methodology please view the Toolkit's Technical Appendices.

Months' Progress	Effective Size From...	...to	Description
0	-0.01	0.01	Very Low or no effect
1	0.02	0.09	Low
2	0.10	0.18	Low
3	0.19	0.26	Moderate
4	0.27	0.35	Moderate
5	0.36	0.44	Moderate
6	0.45	0.52	High
7	0.53	0.61	High
8	0.62	0.69	High
9	0.70	0.78	Very High
10	0.79	0.87	Very High
11	0.88	0.95	Very High
12	0.96	>1.0	Very High

Cost

Cost estimations are based on the approximate cost of implementing an approach in a class of twenty five pupils. Where the approach does not require an additional resource, estimates are based on the cost of training or professional development which may be required. Approaches marked with £££ or less could be funded from the 2012-13 pupil premium allocation of £623 per eligible pupil. For more information about the Toolkit's methodology please view the Toolkit's 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-£5,000 per year per class of 25 pupils, or up to about £170 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. This represents the 2012/13 Pupil Premium allocation (£623).
££££	*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. By 2014/5, the Pupil Premium is projected to rise to approximately £1,200 per pupil.

Evidence

Evidence estimates are based on: the availability of evidence (i.e. the number of systematic reviews or meta-analyses and the quantity of primary studies which they synthesise); the methodological quality of the primary evidence; the magnitude of the impact (in terms of effect size); and the reliability or consistency of this impact across the studies reviewed. For more information about the Toolkit's methodology please view the Toolkit's Technical Appendices.

Rating	Description
★	*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.
★★	*Limited:* At least one meta-analysis or systematic review with quantitative evidence of impact on attainment or cognitive or curriculum outcome measures.
★★★	*Moderate:* Two or more rigorous meta-analyses of experimental studies of school age students with cognitive or curriculum outcome measures.
★★★★	*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.
★★★★★	*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.

Notes on the July 2013 Update

The Toolkit is a live resource which will be updated on a regular basis as findings from EEF-funded projects and other high-quality research become available.

Major updates made to the Toolkit in July 2013 include:

- The addition of three new topics: **Mastery learning, Outdoor adventure learning**, **Repeating a year**.
- The publication of updated **Technical Appendices.**
- An updated entry for **Ability Grouping.**

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 is written by Professor Steve Higgins, Maria Katsipataki and Dr Dimitra Kokotsaki (School of Education, Durham University), Professor Rob Coe (CEM Centre, Durham University), Dr Lee Elliot Major (The Sutton Trust) and Robbie Coleman (Education Endowment Foundation).

Full reference: Higgins, S., Katsipataki, M., Kokotsaki, D., Coleman, R., Major, L.E., & Coe, R. (2013). The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit. London: Education Endowment Foundation.

Using the Toolkit

Like any toolkit, the Teaching and Learning Toolkit will be most useful when in the hands of professionals. The aim of the Toolkit is to support teachers to make their own informed choices and adopt a more 'evidence based' approach. The evidence it contains is a supplement to rather than a substitute for professional judgement; it provides no guaranteed solutions or quick fixes.

We believe that the Toolkit should be used as one step in a decision making process. One possible process is shown in Figure 1, below.

Figure 1. How should the Toolkit be used?

Step 1

Before identifying a new strategy and considering how to evaluate it, it is important to consider your school's context, and what you want to achieve. Much depends on your school, its teachers (their levels of knowledge and experience), and its pupils (their level of attainment and their social background). Internal data and professional judgement should be used to identify priorities.

Step 2

Having identified what you want to achieve, the summaries in the Toolkit can be used to help identify solutions. Crucially, the summaries in the Toolkit combine evidence from a range of different research studies into a single average for each area. **This average will not necessarily be the impact of this approach in your school.** Some of the approaches which are less effective on average might be effective in a new setting or if developed in a new way. Similarly, an approach which tends to be more effective on average may not work so well in a new context. However, we think that evidence of average impact elsewhere will be useful to schools in making a good 'bet' on what might be valuable, or may strike a note of caution when trying out something which has not worked so well in the past.

To take an example we have discussed with many teachers since the Toolkit was launched in 2011, the fact that the average impact of teaching assistants (TAs) is not positive in no way means that TAs cannot have a positive impact on attainment. However, it does imply that schools might want to think carefully about the strategies they use to ensure that their TAs are deployed and supported effectively.

Step 3

As a result of the importance of context, it is crucial to use the Toolkit alongside on-going evaluations of the impact of the decisions you make, to ensure that the approaches you use are having the desired effect. To help with this step the EEF has published a DIY Evaluation Guide which provides advice for schools on how to evaluate new strategies as robustly as possible. Many changes in schools initially feel positive but have little lasting impact on learning so this step is essential.

Finally, it should be noted that the evidence summarised in the Toolkit takes educational attainment as its primary metric. Most of the measures used are traditional measures of attainment such as curriculum tests and examinations. This focus does **not** suggest that all educational aims and outcomes are captured in the literature that we have pulled together. Though we highlight impacts on other outcomes such as aspiration, attendance or behaviour where this information is available, these outcomes are not systematically recorded, or reflected, in the overall summary.

Ability grouping

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

£ £ £ £ £	★★★☆☆	-1
cost per pupil	evidence rating	month

What is it?

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, ability grouping appears to benefit higher attaining pupils and be detrimental to the learning of mid-range and lower attaining learners. On average, ability grouping does not appear to be an effective strategy for raising the attainment of disadvantaged pupils, who are more likely to be assigned to in lower attaining groups. Summer born pupils and students from ethnic minority backgrounds are also likely to be adversely affected by ability grouping.

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 ability 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 three and 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 without ability grouping. 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 ability grouping has benefitted all learners. Further study could be undertaken to understand what happened differently in these examples.

Evidence suggests that the impact of setting is more detrimental to low attaining pupils in mathematics who do better in mixed attainment groups, and that ability grouping particularly affects upper primary and lower secondary education. The effects appear to be less clear-cut in other subjects, though negative effects are reported for low attaining pupils across the curriculum.

Though the average impact of ability grouping on low attaining pupils is negative, evidence suggests that certain types of ability 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 ability 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. Though much of the research into mixed ability grouping is dated and based on studies from overseas, a recent study conducted across 45 secondary schools in the UK showed a similar effect to previous research. This study suggests that setting in mathematics may be a means of raising the attainment of higher attaining pupils in Years 7 to 9, but that the benefits for pupils in high performing groups come at the cost of reducing the attainment of lower attaining pupils, who make better progress in mixed ability classes. Overall the evidence is estimated as moderate.

What are the costs?

Ability grouping is an organisational strategy which has few, if any, financial costs associated with it. Additional resources may be needed to support different groups. Overall the costs are estimated as low.

What do I need to know?

- The key issue is ensuring that any ability grouping benefits all learners, particularly low-attaining or disadvantaged pupils, over both the long and short term.
- It is important to recognise that a measure of current attainment, such as a recent curriculum test, is not the same as a measure of ability or of potential.
- Schools should consider how differences in grouping will support more effective teaching or intensive support for lower attaining pupils.
- The impact of any grouping by attainment should be monitored closely, particularly on attitudes to learning and the engagement of pupils.
- Flexible within-class grouping is preferable to tracking or streaming for low attaining pupils.

After school programmes

Low impact for high cost, based on limited evidence.

£ £ £ £ £ cost per pupil

★ ★ ★ ★ ★ evidence rating

+2 months

What is it?

Children or young people are involved in planned activities which are supervised by adults at the end of the school day. The goals, objectives and approaches of the programmes may vary greatly. Some will have an academic focus and be taught by teachers from the school the pupils attend, others will have a wider variety of activities supported by adults with a range of skills and qualifications.

How effective is it?

Research indicates that participating in after school programmes improves performance on measures of academic achievement. However, the gains are low to moderate on tested attainment of reading or mathematics (with a benefit of about an additional two months progress per year) and there is an inconsistent pattern of impact across studies, which suggests that the quality and focus of the programme is important. In the UK there is evidence that such programmes are linked with GCSE improvement by a third of a level in maths and three-quarters of a level in science. There is evidence that there are wider benefits for low-income students in terms of attendance at school, behaviour and relationships with peers.

Programmes may not be equally effective with all students. At risk children are more likely to benefit as are younger children (5-10 year olds). Positive effects for reading were highest for younger primary pupils and in secondary schools. Maths gains were higher for older primary and secondary pupils. However the research indicates that it is harder to attract and retain pupils in after school programmes at secondary level compared with primary pupils. Programmes which support and encourage children academically while providing stimulating environments and activities will most likely link to engagement. Additionally, teacher's support, promotion of interaction and mutual respect appear to be some of the key elements in enhancing participation.

How secure is the evidence?

There are a number of reviews and a comprehensive meta-analyses, mainly using data from the USA, but with broadly similar findings from less rigorous evaluations undertaken in the UK. Analysis suggests that enthusiasm for after school programmes in the USA has outpaced the research base indicating the need for more rigorous evaluations with outcome measures that demonstrate effectiveness on learning.

What are the costs?

In the UK, official estimates suggest after school clubs cost on average £7 per session, indicating that about 15 weeks of after school provision could be supported by the pupil premium of £600 in 2012-13. The costs of well-qualified and well-trained staff may increase these estimates, particularly if they involve tutoring, so the Toolkit estimates about £10 per session per pupil to take into account the academic focus needed. £10 a day for about half a school year (100 days) comes to about £1,000 per pupil. Costs are therefore estimated as high.

What do I need to know?

- Programmes with greater structure, a strong link to the curriculum, well-qualified and well-trained staff are more clearly related to academic benefits.
- Particularly promising after school activities include one to one or small group tuition.
- Enrichment activities (such as sports or arts engagement) may have positive benefits on attitudes, but these alone will not improve academic learning.
- Particular effort may be required to engage and retain older secondary pupils.
- Booster activities to support revision and test or exam practice are likely to improve results.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/after-school-programmes

Arts participation

Low impact for low cost, based on moderate evidence.

£ £ £ £ £	★ ★ ★ ★ ★	+ 2
cost per pupil	evidence rating	months

What is it?

Arts participation is involvement in terms of performance and creation in artistic and creative extra-curricular activities, such as dance, drama, music, painting, sculpture. 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 on academic learning tends to be low, though greater effects have been identified for younger learners of primary school age in terms of impact on cognitive tests. Wider benefits on attitudes and well-being have also consistently been reported.

There is reasonably consistent but weak evidence that participation in artistic and creative activities is beneficial. Outcomes have been identified from arts participation in terms of impact on English, mathematics and science learning in school at both primary and secondary school level. Specific benefits are linked with some particular activities (such as spatial awareness and music for example). There is some evidence that younger learners may benefit more from these approaches.

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 vary according to the detail of the approach and the age group targeted so the effects are hard to generalise and not conclusive.

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 do I need to know?

- The research evidence shows a wide range of effects from programmes studied, suggesting that achieving learning gains from arts programmes is not straightforward.
- 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.
- The transfer of learning to the classroom is not automatic and needs further exploration, e.g. by encouraging pupils to apply their learning from arts participation in more formal contexts.
- Arts-based approaches may offer a route to re-engage older learners in school.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/arts-participation

Aspiration interventions

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

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

0
months

What is it?

Aspirations are about 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: focusing on parents and families, through work by teachers, and through out-of-school interventions or extra-curricular activities, sometimes involving peers and mentors. Approaches which seek to raise aspirations are very diverse and may seek 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 counter-intuitive – 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 which 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 learning gains can be credited for raising aspirations rather than the additional academic support or increased parental involvement.

How secure is the evidence?

Generally the evidence base on aspiration is weak, and further, more rigorous studies are required, particularly focusing on pupil level rather than school level interventions. There are two systematic and high quality 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, aspirations and attitudes to school. There are no meta-analyses of interventions to raise aspirations which report impact on attainment or learning. Most studies look at the relationship between aspirations and attainment and some find a link, particularly between low aspirations and low attainment. However this does not mean that raising aspirations will raise attainment. 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.

What are the costs?

Costs vary widely, and are hard to estimate precisely. After school programmes typically cost about £5-£10 per session, so a 20 week programme once per week would cost a maximum of £200 per pupil. Parental involvement programmes also vary in costs but again are typically between £200 per child per year where 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 to raising aspiration in the USA have been estimated at $900 per student per year or about £560. Overall the costs are estimated as moderate.

What do I need to know?

- The relationship between aspirations and attainment is not straightforward; in general, approaches to raise aspirations have not translated into increased learning.
- 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.
- The attitudes, beliefs and behaviours that surround aspirations in disadvantaged communities are diverse so generalisations should be avoided.
- Interventions which have positive effects almost always have a significant academic component, suggesting that raising aspirations in isolation will not be effective.
- For pupils or learners with low aspirations, it is important to monitor the impact of any interventions or approaches if the goal is to improve attainment, given the mixed success of interventions to date.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/aspiration

Behaviour interventions

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

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

+4
months

What is it?

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 aims 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 which 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 there is relatively wide variation between alternative 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 (at middle or secondary rather than primary level). 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 alternative to direct behaviour interventions.

School level behaviour approaches are often associated with improvement 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. Evidence mainly comes from studies in the US where problematic behaviours in schools have been studied in more depth than in the UK.

Some caution in interpreting findings is needed as the majority of the meta-analyses on behaviour focus on pupils diagnosed with specific emotional or behavioural disorders. There is also considerable variation in impact between interventions studies, with one meta-analysis of an anger management intervention showing 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 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 (about £500 for professional development, so about £20 per pupil per year). One to one support is much more expensive, but more effective (about £40 per hour, or £640 per pupil for 15 sessions). Overall, costs are estimated as moderate.

What do I need to know?

- Targeted interventions for those diagnosed or at-risk of emotional or behavioural disorders produce the greatest effects.
- Programmes of two to six months seem to produce more long-lasting results.
- The wide variation in impact among evaluated programmes suggests that schools should look for programmes with a proven track record of impact.
- Training of facilitators or professional development improves the impact of programmes
- On average, programmes which involve parent or community involvement show higher effects.

Block scheduling

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

£ £ £ £ £	★★☆☆☆	0
cost per pupil	evidence rating	months

What is it?

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: 4 blocks of 80–90 minute classes in one day, students take 4 subjects in one term
- A/B: classes of 70-90 minutes each for 3/4 different subjects on every alternating day; and
- Hybrid: 5 classes per day, between 55 and 90 minutes in length.

How effective is it?

There is no consistent pattern in the evidence. The most recent systematic review concluded that 4x4 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 the introduction of block scheduling is unlikely to raise attainment by itself. This evidence is insufficient to support the introduction block scheduling in secondary schools to raise attainment by itself. 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. 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 a reasonable number of studies and one systematic review which looks at the quantitative evidence of the impact of timetabling and scheduling changes on students' learning.

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 do I need to know?

- Timetabling changes alone are not sufficient to improve learning
- Teachers need to alter the way that they teach and should plan and organise different kinds of learning activities to obtain benefits.
- Timetabling changes need to be matched to curriculum goals and teaching and learning objectives (such as longer lessons for science experiments).
- One promising approach might be to investigate how longer lessons can increase the amount of feedback that students get from the teacher or from each other (see Feedback).

Collaborative learning

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

£ £ £ £ £ cost per pupil

★★★★☆ evidence rating

+5 months

What is it?

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 a range of approaches to collaborative and co-operative 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 promote 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 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 do I need to know?

- Pupils will need support and practice to work together; this does not happen automatically.
- Tasks need to be designed carefully so that working together is effective and efficient, otherwise some pupils will try to work on their own.
- 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.
- It is particularly important to encourage lower achieving pupils to talk and articulate their thinking in collaborative tasks as they may contribute less.
- Managing effective collaborative group work is challenging so professional development or collaborative professional inquiry is likely to be helpful to support effective use of these approaches.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/collaborative-learning

Digital technology

Moderate impact for high cost, based on extensive evidence.

£ £ £ £ £ cost per pupil

★ ★ ★ ★ ★ evidence rating

+4 months

What is it?

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 practice 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 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 do I need to know?

- Effective use of technology is driven by learning and teaching goals rather than a specific technology: technology is not an end in itself.
- It is important to identify clearly how the introduction of technology will improve learning rather than assuming that new technology will automatically lead to increased attainment; technology without pedagogy is very unlikely to be effective.
- Technology should support pupils to work harder, for longer or more efficiently to improve their learning.
- 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.
- Teachers need support and time to learn to use new technology effectively. This involves more than just learning how to use the technology and should include support to use it for teaching through professional development.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/digital-technology

Early years intervention

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

£ £ £ £ £
cost per pupil

★★★★☆
evidence rating

+6
months

What is it?

Early years or early childhood interventions are approaches which 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.

How effective is it?

Overall, the evidence suggests that early years and pre-school intervention is beneficial with above average levels of impact (a typical impact of six additional months progress). There is some international evidence that these programmes need to be for a whole day (rather than half-day which on average has less impact, though it should be noted the UK's Effective Provision of Pre-School Education (EPPE) study did not find a difference) and of longer duration (up to a year or more) rather than for shorter periods of time.

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 where 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 US 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 Sure Start average cost per child was about £1,000 in 2006, so the estimates are in the region of £1,000-£2,000 per child. This can be compared with the average yearly child-care costs for a child under two at about £5,000. Overall, the costs are estimated as very high.

What do I need to know?

- High quality provision is essential with well-qualified and well trained staff.
- Such 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.
- Extended attendance (1 year+) and starting early (i.e. 3 years) is more likely to have an impact than shorter durations starting later, which on average produce much lower gains.
- Disadvantaged children benefit from good quality programmes, especially where these include a mixture of children from different social backgrounds, and a strong educational component.
- Immediate impact of Home based programmes is associated with a shorter duration.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/early-years-intervention

Extended school time

Low impact for moderate cost, based on limited evidence.

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

+2
months

What is it?

Internationally, two main approaches to extending school time have been implemented and evaluated: 1. Extending the length of the school day; 2. Extending the length of the school year. The are examples of the school day being extended to up to 12 hours per day and the school year being extended by up to five additional weeks. Specific approaches to increasing learning time are included in other sections of the Toolkit, such as *Summer School*, *After School Programmes* and *Early Years Intervention*; this summary is limited to extending core school time.

How effective is it?

Most of the studies find evidence of improved learning compared to shorter days or school years, but this is usually quite small and gains are not consistent across all studies. Unsurprisingly, the amount of improved learning appears to depend heavily on how the time is used and which aspects of teaching and learning are increased. Evidence suggests that it is likely to be cheaper and more efficient to focus on using existing school time more effectively before considering extending school time.

Overall approaches to increasing the length of the school day or the school year add on average two months additional progress to pupils' attainment over the course of a year. Additionally, research based on international comparisons, looking at average times for schooling in different countries is consistent with this conclusion. However, it should also be noted that pupils from disadvantaged backgrounds benefit by, on average, an additional half a month's progress relative to their peers suggesting that extending school time can be an effective means to improve learning for pupils who are most at risk of failure.

Caution should be taken to ensure that any increase in school time does not reduce time for other positive activities either for pupils (e.g. activities which support overall development and well-being, or time to complete homework) or for teachers (e.g. lesson preparation time). To be successful any increases should be supported by both parents and staff, and extreme increases (e.g. above nine hours per day) do not appear to be effective.

How secure is the evidence?

Most of the evaluations of initiatives to extend school time come from the USA and are from wider evaluations of school reform or school improvement models that incorporate an extended school day as one component among a number of changes. This makes attributing any academic gains to either an extended day or an extended year difficult, though what evidence there is suggests that both extending the school day and extending the school year can improve academic attainment, particularly for pupils from disadvantaged backgrounds. More analyses have been undertaken on extending the school year so there is more evidence in this area, but given the current state of the evidence however, it may be better to invest in the quality of teaching and learning in schools in the first instance, rather than the quantity.

What are the costs?

The costs of extending the school day or the school year are rarely explicit in the studies reviewed. It is not clear when teachers were paid more for the additional hours or days worked or that the costs of running the schools for this extra time has been estimated. One US study which aimed to increase school time by 30% worked on a budget of $1,300 per student, per year (about £800). Average costs per pupil in primary schools are about £2,500 and for secondary about £3,500, which is about £13 and £18 per pupil per day. Extending the school year by two weeks would therefore need about £260 per pupil per year for primary schools and about £360 per pupil per year for secondary, if the same spending model is used. The costs are therefore estimated as moderate.

What do I need to know?

- In terms of a longer school day there are indications that smaller increases are associated with greater gains, and with more than three of four hours a day the benefit decreases.
- It is important to look at the quality of teaching and learning in school time as well as the quantity; it might be cheaper and more efficient to attempt to use existing time more effectively before considering extending the school day.
- Staff commitment is vital or any changes may increase staff turnover.
- It may be necessary to do things differently with the extra time, rather than provide more of the same teaching and learning activities.
- Schools should consider what pupils and staff would stop doing because of extended school time.

Feedback

High impact for low cost, based on moderate evidence.

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

+8
months

What is it?

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 by means of ICT. It can come from a teacher or someone taking a teaching role (including pupils acting as teachers) 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 this as an approach. The 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. In general research-based approaches which 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.

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 assessment for learning 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. Estimates of this (including up to 7-10 days cover) are in the region of £2,000-£3,000 per teacher per year or about £100 per pupil. Overall costs are estimated as low.

What do I need to know?

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 (getting a balance between support and challenge) and be given sparingly so that it is meaningful (as too much feedback can stop learners working out what they need to do for themselves).
- provide specific guidance on how to improve and not just tell students when they are wrong.
- be supported with effective professional development for teachers.
- 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).

Homework (Primary)

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

£ £ £ £ £	★★★☆☆	+1
cost per pupil	evidence rating	month

What is it?

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?

It is certainly the case that schools whose pupils do homework tend to be successful schools. 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 some evidence that when homework is used as a short and focused intervention it can be effective in improving students' attainment (with some studies showing up to eight months positive impact on attainment). Overall the general benefits are likely to be modest if homework is more routinely set. There is clear evidence that it is helpful at secondary level, but there is much less evidence of benefit at primary level.

The research strongly suggests that it is more valuable at secondary school level and much less effective for children of primary school age.

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. There is a relatively consistent picture that there is a positive association, but 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 do I need to know?

- Overall, homework in primary schools does not appear to lead to large increases in learning.
- Effective homework is associated with greater parental involvement and support and can be developed to increase parental engagement.
- 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.
- The purpose of homework should be made clear to children.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/homework

Homework (Secondary)

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

£ £ £ £ £	★★★☆☆	+5
cost per pupil	evidence rating	months

What is it?

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 a part of rather 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 1-2 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 be successful schools, 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 do I need to know?

- 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.
- The purpose of homework should be made explicit to learners, e.g. to increase a specific area of knowledge, or fluency in a particular area.
- It should not be used as a punishment or penalty for poor performance.
- A variety of tasks with different levels of challenge is likely to be beneficial.
- The quality of homework is more important than the quantity. Pupils should receive feedback on homework which is specific and timely.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/homework-secondary

Individualised instruction

Low impact for low cost, based on moderate evidence.

£ £ £ £ £	★★★☆☆	+2
cost per pupil	evidence rating	months

What is it?

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 low.

What do I need to know?

- Overall the evidence does not support approaches which individualise instruction at class level.
- 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).
- Individualised instruction runs the risk of the teacher managing diverse activities and learners, without sufficient time to work directly with learners to teach them.
- It might be a more viable strategy in small group or one-to-one settings, where giving learners direct teaching at the same time is still possible.
- 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.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/individualised-instruction

Learning styles

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

£ £ £ £ £	★★★☆☆	+2
cost per pupil	evidence rating	months

What is it?

The idea underpinning learning styles is that we all have different approaches or styles of learning and that learning will therefore be more effective or more efficient if we are taught accordingly. For example a student may prefer words versus pictures or analysis versus listening.

How effective is it?

It has proved difficult to identify reliably any consistent learning 'styles' in young people and evidence suggests that it is unhelpful to assign learners to groups or categories on the basis of a supposed learning style. Learning preferences do change in different situations and over time and there is some evidence that cognitive preference and task type may be connected (e.g. visualisation in some areas of mathematics is particularly valuable). However, studies where targeted learning takes place in conjunction with activities that match an identified learning style have not convincingly shown any benefit, particularly for low attaining pupils. In fact, in some studies the controls did better than the learning styles groups. Overall impacts recorded are low or negative, suggesting that only one or two pupils in a class of 25 might benefit from this approach.

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, rather fostering a belief than that they can succeed through effort, but the lack of impact of learning styles has been documented at all stages of education.

Where gains have been documented these may come from pupils taking responsibility for learning (see *Meta-cognition*) rather than directly from the use of learning styles approaches.

How secure is the evidence?

Overall the picture is consistent and reasonably robust. 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 unreliability of learning styles tests have also been the focus of a number of reviews.

What are the costs?

The costs are relatively low, though some of the available tests of learning styles require purchase.Typically, these about about £5 per pupil.

What do I need to know?

- 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 are still very flexible.
- 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.
- It appears to be more promising to focus on other aspects of motivation to engage pupils in learning activities.
- Pupils should be encouraged to take responsibility for identifying how they can succeed in their learning and develop their own successful strategies and approaches.
- 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.

Mastery learning

Moderate impact for low cost, based on moderate evidence.

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

+5
months

What is it?

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 to 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.

Mastery learning has a long history going back to Benjamin Bloom's work in the 1960s, though aspects of it resonate with more recent developments such as: Assessment for Learning, in terms of explicit outcomes and targets; Carol Dweck's Growth Mindset work, in terms of the expectation that everyone can succeed with effort; and Response to Intervention, in terms of providing structured support. It has also been adopted in other areas of learning such as professional medical education, where it is widely used.

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 here, 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. The evidence suggests that when pupils work at their own pace, as opposed to working as a part of group or whole class, it is 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. For example schools may choose to use mastery learning for particularly challenging topics or important concepts.

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?

There is a large quantity of research on the impact of mastery learning, though it is important to note that much of it is relatively dated and that its 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 recent small study in the US showed that mastery learning approaches can increase achievement by up to six months in mathematics for 13-14 year olds, which is consistent with several older studies. Overall, the evidence base supporting mastery learning is judged to be moderate.

What are the costs?

Costs are hard to estimate as much of the expenditure necessary to make mastery learning work lies in professional development and planning time. Additional small group tuition and one to one support are also likely to be needed. Costs are estimated at about £5,000 per class per year to include professional development, additional resource preparation, and intensive support for up to 20% of the class over the year.

What do I need to know?

- Overall, mastery learning is a learning strategy offering high potential, which appears to be particularly effective for low attaining students.
- Implementing mastery learning effectively is not straightforward, however, requiring a number of complex components and a significant investment in terms of design and preparation.
- 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.
- Providing structured support for pupils who fall behind with a range of interventions, such as peer support and intensive tuition, may help maintain more even progress within classes.
- Incorporating group and team approaches where pupils take responsibility for helping each other within mastery learning appears to be effective.

Mentoring

Low impact for moderate cost, based on moderate evidence.

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

+1
month

What is it?

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 and competence, or to develop resilience and character. Mentors typically build relationships with young people by meeting with them one-to-one for about an hour a week either at school, or at the end of the school day or weekends. Activities will 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. Community and school-based mentoring schemes have expanded rapidly, particularly in the USA. It can be distinguished from coaching or volunteer tutoring where the focus is very much on improving performance in academic subjects, though mentoring is sometimes referred to as 'life-coaching'.

How effective is it?

The impact of mentoring is variable, but on average it has tended to be low in terms of direct effect on academic outcomes. There is some evidence that pupils from disadvantaged backgrounds are likely to benefit more (nearly double the impact). Other positive benefits have been reported in terms of attitudes to school, attendance and behaviour. However, there are also risks associated with unsuccessful mentor pairings which may have a detrimental effect on the mentee, and the negative overall impacts seen in some studies should prompt caution. 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. 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.

How secure is the evidence?

The evidence has been fairly consistent over the last decade or so, and the quality of more recent evaluations from the USA has been higher than in the past. The most recent randomised controlled trials have not been combined in a meta-analysis, but show similar impact to earlier meta-analyses and systematic reviews. Rigorous evaluation of mentoring programmes and approaches in the UK is needed.

What are the costs?

Compared with other professionally delivered interventions and approaches, mentoring is relatively inexpensive. Costs are mainly for mentor training and support and for 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 for space of general support. Estimates in the USA are between $1000-$1500 per student per year or about £600-£850 per pupil per year, some of which appears to pay for the costs of voluntary organisation providing the mentors. Costs are therefore estimated as moderate.

What do I need to know?

- The impact of mentoring varies, but overall, it is likely to have only a small impact on attainment.
- Positive effects tend not to be sustained once the mentoring stops, and often end abruptly, so care must be taken to ensure that benefits are not lost.
- Mentors will benefit from training and support.
- Community based approaches tend to be more successful than school based approaches.
- It is vital that the mentor is reliable, as there is evidence that mentor drop-out can have detrimental effects on their mentee.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/mentoring

Meta-cognition and self-regulation

High impact for low cost, based on extensive evidence.

£ £ £ £ £	★★★★☆	+8
cost per pupil	evidence rating	months

What is it?

Meta-cognitive and self-regulation strategies (sometimes known as 'learning to learn' strategies) are teaching approaches which make learners think about learning more explicitly. This is usually by teaching pupils specific strategies to set goals, monitor and evaluate their own learning. Self-regulation refers to managing one's own motivation towards learning as well as the more cognitive aspects of thinking and reasoning. Overall these strategies involve being aware of one's strengths and weaknesses as a learner, such as by developing self-assessment skills, and being able to set and monitor goals. They also include having a repertoire of strategies to choose from or switch to during learning activities.

How effective is it?

Meta-cognitive and self-regulation approaches have consistently high levels of impact with meta-analyses reporting between seven and nine months additional progress on average. It is usually more effective in small groups so learners can support each other and make their thinking explicit through discussion.

Encouragingly the evidence suggests that teaching meta-cognitive and self-regulation strategies tends to be particularly effective with lower achieving pupils, as well as with older students. Most studies have looked at the impact on English or mathematics, though there is some evidence from other areas such as science, suggesting benefits are likely to be widely applicable.

The potential impact of approaches which encourage learners to plan, monitor and evaluate their learning is very high. However it can be difficult to achieve these gains as this involves pupils in taking greater responsibility for their learning and in developing their understanding of what is involved in being successful. There is no simple strategy 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. A useful metaphor is scaffolding in terms of *removing* the support and dismantling the scaffolding to check that learners are taking responsibility to manage their own learning.

How secure is the evidence?

There are a number of systematic reviews and meta-analyses of programmes and approaches which promote thinking about thinking which have consistently found similar levels of impact.

What are the costs?

Costs are relatively low, though many studies report the benefits of professional development and/or outside support, or an inquiry approach for teachers where they actively evaluate strategies as they use them. A course of sustained professional development or collaborative professional inquiry is estimated at £2-3,000 per year (including some release from classroom teaching) or about £100 per pupil.

What do I need to know?

- Teaching approaches which encourage learners to plan, monitor and evaluate their learning have very high potential, but require careful implementation.
- Teach pupils explicit strategies to plan, to monitor and to evaluate their learning, and give them opportunities to use them with support and then independently.
- When using approaches for planning, ask pupils to identify the different ways that they could plan (general strategies) and about best approach for a particular task (specific technique).
- Monitoring involves identifying the key steps they need to be aware of as they go through a task to keep it on track. (Where might this go wrong? What will be the difficult parts?)
- Evaluating can be part of the process of checking so that it feeds into the current task as it nears completion (Can you make it better? Are you sure this is right?). It can also feed forward into future tasks (What have you learned that will change what you do next time?).

For more information, videos and supporting resources relating to this approach, please visit http://educationendowmentfoundation.org.uk/toolkit/approaches/meta-cognitive-and-self-regulation-strategies

One to one tuition

Moderate impact for high cost, based on extensive evidence.

£ £ £ £ £ cost per pupil

★ ★ ★ ★ ★ evidence rating

+5 months

What is it?

One to one tuition is where an individual pupil is removed from their class and given intensive tuition. It may also be undertaken outside of normal lessons, for example as part of after school programmes or summer schools.

How effective is it?

Evidence indicates that in areas like reading and mathematics one-to-one tuition can enable learners to catch up with their peers. Research has been focused on children who are falling behind their peers, though one-to-one tuition reliably provides benefit. Meta-analyses indicate that pupils might make about 4 or 5 months progress during an intensive programme.

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. However there is no strong evidence that one-to-one is better than paired tuition or intensive small group teaching, and some evidence that pairs make better progress than individual pupils. Evidence also suggests tutoring should be additional or supplemental to normal instruction, rather than as a replacement and that teachers should monitor progress to ensure the tutoring 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. Programmes which used experienced and specifically trained teachers are more effective than those using volunteers or classroom assistants (nearly double the effect). Where tuition is delivered by volunteers or teaching assistants there is some evidence that training is beneficial. The evidence is strongest at primary level and for subjects like reading and mathematics. There are fewer studies at secondary level or for other subjects.

What are the costs?

The costs are high as the support is intensive. 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 will cost in the region of £800 per pupil. Costs could be reduced by using groups of one-to-two or one-to-three (see *Small Group Tuition*). Overall, costs are estimated as high.

What do I need to know?

- One-to-one tuition is very effective in helping learners catch up, but can be relatively expensive.
- To control costs, schools could consider other groupings for intensive support such as one-to-two or even one-to-three.
- Short periods (5-10 weeks) of intensive sessions (up to an hour three or four times a week) tend to have greater impact.
- A qualified teacher is likely to achieve greater progress than support staff or volunteers, and training and professional development are likely to be beneficial for both teachers and support staff.
- Pupils and regular class teachers may need support at the end of the tutoring to ensure the impact is sustained once they return to normal classes and tuition should be explicitly linked to what happens in class.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/one-to-one-tuition

Outdoor adventure learning

Moderate impact for moderate cost, based on limited evidence.

£ £ £ £ £	★★☆☆☆	+3
cost per pupil	evidence rating	months

What is it?

Outdoor adventure learning typically involve 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 months additional progress over the course of a year. 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, though on average both older and more recent studies show a positive impact on academic attainment. Our overall assessment of potential progress is therefore weighted towards more recent studies.

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 do I need to know?

- A wide range of adventure activities are linked with increased academic achievement.
- Experiences of over a week tend to have greater impact and tend to produce effects of a longer duration.
- The main effects tend to be on self-confidence, self-efficacy and motivation and for some pupils explicit links may need to be made to ensure lasting impact on academic attainment.
- 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.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/outdoor-adventure-learning

Parental involvement

Moderate impact for moderate cost, based on moderate evidence.

£ £ £ £ £	★ ★ ★ ☆ ☆	+ 3
cost per pupil	evidence rating	months

What is it?

Actively involving 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 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 is also conflicting evidence which suggests that, at least in terms of Early Years Intervention for example, the involvement of parents does not increase the benefits. This suggests that developing effective parental involvement to improve their children's attainment is challenging and will need effective monitoring and evaluation. The impact of parents' aspirations is also important, though again 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 suggest that increasing parental involvement in primary and secondary schools has 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 which 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 60 Pupil Premium allocations. Costs per pupil are therefore estimated as moderate.

What do I need to know?

- Focused approaches which support parents in working with their children to improve their learning are beneficial. The challenge is in engaging and sustaining such involvement.
- Involvement is often easier to achieve with parents of very young children.
- Parents of older children may appreciate short sessions at flexible times to involve them.
- Schools can be daunting places for parents so it is important to establish a welcoming environment.
- Parents may be anxious about their own educational achievements and it is important to discuss with them the ways in which they can support their children's effort which do not require a high level of ability (e.g. by ensuring that students have an environment where they can work at home, or by asking them to explain what they learned at school and how they learned it.).

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/parental-involvement

Peer tutoring

High impact for low cost, based on extensive evidence.

£ £ £ £ £
cost per pupil

★★★★☆
evidence rating

+6
months

What is it?

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?

The evidence of impact is relatively high (typically equating to about a GCSE grade). The benefits are apparent for both tutor and tutee (particularly in cross-age tutoring), though the approach should be used to supplement or enhance normal teaching, rather than to replace it. There is some evidence that children from disadvantaged backgrounds and low attaining pupils make the biggest gains.

Though both pupils involved gain, cross-age tutoring appears to offer slightly greater benefit for tutor than tutee. A study of cross-age peer tutoring showed that the lowest attaining pairs actually made most progress, and a two-year gap seems to support both tutee and tutor learning. One way of matching pupils across classes is to match the highest attaining pupil in the older class with the highest attaining child in the younger class through to the lowest attaining pupil in the older class being matched with the lowest attaining pupil in the younger class (making adjustments if necessary). This enables the teacher to focus support on lower attaining pairs.

How secure is the evidence?

The evidence is consistent and positive especially for mathematics and reading and at both primary and secondary school levels.

What are the costs?

The direct costs of running peer tutoring in schools are low, as few additional materials required (£10-20 per pupil). Professional development and additional support for staff is recommended, particularly in the early stages of setting up a programme. Estimates are about £3,000-£4,000 per class or £200 per pupil indicating low overall costs.

What do I need to know?

- Activities should be sufficiently challenging for the tutee that they can benefit from the tutor's support but not too difficult that they cannot succeed with support.
- Planning the organisation of tutoring to address the logistical challenges and then training the tutors is a key step. At least a day's professional development for staff is recommended.
- There are several different approaches to peer-tutoring which make different demands on the teacher(s) organising the pairs and on the tutors and tutees.
- Reviewing challenges and successes with tutors will improve their skills and learning.
- Relatively short but intensive periods of tutoring over 4-10 weeks are likely to be more effective than for a longer period with more routine sessions.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/peer-tutoring

Performance pay

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

£ £ £ £ £ — cost per pupil

★ ★ ★ ★ ★ — evidence rating

0 months

What is it?

We know that teachers are the most important part of the education system in terms of improving students' learning. Performance pay is where a direct link is created 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. In the USA, it is sometimes referred to as 'merit pay'. One key issue is how performance is measured and how closely this is linked to outcomes for learners. In the UK, performance measurement was one component in the performance threshold assessment introduced in 2000, but was very loosely connected and at the discretion of the head teacher.

How effective is it?

Estimates based on cross-national comparisons suggest that performance pay could lead to positive impacts of around three months, and one UK study estimates the benefit as about half a GCSE grade, which is a similar effect. However, when more rigorous evaluations are looked at, such as those with experimental trials or with well-controlled groups, within countries such as the USA, 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 in the UK.

As the evaluation of a number of merit pay schemes in the USA 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.

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. It is hard to make causal claims about the efficacy of performance pay. 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".

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 do I need to know?

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.

Performance pay has been tried on a number of occasions, however the evidence of impact on student learning does not support the approach.

Evaluations of the English threshold assessment 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.

Spending on professional development linked to evaluation of better learning by pupils may also offer an alternative to performance pay.

Performance pay may lead to a narrower focus on test performance and restrict other aspects of learning.

Phonics

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

£ £ £ £ £ cost per pupil

★★★★☆ evidence rating

+4 months

What is it?

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 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. The approach tends to be more effective 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 only one part of a successful literacy teaching. For older readers (above Year 5) who are struggling, phonics approaches may be less successful, producing less or no impact and other approaches such as comprehension focused methods may be more effective. In particular, using age appropriate material is likely to be more successful. Furthermore upper primary and lower secondary readers may benefit more from strategy instruction or *Meta-cognitive and self-regulation* strategies to improve their reading skills.

The research suggests that phonics is beneficial for younger learners as they begin to read (4-7 year olds). It is less likely to be helpful for older, less successful learners. Qualified teachers tend to get better results (up to twice the effectiveness of others), suggesting that their expertise is a key component of successful teaching of early reading.

How secure is the evidence?

There have been a number of studies, reviews and meta-analyses which have consistently found that the systematic teaching of phonics is beneficial. There is some evidence that particular approaches such as synthetic phonics may be more beneficial than analytic approaches, 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.

What are the costs?

There are some costs, as specific resources are needed for teaching phonics. 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. Costs for materials and professional development are estimated at £1,200 per teacher or £48 per pupil and therefore very low.

What do I need to know?

Phonics can be an important component in supporting the development of early reading skills, particularly for children from disadvantaged backgrounds. However, it is not a panacea and 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.

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.

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).

Phonics improves the accuracy of children's reading, but not necessarily their comprehension, and as such should be included as part of a wider literacy programme.

As a child's reading skills progress and they become successful with a phonics-based approach, the emphasis should move on to developing children's understanding of what they can read.

For more information, videos and supporting resources relating to this approach, please visit http://educationendowmentfoundation.org.uk/toolkit/approaches/phonics

Physical environment

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

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

0
months

What is it?

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 which seeks to change behaviour and establish new norms (similar to 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. So in terms of sound, if the noise levels are high (such as under the flight path of an airport) 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 though most school environments are within acceptable limits. Likewise, 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), and that classrooms often have poor air quality conditions, with higher CO_2 concentrations than the average recommended levels.

How secure is the evidence?

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

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 do I need to know?

- Most environmental factors have an impact on classrooms only at the extremes.
- 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
- Changes in the environment are opportunities to change people's behaviour as they adjust to the new setting, but are unlikely to have a direct positive impact on learning.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/physical-environment

Reducing class size

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

£ £ £ £ £	★★★☆☆	+3
cost per pupil	evidence rating	months

What is it?

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 explanation for this appears to be whether a 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 have been 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 Key Stage 2). It appears to be very hard to achieve improvements from class size reductions above 20, e.g. from 30 to 25.

There is some evidence that reducing class sizes are 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 do I need to know?

- Smaller classes will not make a difference to learning unless the teacher or pupils do something differently in the smaller class.
- It is likely that the more flexible choices the teacher has for organising learners combined with an increase in the quality or quantity of feedback pupils receive accounts for any gains.
- Small reductions (e.g. from 30 to 25 pupils) are unlikely to be cost-effective relative to other strategies.
- Deploying staff (including teaching assistants) so that teachers can work more intensively with smaller groups may be worth exploring.
- Reducing class sizes for younger children may provide longer term benefits.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/reducing-class-size

Repeating a year

Negative impact for very high cost based on extensive evidence.

£ £ £ £ £
cost per pupil

★★★★☆
evidence rating

-4
months

What is it?

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. In England, repeating a year is currently very uncommon and schools cannot require that students repeat a year without parental consent. However, it is included within the Toolkit as it is a policy which periodically attracts some interest among schools and the media.

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.

Although the overall average impact of repeating a year is negative, some studies suggest that in individual circumstances it can benefit the student, particularly in the short term. However, it does not appear to be easy to identify which students will benefit from repeating a year prior to making a decision, suggesting that choosing to do so represents a significant risk.

There are a number of possible explanations for why repeating a year is so ineffective. One is that in its simplest form repeating a year just provides ‘more of the same’, in contrast to other strategies which provide additional targeted support or involve the use of a new pedagogical approach. In addition, it appears that repeating a year is likely to have a negative impact on the student’s self-confidence and belief that they can be an effective learner.

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.

Some more recent meta-analyses using more rigorous designs have found less severe effects (between zero effect and negative 1 month). However, these studies have also been consistent with earlier research in showing that detrimental effects of repeating a year increases over time and that repeating a year has a disproportionately negative effect on pupils from disadvantaged backgrounds. Overall, the evidence is extensive and reasonably consistent and is therefore estimated as strong.

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 do I need to know?

Negative effects are rare for educational interventions, so the extent to which pupils who repeat a year go backwards is striking.

The negative effects are disproportionately greater for disadvantaged pupils, for pupils from ethnic minorities and for children born in the summer months.

Alternative interventions such as intensive tuition or one to one support are considerably cheaper and may make repeating a school year unnecessary. As a result these appear to be better bets in the first instance.

The negative effects tend to increase with time and repeating more than one school year significantly increases the risk of pupils dropping out and not completing their schooling.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/repeating-a-year

School uniform

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

£ £ £ £ £ cost per pupil

★ ★ ★ ★ ★ evidence rating

0 months

What is it?

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 the UK 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 which have information about these outcomes linked to the introduction of a school uniform policy, but this 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 which look at the relationship between schools with uniforms compared with those without or the performance of schools before and after the introduction of uniforms and the school's subsequent trajectory of improvement. 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 do I need to know?

- When combined with the development of a school ethos and the improvement of behaviour and discipline, the introduction or enforcement of a school uniform can be successfully included as part of this process.
- Wearing a uniform is not, on its own, going to improve learning.
- The commitment of staff to uphold and enforce a behaviour policy is crucial to its success.
- Improved behaviour, on its own, does not necessarily lead to better learning, though it may be an important precondition (see Behaviour).

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/school-uniforms

Small group tuition

Moderate impact for moderate cost, based on limited evidence.

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

+4
months

What is it?

Intensive tuition in small groups is usually provided to support lower attaining learners or those who are falling behind, though it can also be used as a more general strategy to ensure effective progress, or to teach challenging topics or skills. The most familiar approach is one teacher with one pupil (see One to one). However other approaches to provide for intensive support are possible, such as teaching pupils in pairs or small groups of 3-5. For the purposes of the Toolkit '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.

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 explain 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. Likewise, in the evaluation of Every Child Counts in the UK, one to one, paired teaching and groups of three were almost equally effective. 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 pupils 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 the outcomes.

Given the closeness in impact between various forms of small group tuition and its much lower cost, it may be useful for schools to trial small group tuition as an alternative option to one to one tuition.

How secure is the evidence?

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 less 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.

What are the costs?

Costs decrease with group size as the majority of the costs are for teaching time. We have estimated the cost of one to two tuition as £400 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 (£270 per pupil). Costs are therefore estimated as moderate.

What do I need to know?

- Intensive tuition in small groups is very effective.
- Pupils are usually grouped according to current level of attainment or specific need.
- It is important to assess pupils' needs accurately and provide work at a challenging level with effective feedback and support.
- The cost effectiveness of one-to-two and one-to-three indicates that greater use of these approaches would be productive in schools.
- Professional development and evaluation are likely to increase the effectiveness of small group tuition.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/small-group-tuition

Social and emotional learning

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

£ £ £ £ £ cost per pupil

★★★★☆ evidence rating

+4 months

What is it?

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. As with behaviour, three broad categories of 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.

SEL interventions seek to improve the ways in which pupils work with and alongside their peers, teachers, family and community. In 2005, a national SEL programme was introduced to support effective learning, positive behaviour, attendance, and emotional well-being, first in primary schools 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 (on average around three to four months additional progress).

However, though SEL interventions almost always improve emotional or attitudinal outcomes, not all interventions are equally effective at raising attainment. In particular, evidence from the nationwide SEL programme introduced in 2005 suggests that benefits on learning will not be automatically achieved. A quasi-experimental evaluation of the impact of the secondary programme did not find a significant impact on attainment in the SEL schools.

Improvements seem 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 were committed to the approach appeared 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.

What are the costs?

Social and emotional interventions targeted at individuals are the most expensive (see also Behaviour interventions). Estimates from the US suggest targeted programs cost about $4,600 per student (about £2,800) per year and involve professional counselling services. However, the costs of training school staff and implementing and evaluating the impact are estimated at £1,000 per teacher for professional development and in-school support. Overall the costs per pupil are therefore estimated as low at about £40 per pupil per year, assuming a school-based, whole class approach.

What do I need to know?

- Skills should be taught purposefully and explicitly linked to direct learning in schools, encouraging pupils to apply the skills they learn.
- Teachers and other school staff can effectively support these approaches, particularly with appropriate professional development.
- Staff commitment to the programme and support for the consistent application of the skills more widely are likely to be important features of successful approaches.
- Sensitive and targeted intervention may benefit at risk or more vulnerable pupils.
- It is important to evaluate the impact of any initiative to improve learning based on social and emotional aspects of learning as the impact on attainment is not found consistently.

For more information, videos and supporting resources relating to this approach, please visit http://educationendowmentfoundation.org.uk/toolkit/approaches/social-and-emotional-aspects-of-learning

Sports participation

Moderate impact for moderate cost based on moderate evidence.

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

+2
months

What is it?

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 as a means 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' acts as an enticement to undertake additional instruction.

The variability in effects suggest 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, including a recent systematic review for the Department for Culture, Media and Sport (DCMS). There is, however, considerable variation in impact, including some studies which show negative effects.

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 do I need to know?

Being involved in extra-curricular sporting activities may increase attendance and retention.

Participation in sports does not straightforwardly transfer to academic learning.

Planned extra-curricular activities which include short regular 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.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/sports-participation

Summer schools

Moderate impact for moderate cost based on limited evidence.

£ £ £ £ £
cost per pupil

★ ★ ★ ★ ★
evidence rating

+3
months

What is it?

Summer schools are lessons or classes during the summer holidays, often run as catch-up or enrichment lessons. Some summer 'schools' do not have an academic focus and concentrate on sports or other non-academic activities. Others may be targeted at either low or high performing students for under-achieving or gifted and talented students.

How effective is it?

The effects are reasonably consistent (with an average impact of about three months progress), though usually higher for higher attaining pupils and less effective for low income pupils. Programmes are usually more effective in mathematics, when they are specifically tailored to students' needs, and when parents are involved, such as by attending conferences with teachers, observing their children in class and reading with them at home. Summer schools which do not have a clear academic component are not usually evaluated for, or associated with, learning gains. Other variables seem to make less difference, such as whether the teacher is one of the student's usual teachers.

The impacts vary according to the focus of the summer school, with more academic benefits linked to those with teaching or tutoring. Benefits have been identified in a range of subjects, particularly for secondary school pupils but are not consistent across all programmes. This indicates that it is important to have a clear focus on learning.

How secure is the evidence?

There are a number of meta-analyses, finding broadly similar effects, though mostly based on studies in the USA. As mentioned, a crucial factor is whether the summer school has an academic focus.

What are the costs?

The costs involved are the employment of teachers for the duration of the summer school, with associated venue and resource costs (books, photocopying etc.). Courses are in the region of £250 per week per student. A two week summer school would cost about £500 per pupil and are therefore estimated as moderate.

What do I need to know?

Summer school provision which aims to improve learning needs to have an academic component.

Qualified and experienced teachers are more likely to support improvement in literacy or mathematics (or other subjects) than less well-qualified staff.

Intensive tutoring (one-to-one or small group) can be productively included in summer school provision.

Summer schools can also provide support for the highly able and transition to university.

As with After School Programmes providing a stimulating environment, teacher's support and promoting interaction appear to increase participation.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/summer-schools

Teaching assistants

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

£ £ £ £ £ — cost per pupil

★ ★ ☆ ☆ ☆ — evidence rating

0 months

What is it?

A teaching assistant (TA) is someone who supports a teacher in the classroom. Their duties can differ dramatically from school to school, though the main tasks tend to be working with small groups of children who need extra support in an area of the curriculum such as literacy or numeracy. They are also often responsible for hearing children read and helping teachers with administrative tasks.

How effective is it?

Overall, research shows 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 recorded impacts; in some cases teachers and teaching assistants have worked together effectively leading to increases in attainment, while in others pupils (particularly those who are lower attaining) have performed worse in classes with teaching assistants present compared to those without.

One clear implication of this surprising finding is that schools should think carefully about the deployment, training (both of the teacher and TA) and evaluation of their TAs if they hope to achieve positive impacts in terms of attainment. Comparisons with qualified teachers suggest that TAs are consistently less effective in terms of raising attainment (achieving about half the gains). It is suggested that where overall negative impacts have been recorded TAs have effectively been substitutes rather than supplementary to teaching from teachers.

There is some evidence of greater impact when TAs are given a well-defined pedagogical role or responsibility for delivering specific interventions, particularly when training and support are provided. Evidence suggests that impact is similar across subjects and at both primary and secondary level. There is also evidence that pupils' perceptions and attitudes may be more positively affected, and also of positive effects in terms of teacher morale and reduced stress of working with a TA.

How secure is the evidence?

There are a number of systematic reviews of the impact of support staff in schools, though there are no meta-analyses specifically looking at the impact of TAs on learning. However, there have been a number of reviews internationally which have consistently found broadly similar effects. The most recent study in the UK suggests that on average low attaining pupils do less well with a TA supporting them. The research literature does not distinguish between different levels or grades of teaching assistants.

What are the costs?

The average teaching assistant's salary is about £17,000 per annum or about half of an average teaching salary. Costs overall are estimated as high.

What do I need to know?

Teaching assistants undoubtedly contribute to the effective management and organisation of a school. On average, however, they do not seem to add to the learning of the children and the classes that they support. More research must be done to determine the best ways for teachers and teaching assistants to work together, but likely best bets include:

- Identifying activities where TAs can support learning, rather than simply managing tasks.
- Providing support and training for teachers and TAs so that they understand how to work together effectively, e.g. by making time for discussion to talk before and after lessons.
- Ensuring that teachers do not reduce their support or input to the pupils supported by TAs that TAs are focused on learning as opposed to just ensuring that pupils finish their work.
- Evaluating the impact of different strategies for deploying TAs.
- Ensuring that teachers do not reduce their support or input to the pupils supported by TAs and that TAs are focused on learning as opposed to just ensuring that pupils finish their work.

For more information, videos and supporting resources relating to this approach, please visit
http://educationendowmentfoundation.org.uk/toolkit/approaches/teaching-assistants

References: Ability grouping

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

Nomi, T. (2009). The Effects of Within-Class Ability Grouping on Academic Achievement in Early Elementary Years, Journal of Research on Educational Effectiveness, 3:1, pp 56-92.

Rogers, K. B. (2007). Lessons Learned About Educating the Gifted and Talented A Synthesis of the Research on Educational Practice. *Gifted Child Quarterly*, 51(4), 382-396.

Slavin R.E. (1990). Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis. Review of Educational Research, 60 (3), 471-499.

Steenbergen-Hu, S., & Moon, S. M. (2011). The effects of acceleration on high-ability learners: A meta-analysis. Gifted Child Quarterly, 55(1), 39-53.

Vaughn, V. L., Feldhusen, J. F., & Asher, J. W. (1991). Meta-analyses and review of research on pull-out programs in gifted education. Gifted Child Quarterly, 35(2), 92-98.

Walsh, R. L., Kemp, C. R., Hodge, K. A., & Bowes, J. M. (2012). Searching for Evidence-Based Practice A Review of the

Research on Educational Interventions for Intellectually Gifted Children in the Early Childhood Years. Journal for the Education of the Gifted, 35(2), 103-128

Summary of effects

Study	*Effect size*
Kulik & Kulik 1982 (on secondary pupils)	
Kulik & Kulik, 1984 (on elementary/primary pupils)	
Lou et al., 1996 (on low attainers)	
Slavin, 1990 (on low attainers)	
Indicative effect size	**-0.0**

For more information about the effect sizes in the Toolkit, click here. (#http://educationendowmentfoundation.org.uk/toolkit/about-the-toolkit/AverageImpact)

Meta-analyses abstracts

Study	*Abstract*
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 honors 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 also showed that ability grouping has trivially small effects on the self-concepts of elementary school pupils.
Kulik & Kulik (1987).	No abstract available.
Kulik & Kulik,	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

(1992). 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.

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, S., & Moon, S. M. (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

References: After school programmes

Full references

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., 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.

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).

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.

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.

Posner, J.K. & Vandell, D.L. (1994). Low-Income Children's After-School Care: Are There Beneficial Effects of After-School Programs? Child Development, 65.2 pp 440-456.

Scott-Little, C., Hamann, M.S. & Jurs, S.G. (2002). Evaluations of After-School Programs: A Meta-Evaluation of Methodologies and Narrative Synthesis of Findings. American Journal of Evaluation, 23.4 pp 387-419.

Woodland, M.H. (2008). Whatcha Doin' After School? A Review of the Literature on the Influence of After-School Programs on Young Black Males. Urban Education 43.5 pp 537-560.

Zief, S.G., Lauver, S., Maynard, R.A. (2006). Impacts of After-School Programs on Student Outcomes Oslo: The Campbell Collaboration.

Summary of effects

Study	*Effect size*
Crawford, 2011	0.40
Durlak & Weissberg, 2007	0.16
Fashola, 1998	NPE
Lauer, Akiba & Wilkerson, 2006	0.07
Scott-Little et al., 2002	NPE
Zief et al. 2006 (on reading)	0.03
Zief et al. 2006 (on GPA)	0.08
Indicative effect size	**0.10**

Meta-analyses

abstracts

Study	*Abstract*
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.
	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,

Fashola, (1998).	• 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.
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).	No abstract provided.

References: Arts participation

Full references

Lewis C.P. (2004) The Relation between Extracurricular Activities with Academic and Social Competencies in School Age Children: A Meta-Analysis. College Station: Texas A&M University.

Newman M., Bird K., Tripney J., Kalra N., Kwan I., Bangpan M., & Vigurs C (2010a). 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.

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.

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.

Summary of effects

Study	*Effect size*
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
Standley, 2008	0.32
Winner & Cooper, 2000 (maths)	0.04
Indicative effect size	**0.15**

Meta-analyses abstracts

Study	*Abstract*
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. (2010a). No abstract provided.

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 behavior (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). No abstract provided.

References: Aspiration interventions

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 www.jrf.org.

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.

Gorard, S., See, B. H., & Davies, P. (2012). The impact of attitudes and aspirations on educational attainment and participation. York: Joseph Rowntree Foundation www.jrf.org.

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 http://www.nationalserviceresources.org/online-library/items/m3588.

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 www.jrf.org.

Strand, S., & Winston, J. (2008). Educational aspirations in inner city schools. Educational Studies, 34(4), 249-267.

Summary of effects

Study	*Effect size*
No meta-analyses of impact of raising aspirations on learning outcomes.	
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 0.032-0.092 for extra-curricular activities. However these effects are associated with direct influences on learning such as parental involvement in reading or academic mentoring.	
Indicative effect size	**0.00**

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.

References: Behaviour interventions

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.

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)
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)
Indicative effect size	**0.32**

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-analyzed. 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 behaviors, 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 behavioral 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 behaviors, irrationality, self-concept, and grade point average. In terms of magnitude, the largest positive mean effect of REBT was on disruptive behaviors. 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.
	Many programs designed for youth with Emotional or Behavioral 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

Quinn et al. (1999). 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 analyzed 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 behavior 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 behavior were obtained. Effect sizes and study characteristics were coded from these studies and analyzed. Positive overall intervention effects were found on aggressive and disruptive behavior 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 behavior 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., behavioral, 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.

References: Block scheduling

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.

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.

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.

Veal, W.R. & Flinders, D.J. (2001). How Block Scheduling Reform Effects Classroom Practice. The High School Journal, 84.4 pp 21-31.

Zepeda, S.J. & Mayers, R.S. (2006). An Analysis of Research on Block Scheduling. Review of Educational Research, 76.1 pp 137-170.

Summary of effects

Study	*Effect size*
Dickson et.al, 2010	achievement: 0.11, mathematics: -0.02, science: 0.20
Lewis et.al, 2005	mathematics: -0.10, English: -0.17, science: -0.12
Indicative effect size	**0.00**

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.

References: Collaborative learning

Full references

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.

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.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). Cooperative learning methods: A meta-analysis. Minneapolis, MN: University of Minnesota. http://www.tablelearning.com/uploads/File/EXHIBIT-B.pdf.

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.

Summary of effects

Study	*Effect size*
Romero, 2009	0.40
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
Indicative effect size	**0.42**

Meta-analyses abstracts

Study *Abstract*

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

Romero, (2009).	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. 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.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
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.

References: Digital technology

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.

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.

Morphy P. & Graham S. (2012). Word Processing Programs and Weaker Writers/Readers: A Meta-Analysis of Research Findings, Reading and Writing, 25, 641-678.

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.

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

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.

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.

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**

Meta-analyses abstracts

Study	*Abstract*
Bayraktar (2001).	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.
Blok et al. (2002).	How effective are computer-assisted instruction (CAI) programs in supporting beginning readers? This article reviews 42 studies published from 1990 onward, comprising a total of 75 experimental comparisons. The corrected overall effect size estimate was d = 0.19 (± 0.06). Effect sizes were found to depend on two study characteristics: the effect size at the time of pre-testing and the language of instruction (English or other). These two variables accounted for 61 percent of the variability in effect sizes. Although an effect size of d = 0.2 shows little promise, caution is needed because of the poor quality of many studies.
Camnalbur & Erdogan (2008).	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.
Cassil (2005).	Statistical meta analyses performed for this study included 32 primary studies conducted between 1993–2005. Two independent meta analyses were conducted regarding student attitudes and academic outcomes. The overall meta analysis mean by author was .23, indicating that student use of mobile computers had a small and positive effect on student attitudes and academic outcomes. The consistent pattern of positive effect size results indicated that student use of mobile computers was effective in improving student attitudes and academic outcomes. The small number of samples in the independent meta analyses suggests a need for further research

	regarding mobile computers.
Cheung & Slavin (2011).	A total of 75 qualifying studies were included in our final analysis with a total sample size of 56,886 K-12 students: 45 elementary studies (N=31,555) and 30 secondary studies (N=25,331). The overall weighted effect size is +0.15. Types of intervention. With regards to intervention types, the studies were divided into three major categories: Computer-Managed Learning (CML) (N=7), Comprehensive Models (N=8), and Supplemental CAI Technology (N=37). Over 70% of all studies fell into the supplemental program category, which consists of individualized computer-assisted instruction (CAI). These supplemental CAI programs, such as Jostens, PLATO, Larson Pre-Algebra, and SRA Drill and Practice, provide additional instruction at students' assessed levels of need to supplement traditional classroom instruction. Computer-managed learning systems included only Accelerated Math, which uses computers to assess students' mathematics levels, assign mathematics materials at appropriate levels, score tests on this material, and chart students' progress. One of the main functions of the computer in Accelerated Math is clerical (Niemiec et al., 1987). Comprehensive models, such as Cognitive Tutor and I Can Learn, use computer assisted instruction along with non-computer activities as the students' core approach to mathematics.
Christmann & Badgett (2003).	This meta-analysis compared the academic achievement of elementary students who received either traditional instruction or traditional instruction supplemented with CAI. From the 68 effect sizes, an overall mean effect size of 0.342 was calculated, indicating that, on average, students receiving traditional instruction supplemented with CAI attained higher academic achievement than did 63.31% of those receiving only traditional instruction. However, a -0.463 correlation between effect size and years indicates that the effect of CAI on academic achievement has declined between the years 1969 and 1998.
Goldberg et al. (2003).	Meta-analyses were performed including 26 studies conducted between 1992–2002 focused on the comparison between k–ı2 students writing with computers vs. paper-and-pencil. Significant mean effect sizes in favor of computers were found for quantity of writing (d=.50, n=14) and quality of writing (d=.41, n=15). Studies focused on revision behaviours between these two writing conditions (n=6) revealed mixed results. Other studies collected for the meta-analysis which did not meet the statistical criteria were also reviewed briefly. These articles (n=35) indicate that the writing process is more collaborative, iterative, and social in computer classrooms as compared with paper-and-pencil environments. For educational leaders questioning whether computers should be used to help students develop writing skills, the results of the meta-analyses suggest that, on average, students who use computers when learning to write are not only more engaged and motivated in their writing, but they produce written work that is of greater length and higher quality.
Graham & Perin (2007)	There is considerable concern that the majority of adolescents do not develop the competence in writing they need to be successful in school, the workplace, or their personal lives. A common explanation for why youngsters do not write well is that schools do not do a good job of teaching this complex skill. In an effort to identify effective instructional practices for teaching writing to adolescents, the authors conducted a meta-analysis of the writing intervention literature (Grades 4–12), focusing their efforts on experimental and quasi-experimental studies. They located 123 documents that yielded 154 effect sizes for quality of writing. The authors calculated an average weighted effect size (presented in parentheses) for the following 11 interventions: strategy

	instruction (0.82), summarization (0.82), peer assistance (0.75), setting product goals (0.70), word processing (0.55), sentence combining (0.50), inquiry (0.32), prewriting activities (0.32), process writing approach 0.32), study of models (0.25), grammar instruction (– 0.32).
Kulik (2003).	This report reviews findings from controlled evaluations of technology applications in elementary and secondary schools published since 1990 located through computer searches of library databases… and summarises reviews of studies published before 1990.
LeJeune (2002).	The purpose of this study was to synthesize the findings from existing research on the effects of computer simulated experiments on students in science education. Results from 40 reports were integrated by the process of meta-analysis to examine the effect of computer-simulated experiments and interactive videodisc simulations on student achievement and attitudes. Findings indicated significant positive differences in both low-level and high-level achievement of students who use computer-simulated experiments and interactive videodisc simulations as compared to students who used more traditional learning activities. No significant differences in retention, student attitudes toward the subject, or toward the educational method were found. Based on the findings of this study, computer-simulated experiments and interactive videodisc simulations should be used to enhance students' learning in science, especially in cases where the use of traditional laboratory activities are expensive, dangerous, or impractical.
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.
Liao (2005).	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. 52 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.
	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

Lou et al. (2001).	(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.The results of Hierarchical Regression Model development indicate that the effects of small group learning with CT on individual achievement were significantly larger when: (a) students had group work experience or specific instruction for group work rather than when no such experience or instruction was reported; (b) cooperative group learning strategies were employed rather than general encouragement only or individual learning strategies were employed; (c) programs involved tutorials or practice or programming languages rather than exploratory environments or as tools for other tasks; (d) subjects involved social sciences or computer skills rather than mathematics, science, reading, and language arts; (e) students were relatively low in ability rather than medium or high in ability; and (f) studies were published in journals rather than not published. When all the positive conditions were present, students learning in small groups could achieve 0.66 standard deviation more than those learning individually. When none of the positive conditions were present, students learning individually could learn 0.20 standard deviation more than those learning in groups.
Means et al. (2009).	A systematic search of the research literature from 1996 through July 2008 identified more than a thousand empirical studies of online learning. Analysts screened these studies to find those that (a) contrasted an online to a face-to-face condition, (b) measured student learning outcomes, (c) used a rigorous research design, and (d) provided adequate information to calculate an effect size. As a result of this screening, 51 independent effects were identified that could be subjected to meta-analysis. The meta-analysis found that, on average, students in online learning conditions performed better than those receiving face-to-face instruction. The difference between student outcomes for online and face-to-face classes—measured as the difference between treatment and control means, divided by the pooled standard deviation—was larger in those studies contrasting conditions that blended elements of online and face-to-face instruction with conditions taught entirely face-to-face. Analysts noted that these blended conditions often included additional learning time and instructional elements not received by students in control conditions. This finding suggests that the positive effects associated with blended learning should not be attributed to the media, per se. An unexpected finding was the small number of rigorous published studies contrasting online and face-to-face learning conditions for K–12 students. In light of this small corpus, caution is required in generalizing to the K–12 population because the results are derived for the most part from studies in other settings (e.g., medical training, higher education). Few rigorous research studies of the effectiveness of online learning for K–12 students have been published. A systematic search of the research literature from 1994 through 2006 found no experimental or controlled quasi-experimental studies comparing the learning effects of online versus face-to-face instruction for K–12 students that provide sufficient data to compute an effect size. A subsequent search that expanded the time frame through July 2008 identified just five published studies meeting meta-analysis criteria.
	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 for middle school students demonstrate that technology can have a positive effect on reading comprehension (weighted effect size of 0.489). Very little research has focused on the effect of technology on

Moran et al. (2008).	other important aspects of reading, such as metacognitive, affective, and dispositional outcomes. The evidence permits the conclusion that there is reason to be optimistic about using technology in middle-school literacy programs, but there is even greater reason to encourage the research community to redouble its efforts to investigate and understand the impact of digital learning environments on students in this age range and to broaden the scope of the interventions and outcomes studied.
Morphy & Graham (2012).	Since its advent word processing has become a common writing tool, providing potential advantages over writing by hand. Word processors permit easy revision, produce legible characters quickly, and may provide additional supports (e.g., spellcheckers, speech recognition). Such advantages should remedy common difficulties among weaker writers/readers in grades 1–12. Based on 27 studies with weaker writers, 20 of which were not considered in prior reviews, findings from this meta-analysis support this proposition. From 77 independent effects, the following average effects were greater than zero: writing quality (d = 0.52), length (d = 0.48), development/organization of text (d = 0.66), mechanical correctness (d = 0.61), motivation to write (d = 1.42), and preferring word processing over writing by hand (d = 0.64). Especially powerful writing quality effects were associated with word processing programs that provided text quality feedback or prompted planning, drafting, or revising (d = 1.46), although this observation was based on a limited number of studies (n = 3).
Onuoha (2007).	The purpose of this research study was to determine the overall effectiveness of computer-based laboratory compared with the traditional hands-on laboratory for improving students' science academic achievement and attitudes towards science subjects at the college and pre-college levels of education in the United States. Meta-analysis was used to synthesis the findings from 38 primary research studies conducted and/or reported in the United States between 1996 and 2006 that compared the effectiveness of computer-based laboratory with the traditional hands-on laboratory on measures related to science academic achievements and attitudes towards science subjects. The 38 primary research studies, with total subjects of 3,824 generated a total of 67 weighted individual effect sizes that were used in this meta-analysis. The study found that computer-based laboratory had small positive effect sizes over the traditional hands-on laboratory (ES = +0.26) on measures related to students' science academic achievements and attitudes towards science subjects (ES = +0.22). It was also found that computer-based laboratory produced more significant effects on physical science subjects compared to biological sciences (ES = +0.34, +0.17).
Rosen & Salomon (2007).	Different learning environments provide different learning experiences and ought to serve different achievement goals. We hypothesized that constructivist learning environments lead to the attainment of achievements that are consistent with the experiences that such settings provide and that more traditional settings lead to the attainments of other kinds of achievement in accordance with the experiences they provide. A meta-analytic study was carried out on 32 methodologically-appropriate experiments in which these 2 settings were compared. Results supported 1 of our hypotheses showing that overall constructivist learning environments are more effective than traditional ones (ES = .460) and that their superiority increases when tested against constructivist-appropriate measures (ES = .902). However, contrary to expectations, traditional settings did not differ from constructivist ones when traditionally-appropriate measures were used. A number of possible interpretations are offered among them the possibility that traditional settings have come to incorporate some constructivist elements. This possibility is supported by

	other findings of ours such as smaller effect sizes for more recent studies and for longer lasting periods of instruction.
Sandy-Hanson (2006).	Meta-analytical research has shown that computer technology can play a significant role in increasing positive learning outcomes of students. Research on this topic has resulted in conflicting findings on academic achievement and other measures of student outcomes. The current meta-analysis sought to assess the level of differences that existed between students being instructed with computer technology versus the academic achievement outcomes of students instructed with traditional methods. Based on specified selection criteria, 31 studies were collected and analyzed for homogeneity. From this original group, 23 studies were systematically reviewed under standard meta-analytical procedures. According to Cohen's (1988) classification of effect sizes in the field of education, the obtained weighted mean effect size of .24 shows a medium difference. This finding indicates that students who are taught with technology outperform their peers who are taught with traditional methods of instruction. In addition, five secondary analyses were conducted on higher-order thinking skills, ES = .82, motivation, ES = .17, retention-attendance rates, ES = .16, physical outcomes, no data were found, and social skills, ES = .21. Eleven ancillary analyses were then conducted to assess study findings across various dimensions including duration of study, type of technology used, and grade-level analyzed.
Seo & Bryant (2009).	The purpose of this study was to conduct a meta-study of computer-assisted instruction (CAI) studies in mathematics for students with learning disabilities (LD) focusing on examining the effects of CAI on the mathematics performance of students with LD. This study examined a total of 11 mathematics CAI studies, which met the study selection criterion, for students with LD at the elementary and secondary levels and analyzed them in terms of their comparability and effect sizes. Overall, this study found that those CAI studies did not show conclusive effectiveness with relatively large effect sizes. The methodological problems in the CAI studies limit an accurate validation of the CAI’s effectiveness. Implications for future mathematics CAI studies were discussed.
Sisson (2008).	There has been contradictory evidence concerning the validity of auditory temporal processing deficits as a cause for reading and language problems. In spite of the controversy, Merzenich and Tallal helped develop a popular computer-based intervention, Fast ForWord (Scientific Learning Corporation [SLC], 2006). Although a variety of studies have examined the effectiveness of FFW on academic performance, the findings have been inconsistent, creating the need to quantitatively synthesize findings of experimental studies on Fast ForWord. Thirty-one studies met the stipulated inclusion criteria, which generated 163 effect sizes aggregated across academic skills (e.g., reading, language, phonological processing). The overall mean effect size was in the small to medium range, and no particular reading, language, or phonological processing skill appeared to be significantly more responsive to FFW than another skill. All mean effect sizes were associated with sizable variability, often equal to or exceeding effect size, which decreased the confidence one could place in the "true" effect of FFW. Aggregations were also made across moderator variables (e.g., grade, ethnicity, diagnostic category). This paper provides supporting evidence on the need for the study, a review of the related auditory temporal processing literature, and the purpose, procedure, and findings of the meta-analysis.
Soe et al.	Whether computer-assisted instruction (CAI) can improve reading achievement of students has been a crucial question addressed by studies in the past. This meta-analysis reviewed 17

(2000).	research studies based on students K-12 and revealed that CAI does have a positive effect on reading achievement. Although the effects of CAI in 17 studies were not homogeneous, there seems to be no particular study characteristic that might have caused the heterogeneity.
Strong et al. (2011).	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: M-eta-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.
Tokpah (2008).	This meta-analysis sought to investigate the overall effectiveness of computer algebra systems (CAS) instruction, in comparison to non-CAS instruction, on students' achievement in mathematics at pre-college and post-secondary institutions. The study utilized meta-analysis on 31 primary studies (102 effect sizes, N= 7,342) that were retrieved from online research databases and search engines, and explored the extent to which the overall effectiveness of CAS was moderated by various study characteristics. The overall effect size, 0.38, was significantly different from zero. The mean effect size suggested that a typical student at the 50th percentile of a group taught using non-CAS instruction could experience an increase in performance to the 65th percentile, if that student was taught using CAS instruction. The fail-safe N, Nfs, hinted that 11,749 additional studies with nonsignificant results would be needed to reverse the current finding. Three independent variables (design type, evaluation method, and time) were found to significantly moderate the effect of CAS. The current results do not predict future trends on the effectiveness of CAS; however, these findings suggest that CAS have the potential to improve learning in the classroom. Regardless of how CAS were used, the current study found that they contributed to a significant increase in students' performance.
Torgerson & Zhu (2003).	What is the evidence for the effectiveness of ICT on literacy learning in English, 5-16? Studies were retrieved from the three electronic databases. PsycInfo and ERIC were the richest sources for retrieving RCTs for this review. 5.1.2 Mapping of all included studies Forty-two RCTs were identified for the effectiveness map. 5.1.3 Nature of studies selected for effectiveness in-depth review The 12 included RCTs were assessed as being of 'medium' or 'high' quality in terms of internal quality: 'high' quality in terms of relevance to the review; 'medium' or 'high' in terms of the relevance of the topic focus; and 'medium' or' high' for overall weight of evidence. All 12 studies were undertaken in the USA with children aged between 5 and 14. Seven of the RCTs included samples where all or half of the participants experienced learning disabilities or difficulties or specific learning disabilities. All 12 studies focused on the psychological aspects or representations of literacy.

Torgerson and Elbourne (2002).

Recent Government policy in England and Wales on Information and Communication 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 =-0.02 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.

Vogel et al. (2006).

Substantial disagreement exists in the literature regarding which educational technology results in the highest cognitive gain for learners. In an attempt to resolve this dispute, we conducted a meta-analysis to decipher which teaching method, games and interactive simulations or traditional, truly dominates and under what circumstances. It was found that across people and situations, games and interactive simulations are more dominant for cognitive gain outcomes. However, consideration of specific moderator variables yielded a more complex picture. For example, males showed no preference while females showed a preference for the game and interactive simulation programs. Also, when students navigated through the programs themselves, there was a significant preference for games and interactive simulations. However, when teachers controlled the programs, no significant advantage was found. Further, when the computer dictated the sequence of the program, results favored those in the traditional teaching method over the games and interactive simulations. These findings are discussed in terms of their implications for exiting theoretical positions as well as future empirical research.

Waxman et al. (2002).

To estimate the effects of teaching and learning with technology on students' cognitive, affective, and behavioral outcomes of learning, 138 effect sizes were calculated using statistical data from 20 studies that contained a combined sample of approximately 4,400 students. The mean of the study-weighted effect sizes averaging across all outcomes was .30 (p < .05), with a 95-percent confidence interval (CI) of .004 - .598. This result indicates that teaching and learning with technology has a small, positive, significant (p < .05) effect on student outcomes when compared to traditional instruction. The mean study-weighted effect size for the 13 comparisons containing cognitive outcomes was .39, and the mean study-weighted effect size for the 60 comparisons that focused on student affective outcomes was .208. On the other hand, the mean study-weighted effect size for the 30 comparisons that contained behavioral outcomes was -.154, 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.

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

Waxman et al. (2003).

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.

References: Early years intervention

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.

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 pp 1-26.

Barnett, W.S. (2011). Effectiveness of Early Educational Intervention Science 333, pp 975-978.

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.

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 pp 684-698.

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.

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. doi: 10.1521/scpq.16.1.9.19163.

Karoly, L., Kilburn, R., Cannon, J.S. (2005). Early Childhood Interventions: Proven Results, Future Promise. Rand Corporation. Available at: http://www.rand.org/pubs/monographs/2005/RAND_MG341.pdf (30/6/12).

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 pp 443-484.

Lewis, R.J. & Vosburgh, W.T. (1988). Effectiveness of Kindergarten Intervention Programs: A Meta-Analysis. School Psychology International, 9(4), 265-275.

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 Children and Youth Services Review 32, 506–519.

Melhuish, E., Belsky, J., Leyland, A.H., Barnes, J (2008). The Impact of Sure Start Local Programmes on Three Year Olds and Their Families: The National Evaluation of Sure Start NESS Research Report: 2008/FR/027 London: University of London, Institute for the Study of Children, Families and Social Issues.

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.

Nelson, G., Westhues, A. & MacLeod, J. (2003). A Meta-analysis of Longitudinal Research on Preschool Prevention Programs for Children. Prevention and Treatment, 6(2), 46-50.

The National Evaluation of Sure Start (NESS) Team (2010). The impact of Sure Start Local Programmes on five year olds and their families Research Report DFE-RR067 London: Department for Education.

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Elliot, K. (2004). The Effective Provision of Pre-School Education [EPPE] Project: Technical Paper 8a Measuring the Impact of Pre-School

on Children's Cognitive Progress over the Pre-School Period London: Institute of Education, London University.
Sylva, K. and MacPherson, K. (2002). Parents and Parenting in the Early Years: Research evidence. Oxford: Department of Educational Studies, University of Oxford.
Sylva, K. and Pugh, G. (2005). Transforming the early years in England Oxford Review of Education 31.1 pp 11-27.

Summary of effects

Study	*Effect size*
Anderson et.al, 2003	0.35
Camilli, Vargas, Ryan & Barnett, 2010	0.23
Gilliam & Zigler, 2000	NPE
Gorey, 2001 (estimate on long term impact)	0.55
Karoly, Kilbirn & Cannon, 2005	0.28
LaParo & Pianta, 2000	0.51
Lewis & Vosburgh, 1988	0.41
Manning et al. 2010 (on adolescent education)	0.53
Nelson et.al 2003	0.52
Indicative effect size	**0.45**

Meta-analyses abstracts

Study	*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.
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).	No abstract provided.
	School readiness screening is prevalent throughout the US. 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

La Paro & Pianta, (2000).	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 form 70 longitudinal studies that reported correlations between academic/cognitive and social/behavioral 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 outcomes. Effect sizes varied considerably across individual studies and samples. Findings are discussed in terms of assessment and conceptualization of school readiness, the role of school and classroom experiences in contributing 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

Nelson, et.al., (2003).

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.

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.

References: Extended school time

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.

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.

Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the School Day or School Year A Systematic Review of Research (1985–2009). Review of Educational Research, 80(3), 401-436.

Pennington, H. (2007) The Massachusetts Expanding Learning Time To Support Student Success Initiative Washington DC: Center for American Progress http://zdvlc3test.techprogress.org/issues/2007/01/pdf/MALearningTimeReport.pdf.

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.

Symonds, J. E., & Hagell, A. (2011). Adolescents and the organisation of their school time: a review of changes over recent decades in England. Educational Review, 63(3), 291-312.

Summary of effects

Study	*Effect size*
Cooper et al., 2003 (district level comparison)	0.06 (with comparison group) 0.11 (with well matched controls)
Baker et.al. 2004 (international comparison)	0.12 (maths in the UK)
Indicative effect size	**0.11**

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. 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

Baker et.al. (2004). 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.

References: Feedback

Full references

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.

Bennett, R.E. (2011). Formative assessment: a critical review, Assessment in Education: Principles, Policy & Practice, 18: 1, 5-25.

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.

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability 21 1, pp 5-31.

Black P. & Wiliam, D. (1998). Assessment and classroom learning, Assessment in Education, 5, pp. 7-73.

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.

Hattie, J. and Timperley, H. (2007). The Power of Feedback. Review of Educational Research 77.1 pp 81–112.

Kluger, A.N. & De Nisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin, 119.2 pp 254-284.

Kingston, N. & Nash, B. (2011). Formative Assessment: A Meta-Analysis and Call for Research. Educational Measurement: Issues and Practice. 30.4 pp 28-37.

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.

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.

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.

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.

Walberg, H.J. (1982). What makes schooling effective? Contemporary Education Review, 1, 1–34.

Summary of effects

Study	*Effect size*
Bangert-Drowns et al., 1991	0.26
Fuchs & Fuchs, 1986	0.72
Kingston & Nash, 2011 (AfL)	0.20
Kluger & DeNisi, 1996	0.41
Lysakowski & Walberg, 1989	0.97
Tenenbaum & Goldring, 1989	0.72
Walberg, 1982	0.81

Indicative effect size **0.62**

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 pretests 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 behavior 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.
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.
	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

Kluger & De Nisi (1996).	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.
Walberg (1982).	Abstract not available.

References: Homework (Primary)

Full references

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.

Farrow, S., Tymms, P. & Henderson, B. (1999). Homework and attainment in primary schools. British Educational Research Journal, 25.3 pp 323-341.

Muhlenbruck, L., Cooper, H., Nye, B., & Lindsay, J.J. (1999). Homework and achievement: Explaining the different strengths of relation at the elementary and secondary school levels. Social Psychology of Education, 3, 295-317.

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.

Trautwein, U. (2007). The homework-achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. Learning and Instruction, 17, 372–388.

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.

Summary of effects

Study	*Effect size*
Cooper et al., 2006	0.60 (all effects)
Paschal et al., 1984	0.36 (all effects)
Indicative effect size	**0.07 (primary only)**

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).

References: Homework (Secondary)

Full references

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.

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.

Farrow, S., Tymms, P., & Henderson, B. (1999). Homework and attainment in primary schools, British Educational Research Journal, 25(3), 323-341.

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.

Trautwein, U. (2007). The homework-achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. Learning and Instruction, 17, 372–388.

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.

Tymms, P. B. and C. T. Fitz-Gibbon (1992). The relationship of homework to A-level results. Educational Research, 34(1): 3-19.

Summary of effects

Study	*Effect size*
Cooper et al., 2006	0.60
Paschal et al., 1984	0.36
Indicative effect size	**0.44**

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. 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

Paschal et.al. (1984). 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).

References: Individualised instruction

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**

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 analyzed 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.
	This meta-analytic synthesis of findings from 51 studies indicated that use of an individualized

Bangert et.al. (1983). 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.

References: Learning styles

Full references

Arnold, R.D. (1968). Four Methods of Teaching Word Recognition to Disabled Readers. The Elementary School Journal, 68. 5 pp. 269-274.

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.

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.

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.

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).

Mayer, R.E. (2011). Does styles research have useful implications for educational practice? Learning and Individual Differences 21 pp 319-320.

Martin, S. (2010). Teachers using learning styles: Torn between research and accountability? Teaching and Teacher Education 26 pp 1583-1591.

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.

Roberts R.W. & Coleman J.C. (1958). An Investigation of the Role of Visual and Kinesthetic Factors in Reading Failure. The Journal of Educational Research, 51. 6 pp. 445-451.

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).

Summary of effects

Study	*Effect size*
Kavale & Forness, 1987	0.14
Garlinger & Frank, 1986	-0.03
Lovelace, 2005	0.67
Slemmer, 2002	0.13
Indicative effect size	**0.10**

Meta-analyses abstracts

Study	*Abstract*
Kavale &	A literature search identified 39 studies assessing modality preferences and modality teaching. The

Forness, (1987). 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). The author performed a quantitative synthesis of experimental research conducted between 1980 and 2000, in which the Dunn and Dunn Learning-Style Model (R. Dunn & K. Dunn, 1993, 1999) was used. Of the 695 citations elicited by the database and reference section searches, 76 original research investigations met the established inclusion criteria.The 7,196 participants provided 168 individual effect sizes. The mean effect-size values calculated and interpreted through this meta-analysis provided evidence for increased achievement and improved attitudes when responsive instruction was available for diagnosed learning-style preferences.Three indicators rejected homogeneity for achievement and attitude effect sizes. Mean effect sizes for achievement and attitude were variable enough to be described as heterogeneous. Therefore, the author searched for variables that moderated the effect sizes; 6 were found. The author compared this investigation and a previous meta-analysis conducted by M. Sullivan (1993) and reported in The Journal of Educational Research (R. Dunn, S. A. Griggs, J. Olson, B. Gorman, & M. Beasley, 1995) and in the National Forum of Applied Educational Research Journal (M. Sullivan, 1996–1997). Mean effect-size results for achievement from the present and previous meta-analyses were consistent. The author suggested that, on average, learning-styles responsive instruction increased the achievement or improved the attitudes toward learning, or both, of all students. Although several moderating variables influenced the outcome, results overwhelmingly supported the position that matching students' Learning-style preferences with complementary instruction improved academic achievement and student attitudes toward learning. The Dunn and Dunn model had a robust moderate to large effect that was 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.

References: Mastery learning

Full references

Bangert, R. L., Kulik, J. A., & Kulik, C. L. C. (1983). Individualized systems of instruction in secondary schools. Review of Educational Research, 53(2), 143-158.

Guskey, T. R., & Pigott, T. D. (1988). Research on group-based mastery learning programs: A meta-analysis. The Journal of Educational Research, 197-216.

Guskey, T.R. (2007). Closing Achievement Gaps: Revisiting Benjamin S. Bloom's "Learning for Mastery. Journal of Advanced Academics. 19, 8-31.

Hattie, J. (2008). Visible Learning. London: Routledge.

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.

Kulik, J. A., Kulik, C. L. C., & Cohen, P. A. (1979). A meta-analysis of outcome studies of Keller's personalized system of instruction. American Psychologist, 34(4), 307-318.

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.

Miles, K. (2010). Mastery Learning and Academic Achievement. Ed.D. Dissertation, Walden University Document URL ProQuest Dissertations & Theses: http://search.proquest.com/docview/193327442.

Summary of effects

Study	*Effect size*
Guskey & Piggott, (1988)	0.60
Kulik, Kulik & Bangert-Drowns, (1990)	0.52
Bangert, Kulik & Kulik, (1983).	0.10
Waxman et.al. (1985).	0.39
Indicative effect size	**0.40**

Meta-analyses abstracts

Study	*Abstract*
Guskey & Piggott, (1988).	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. A meta-analysis of findings from 108 controlled evaluations showed that mastery learning programs

Kulik, Kulik & Bangert-Drowns, (1990). 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 increases student time on instructional tasks. In addition, self-paced mastery programs often reduce the completion rates in college classes.

Bangert, Kulik & Kulik, (1983). This meta-analytic synthesis of findings from5 1 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.

Kulik, Kulik & Cohen, (1979). Meta-analysis, the application of statistical methods to results from a large collection of individual studies, may prove useful to social scientists trying to draw reliable and general conclusions from a diverse and voluminous literature. This article describes a meta-analysis of 75 comparative studies of an innovative method of college teaching, Keller's personalized system of instruction (PSI). The analysis establishes that PSI generally produces superior student achievement, less variation in achievement, and higher student ratings in college courses, but does not affect course withdrawal or student study time in these courses. The analysis also shows that PSI's superiority can be demonstrated in a variety of course settings with a number of different research designs. Certain settings and research designs, however, produce especially sharp differences between PSI and conventional courses.

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 adaptive ness, student outcomes, social contexts and methodological rigor of the studies.

References: Mentoring

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.

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.

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.

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.

Wheeler, M. E., Keller, T. E., & DuBois, D. L. (2010). Review of Three Recent Randomized Trials of School-Based Mentoring. Social Policy Report, 24(3),1-27.

Wood, S., & Mayo-Wilson, E. (2012). School-Based Mentoring for Adolescents A Systematic Review and Meta-Analysis. Research on Social Work Practice, 22(3), 257-269.

Summary of effects

Study	*Effect size*
Bernstein et al., 2009	0.05 (maths) -0.04 (reading) -0.03 (science)
DuBois et.al., 2002	0.11 (academic)
Wheeler, Keller & DuBois. 2010	-0.02 (maths) -0.01 (reading)
Wood & Mayo-Wilson, 2012	-0.01 (academic performance)
Indicative effect size	**0.05**

Meta-analyses 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.
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 endeavor 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, behavior, 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.

References: Meta-cognition and self-regulation

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.

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.

Dignath, C, Buettner,G and 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.

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.

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.

Klauer, K.J. & Phye, G.D. (2008). Inductive Reasoning: A Training Approach. Review of Educational Research, 78.1 pp 85-123.

Schunk, D.H. (2008). Metacognition, self-regulation, and self-regulated learning: Research recommendations. Educational Psychology Review, 20.4 pp 463-467.

Summary of effects

Study	*Effect size*
Abrami et.al., 2008	0.34
Chiu, 1998	0.67
Dignath et.al., 2008	0.62
Haller et.al., 1988	0.71
Higgins et.al., 2005	0.62
Klauer & Phye, 2008	0.69
Indicative effect size	**0.62**

Meta-analyses abstracts

Study *Abstract*

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

Abrami et.al. (2008).	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.
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). No abstract provided!

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.

References: One to one tuition

Full references

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.

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: pp. 237-248.

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.

Jun, S.W., Ramirez, G., & Cumming, A. (2010). Tutoruing Adolescents in Literacy: A Meta-Analysis. Journal of Education, 45 (2), 219-238.

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.

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.

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.

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.

Summary of effects

Study	*Effect size*
Cohen, Kulik and Kulik, 1982 (on tutees)	0.40
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
Wasik & Slavin, 1993	NPE
Indicative effect size	**0.44**

Meta-analyses abstracts

Study	*Abstract*
Cohen et.al.	A meta-analysis of findings from 65 independent evaluations of school 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-

(1982). 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.

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.

Wasik & Slavin (1993). This article reviews research on one-to-one tutoring models that have been used to improve the reading skills of first graders who are at risk for reading failure. Five models were identified: Reading Recovery, Success for All, Prevention of Learning Disabilities, The Wallach Tutoring Program, and Programmed Tutorial Reading. Sixteen studies evaluating these models found substantial positive effects of tutoring in comparison to traditional methods. Follow-up studies found that effects of

tutoring were generally lasting. Results were more positive when reading instruction was based on a more comprehensive model of reading and when certified teachers (rather than paraprofessionals) were the tutors. The cost effectiveness of tutoring and the meaning of the findings for remedial and special education are discussed.

References: Outdoor adventure learning

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**

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 (ĆES = .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.

References: Parental involvement

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.

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.

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
Van-Steensel et.al., 2011 (family literacy)	0.18
Indicative effect size	**0.26**

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 behaviors. 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 behavior, family functioning, parental mental health and health risk behaviors, 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).	No abstract provided!
	This review focuses on intervention studies that tested whether parent–child reading activities would

Senechal & Young, (2008). 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.

Van-Steensel et.al. (2011). This meta-analysis examines the effects of family literacy programs on children's literacy development. It analyzes 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.

References: Peer tutoring

Full references

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: pp. 237-248.

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.

Jun, S.W., Ramirez, G., & Cumming, A. (2010). Tutoruing Adolescents in Literacy: A Meta-Analysis. Journal of Education, 45 (2), 219-238.

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.

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.

Topping, K.J. (2005). Trends in Peer Learning Educational Psychology 25.6.pp 631-645.

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.

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.

Summary of effects

Study	*Effect size*
Cohen, Kulik and Kulik, 1982 (on tutees)	0.40
Cohen, Kulik and Kulik, 1982 (on tutors)	0.33
Ginsburg-Block et.al., 2006	0.48
Jun, Ramirez & Cumming, 2010 (cross age peer tutoring)	1.05
Ritter et.al., 2009	0.30
Rohrbeck et.al., 2003	0.59
Indicative effect size	**0.48**

Meta-analyses abstracts

Study	*Abstract*
Cohen et.al.	A meta-analysis of findings from 65 independent evaluations of school 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-

(1982). 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.

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.

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.

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.

References: Performance pay

Summary of effects

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

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.

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.

References: Phonics

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.

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.

Connor, C.M., Morrison, F.J., Fishman, B.J., Schatschneider, C. & Underwood, P. (2007). Algorithm-guided individualized reading instruction. Science, 315, 464–465.

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.

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.

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.

Stuebing, K.K., Barth, A.E., Cirino, P.T., Francis, D.J. & Fletcher, J.M. (2008). A response to recent reanalyses of the National Reading Panel report: Effects of systematic phonics instruction are practically significant. Journal of Educational Psychology, 100.1, 123-134.

Swanson, H.L. (1999). Reading Research for Students with LD: A Meta-Analysis of Intervention Outcomes Journal of Learning Disabilities 32, 504-532.

Torgerson C.J., Brooks, G. & Hall, J. (2008). A Systematic Review of the Research Literature on the use of Phonics in the Teaching of Reading and Spelling. Research Report, No.711 London: DfEE.

Summary of effects

Study	*Effect size*
Camilli, Vargas & Yurecko, 2003	0.24
Ehri, Nunes, Stahl & Willows, 2001	0.41
Jeynes, 2008	0.30
Torgeson, Brooks & Hall, 2006	0.27
Slavin et al. 2011 (One-to-one phonics tutoring)	0.62
Slavin et al. 2011 (Small group phonics)	0.35
Indicative effect size	**0.35**

Meta-analyses abstracts

Study	*Abstract*
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.

Torgerson et.al. (2008). No abstract provided.

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.

References: Physical environment

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://www.cfbt.com/PDF/91078.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.

Xie, H., Kang, J., & Tompsett, R. (2011). The impacts of environmental noise on the academic achievements of secondary school students in Greater London. Applied Acoustics, 72(8), 551-555.

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**

Meta-analyses abstracts

Study	*Abstract*
See above.	

References: Reducing class size

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 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 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**

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.

References: Repeating a year

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 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).

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**

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, behavior, 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 favor of a more productive course of action in the new millennium.

Yoshiba (1989).

Current interest in the grade-standards promotion policy and grade retention, resulting from the minimum competency testing movement, emphasizes the need for practical research information on the differential effects of promotion policies in a way that can assist policy makers. Meta-analysis was employed to conduct an integrative review of the research literature, and to analyze relationships between substantive and methodological variables of the sample studies and study results. The substantive variables included sex, ethnicity, SES, and grade level of the pupils retained; the academic interventions used during the year of retention; and the measures or criteria used to determine academic progress subsequent to retention. The methodological variables were the quality of the study; the era of publication, that is, the promotion policy in vogue when the study was published; and the time elapsed between retention and the measurement of the effects.^ Thirty-four studies were drawn from dissertation abstracts, journal articles, ERIC documents, narrative reviews, and education references. They met these criteria: (a) they investigated the effects of grade retention in the elementary grades on subsequent student achievement, (b) they included two groups of students, retained students and promoted students, (c) they took place in the United States, and (d) they reported data from which an effect size, that is, a normally distributed statistic defined as the mean of the treatment group minus the mean of the control group divided by the standard deviation of the control group, could be computed.^ The results of the meta-analysis indicated that grade retention has a negative impact overall on subsequent academic achievement across studies with different designs and methodologies. Most studies did not define what occurred during the year of retention. However, the value of an individualized educational program (IEP) for each retained student needs further evaluation because retained students appeared to do better in arithmetic with an IEP in two studies which specified their use during the year of retention. No significant differences were found for grade level of retention. The sampled studies contained insufficient data for comparisons based upon sex, ethnicity, and SES.^ Suggestions for future

research include: (a) random assignment of low achieving student to a retention or promotion condition, (b) operational definition of the retention treatment, and (c) development of a policy evaluation program that includes variables such as the promotion policy as written, the policy as implemented in specific sites, and the effects of the implemented policy on groups of students and individual students with different characteristics.

References: School uniform

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**

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 center. 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' behavior; 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.

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

Sowell, (2012). with disabilities status, School A (with uniforms) had significantly better attendance and somewhat fewer minor behavior infractions, but trended lower in standardized math scores and more intermediate and major behavioral 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.

References: Small group tuition

Full references

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.

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.

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.

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.

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.

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.

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.

Summary of effects

Study	*Effect size*
Elbaum et al. 2000	0.40 (pairs)
	1.61 (small group – NB only one study)
	0.16 (individual) (CI 0.12 to 0.20)
Lou et al. 2001 (with ICT)	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
Indicative effect size	**0.34**

Meta-analyses abstracts

Study	*Abstract*
Elbaum et.al. (2000).	No abstract available.

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

Lou et.al. (2001). 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.

References: Social and emotional aspects of learning

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.

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.

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.

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.

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).

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.

Summary of effects

Study	*Effect size*
Durlak et al. 2011	0.27
Multon et al. 1991	0.26
Payton et al. 2008	0.28
Indicative effect size	**0.27**

Meta-analyses 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, behavior, 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.
	This article reports on meta-analyses of the relations of self-efficacy beliefs to academic

Multon et al. (1991). performance and persistence. Results revealed positive and statistically significant relationships between self-efficacy beliefs and academic performance and persistence outcomes across a wide variety of subjects, experimental designs, and assessment methods. The relationships were found to be heterogeneous across studies, and the variance in reported effect sizes was partially explained by certain study characteristics. Implications for further research and for intervention are discussed.

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 across the three reviews, indicating that they offer students a practical educational benefit. Given these positive findings, we recommend that federal, state, and local policies and practices encourage the broad implementation of well-designed, evidence-based SEL programs during and after school.

References: Sports participation

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**

Meta-analyses abstracts

Study	*Abstract*
Newman et.al. (2010).	No abstract provided!

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

Lewis, (2004).	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.

References: Summer schools

Full references

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.

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.

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.

Summary of effects

Study	*Effect size*
Lauer et.al., 2006.	0.16
Cooper et al., 2000.	0.26
Lewis, 2010.	0.10
Indicative effect size	**0.19**

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.

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.

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.

References: Teaching assistants

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.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.

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.

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.

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.

Summary of effects

Study	*Effect size*
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
Indicative effect size	**0.00**

Study abstracts

Study *Abstract*

Blatchford et.al. (2009). Not available.

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,

Gerber et.al. (2001). 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.

The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit: Technical Appendices

Professor Steve Higgins

Maria Katsipataki

(School of Education, Durham University)

Dr Dimitra Kokotsaki

Professor Robert Coe

(CEM Centre, Durham University)

Dr Lee Elliot Major (The Sutton Trust) and Robbie Coleman (Education Endowment Foundation)

Summer 2013

Contents

Acknowledgements

The first version of 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 is written by Professor Steve Higgins, Maria Katsipataki (School of Education, Durham University), Dr Dimitra Kokotsaki, Professor Rob Coe (CEM Centre, Durham University), Dr Lee Elliot Major (The Sutton Trust). and Robbie Coleman (Education Endowment Foundation). The authors also thank Isabella McDonald, Laura Evans and Sarah Whiteway for their help producing the January 2013 update.

Suggested reference: Higgins, S., Katsipataki, M., Kokotsaki, D., Coe, R., Major, L.E. & Coleman, R. (2013). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit: Technical Appendix.* London: Education Endowment Foundation.

Overview

The aim of these appendices is to set out some of the methods and assumptions used in the comparative synthesis of effect sizes in the Sutton Trust-EEF Teaching and Learning Toolkit. The primary aim of the Toolkit is to provide schools with evidence from education research which will help them to make informed decisions about spending to support the learning of disadvantaged pupils. Our emphasis is on identifying comparative messages from existing research. In summarising each particular field a number of judgements have had to be made about the applicability of the research evidence to the challenge of supporting learners from disadvantaged backgrounds in schools. This appendix sets out the rationale and sources of evidence for these decisions.

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 possible strategies, and in implementing a strategy as effectively as possible. However there are, of course, some limitations and caveats to the meta-analytic approach we have taken and these are also set out here. The quality of the evidence within any area is variable and one of the issues in meta-analysis is that some of the subtleties of these issues are lost in aggregation. There is also considerable variation in each of the themes that have been summarised for the Toolkit. There are examples within each area where interventions have been successful in improving attainment and others that have been unsuccessful. The most successful approaches on average have had their failures and the least successful their triumphs. What we are saying is that the existing evidence so far suggests provides information and insight which we believe is useful to schools as they make decisions about their spending and teaching priorities. What we are not saying is that approaches which are unsuccessful on average can *never* work or that approaches like feedback and metacognitive approaches will *always* work in a new context, with different pupils, a different curriculum and undertaken by different teachers.

Overall we think that the messages in the Toolkit are encouraging for teachers. The evidence summarised in the Toolkit shows that they can make a difference and that they are the most important people in the education system who are able make that difference to children and young people’s learning. However, we think that the evidence indicates that that the challenge is to get the *pupils* to work harder, not the teachers. Learners need to engage in activities which make them think harder, more deeply and more frequently. They also need to learn what is expected in different subjects and to develop strategies to help them when they get stuck. Above all they should believe they should succeed through effort and that they should be able to seek and respond to feedback to improve their learning.

We should also make it clear that we do not believe that there are any guarantees from the evidence. Teachers and schools will need to try out these ideas and evaluate their usefulness in improving learning. Sometimes this needs perseverance or effort to create the conditions in which learners can respond to feedback or take more responsibility for their learning. Another way of looking at these approaches is seeing them as means to set up a context in which learning is more or less likely to improve. The actual improvement will depend on the extent to which learners actually think harder, more deeply or more frequently about what is being learned and their teachers can support, challenge, extend and develop this thinking.

Section 1: Resources and pupil learning

It is difficult to establish a clear link between educational expenditure and pupils' learning in schools. Analysis of spending per pupil and scores on the Third International Maths and Science Study (TIMSS) found 'no association between spending levels and average academic achievement' even after controlling for variables such as family background and school characteristics' (Hanushek & Woessman, 2010). However, most of the studies have been undertaken at the system level (e.g. whole countries, states or local authorities) where the relationship between allocation of resources and differences in schools, teachers and pupils is highly complex. It may seem obvious that more money offers the possibilities for a better or higher quality educational experience, but the evidence suggests that it is not simply a question of spending more to get better results. This may be because in the UK and other developed countries we broadly spend reasonably efficiently, and increased effectiveness comes at much greater cost (Steele et al., 2007). Much of the early research in this area failed to find a convincing connection for a range of reasons (Burtless, 1996), though meta-analyses of such studies indicated there was a sufficient connection to warrant increased spending (e.g. Greenwald et al. 1998). More recent research suggests that there is a link between spending and outcomes, but that it is a complex picture (e.g. Vignoles et al., 2000) and that higher quality data sets are required to understand the mechanisms by which spending and learning are associated (Levačić & Vignoles, 2002). Some analyses suggest that the effects of greater spending tend to influence mathematics and science more than English in UK secondary schools (Steele et al., 2007) and that disadvantaged pupils may benefit more (Holmund et al. 2008; Pugh et al. 2011).

Over the period 1997-2011 per capita spending in England increased by 85% in real terms (based on projections in DCSF, 2009). During the same period improvements pupil outcomes were marginal on most international and comparative measures (e.g. Tymms, 2004; Tymms and Merrell, 2007; NFER, 2011; OECD, 2011).

Investing for better learning, or spending so as to improve learning, is therefore not easy, particularly when the specific aim is to support disadvantaged learners whose educational trajectories are harder to influence. Much depends on the context, the school, the teachers (their levels of knowledge and experience), the learners (their level of attainment and their social background) and the educational outcomes that you want to improve (knowledge, skills or dispositions). Improving test scores in arithmetic in the short term, for example, may not raise students' aspirations for what further learning in mathematics may accomplish for them. There is some evidence where interventions have been costed that spending can be used effectively to bring about measurable improvement. However these estimates vary considerably. Wiliam (2002), for example, estimated the cost of a formative assessment project with an effect size of 0.32 on pupil attainment was about £2,000 *per teacher* per year. A recent evaluation of Every Child a Reader (Tanner et al., 2011) estimated costs of £3,100 in the first year and £2,600 per year subsequently *per child* with an average reading gain of 13% (non-significant, p142) (estimated at an effect size of about 0.14: Glass, McGaw & Smith, 1981: 136).

We interpret the lack of a clear causal link between general additional spending and learning to mean that it is difficult to spend additional resource effectively to improve learning and to increase attainment, but that there must be some areas which offer better prospects than others. This is what this *Toolkit* seeks to provide. We also think that the evidence shows that if schools want to use any additional resource, such as the Pupil Premium, to benefit disadvantaged learners they should not assume that any increased allocation alone will improve learning, but they will need to decide specifically and deliberately how it should be spent, and then evaluate the impact of this for themselves. The existing research indicates that this is a challenging but achievable task.

Section 2: Cost effectiveness estimates

Cost estimates are based on the likely costs of adopting or implementing an approach with a class of twenty-five pupils. Where an approach does not require an additional resource, estimates are based on the cost of training or professional development which may be required to support establishing new practices. Approaches marked with £££ or less could be funded from the 2012-13 pupil premium allocation of £623 per eligible pupil. For example, at least 40 pupils receiving the Pupil Premium will be needed to employ an additional teacher in 2012-13 (assuming Main Pay Scale 3 (£25,168) or Outer London MPS1 (£25,117). This drops to 28 pupils eligible for the £900 Pupil Premium in 2013-4. If the Pupil Premium eventually increases to £1,200, this will be reduced to about 20 pupils.

In terms of cost effectiveness it may also be useful for schools to consider the kind of investment they are making. Reducing class sizes only last for as long as the funding maintains smaller classes. Technology equipment typically lasts for up to five years or so (with some maintenance costs). Developing teachers' feedback skills through professional development is potentially more valuable, as it may make a more lasting change in their effectiveness and build capacity within the school.

The scale used in the costing assumptions is as follows:

£	*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-£5,000 per year per class of 25 pupils, or up to about £170 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. This represents the 2012/13 Pupil Premium allocation (£623).
££££	*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. By 2014/5, the Pupil Premium is projected to rise to approximately £1,200 per pupil.

Other estimates, based on costs per class or per teacher are as follows:

Expenditure	Rate	Cost estimate
Teacher	£25-£30k per year (Scale point 3 England & Wales, Inner London Scale Point 3)	£27,500 per year
Teaching Assistant	£16-20k per year	£18,000 per year
Supply cover	£150-£200 per day	£175 per day
Computer	Total cost of ownership estimated at £3,000	£600 per year
CPD day course	£60-£500 per day	£200 per day
CPD programme	Training, support and cover for a 5 day programme with classroom development	£2,000 per year

Section 3: Effect size: what it is, what it means and how it is calculated

What is an effect size?

An effect size[1] is a key measure in intervention research and an important concept in the methodology of the *Toolkit*. It is basically a way of measuring the *extent* of the difference between two groups. It is easy to calculate, readily understood and can be applied to any measured outcome for groups in education or in research more broadly.

The value of using an effect size is that it quantifies the effectiveness of a particular intervention, relative to a comparison group. It allows us to move beyond the simplistic, 'Did it work (or not)?' to the far more important, 'How *well* did it work across a *range* of contexts?' It therefore supports a more scientific and rigorous approach to the accumulation of knowledge, by placing the emphasis on the most important aspect of the intervention – the size of the effect – rather than its statistical significance, which conflates the effect size and sample size. For these reasons, effect size is the most important tool in reporting and interpreting effectiveness, particularly when drawing comparisons about *relative* effectiveness of different approaches.

The basic idea is to compare groups, relative to the distribution of scores. This is the standardised mean difference between two groups. There has been some debate over the years about exactly how to calculate the effect size (see below), however in practice most of the differences in approaches are small in the majority of contexts where effect sizes are calculated using data on pupils' learning. It is important to remember that, like with many other statistics, the effect size is based on the average difference between two groups. It does not mean that all of the pupils will show the same difference.

For those concerned with statistical significance, it is still readily apparent in the confidence intervals surrounding an effect size. If the confidence interval includes zero, then the effect size would be considered not to have reached conventional statistical significance. The advantage of reporting effect size with a confidence interval is that it lets you judge the size of the effect first and then decide the meaning of conventional statistical significance. So a small study with an effect size of 0.8, but with a confidence interval which includes zero, might be more interesting educationally that a larger study with a negligible effect of 0.01, but which is statistically significant.

What does it mean?

As an example, suppose we have two classes of 25 students, one class is taught using a feedback intervention, the other is taught as normal. The classes are equivalent before the intervention. The intervention is effective with an effect size of 0.8. This means that the average person in the class receiving the feedback intervention (i.e. the one who would have been ranked 12th or 13th in their class) would now score about the same as the person ranked 6th in a control class which had not received the intervention. Visualising these two individuals provides a valuable interpretation of the difference between the two effects (see Figure 1).

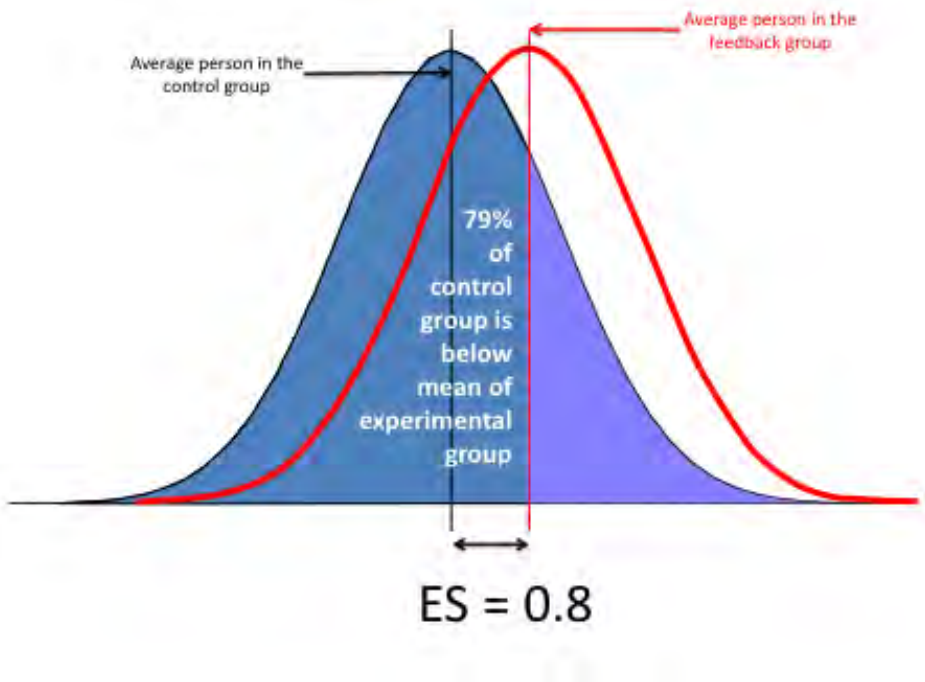
Another way to interpret effect sizes is to compare them with effect sizes of differences that are familiar. For example, Cohen (1969, p23) describes an effect size of 0.2 as 'small', and gives to illustrate the point an example that the difference between the heights of 15 year old and 16 year old girls in the US corresponds to an effect of this size. An effect size of 0.5 is described as 'medium' and is 'large enough to be visible to the naked eye'. A 0.5 effect size corresponds to the difference between the heights of 14 year old and 18 year

[1] Effect sizes can be thought of in two broad categories: first, those which compare the extent of the differences between two groups or standardised mean differences, such as Cohen's d or Hedges g; and second, variance-accounted for effect sizes, such as η^2, ω^2 or R^2) which report the extent to which overlap of key variables is explained. A third group of corrected effect sizes (Vacha-Haase and Thompson, 2002), are variations of these two, but which attempt to adjust for sampling issues. Some effect sizes can be converted mathematically into others (d to r, for example). However, it is important to bear in mind the research design from which data is analysed and the precise calculation method used in understanding the comparability of particular effect size measures (Coe, 2004). The Toolkit focuses on standardised mean difference as a measure of the impact of different interventions.

old girls. Cohen describes an effect size of 0.8 as 'grossly perceptible and therefore large' and equates it to the difference between the heights of 13 year old and 18 year old girls.

As a further example he states that the difference in IQ between holders of the PhD and 'typical college freshmen' is comparable to an effect size of 0.8.

FIGURE 1: AN EFFECT SIZE OF 0.8



Although this labelling also corresponds with the overall distribution of effects found in education research with an average around 0.4 (Sipe and Curlette, 1997; Hattie and Timperley, 2007), a ‘small’ effect may be educationally important if, for example, it is easy or cheap to attain or is achievable with groups who are otherwise hard to influence. Similarly a large effect size may not be as important if is unrealistic to bring about in normal circumstances. Cohen does acknowledge the danger of using terms like 'small', 'medium' and 'large' out of context. Glass and colleagues (1981, p104) are particularly critical of this approach, arguing that the effectiveness of a particular intervention can only be interpreted in relation to other interventions that seek to produce the same effect. They also point out that the practical importance of an effect depends entirely on its relative costs and benefits. In education, if it could be shown that making a small and inexpensive change would raise academic achievement by an effect size of even as little as 0.1, then this could be a very significant improvement, particularly if the improvement applied uniformly to all students, and even more so if the effect were cumulative over time.

As a standardised metric an effect size can also be converted to other measures for comparison: e.g. “students at Phoenix Park outperformed those at Amber Hill in the national school-leaving examination (the General Certificate of Secondary Education, or GCSE) by, on average, one third of a grade, equivalent to a standardized effect size of 0.21” (Wiliam et al. 2004: 50). So using this conversion, an effect size of 0.8 would be equivalent to an improvement of just over one GCSE grade.

In the Toolkit we have equated school progress in months to effect size as a crude but meaningful equivalent (see Table 1, below). We have assumed that a year of progress is about equivalent to one standard deviation per year and corresponds with Glass’ observation that “the standard deviation of most achievement tests in elementary school is 1.0 grade equivalent units; hence the effect size of one year’s instruction at the

elementary school level is about +1" (Glass, 1981: 103). However, in is important to note that the correspondence of one standard deviation to one year's progress can vary considerably for different ages and types of test.

It is also the case that effect size difference reduces with age. Bloom and colleagues (2008) estimate annual progress on tests drops from 1.52 to 0.06 for reading and from 1.14 to 0.01 for mathematics in the US from Kindergarten to Grade 12. Wiliam (2010) estimates "apart from the earliest and latest grades, the typical annual increase in achievement is between 0.3 and 0.4 standard deviations". In the UK, data[2] from National Curriculum tests (DfES, 2004) indicates annual gains representing an effect size of about 0.8 at age 7 (at the end of Key Stage 1), falling to 0.7 at 11 (at the end of Key Stage 2) and only 0.4 at age 14 (end of Key Stage 3). One implication of this is that our estimates of improvement may underestimate the gains achievable for older pupils. If 11 year old pupils tend to make 0.7 standard deviations progress over a year, then the potential gain in terms of months estimated from meta-analytic effect sizes would increase by nearly a third. However, we think this would overestimate the gains achievable for younger children, particularly when effect sizes are re-estimated as months of possible additional progress. On the other hand, part of the reason that the same effect corresponds to more 'months gain' in older pupils is that their overall rate of gain slows down. By the end of secondary school age, the difference between the attainments of successive age groups is relatively small, especially compared with the spread within each. For these older pupils it may be a bit misleading to convert an effect size into typical months' gain: one month's gain is typically such a small amount that even quite a modest effect appears to equate to what would be gained in a long period of teaching.

TABLE 1: CONVERTING EFFECT SIZE TO MONTHS' PROGRESS

Months' progress	Effect Size from …	... to	Description
0	-0.01	0.01	Very low or no effect
1	0.02	0.09	Low
2	0.10	0.18	Low
3	0.19	0.26	Moderate
4	0.27	0.35	Moderate
5	0.36	0.44	Moderate
6	0.45	0.52	High
7	0.53	0.61	High
8	0.62	0.69	High
9	0.70	0.78	Very high
10	0.79	0.87	Very high
11	0.88	0.95	Very high
12	0.96	>1.0	Very high

There are other reasons for preferring a more conservative estimate of what it likely to be achievable in practice. One problem is that estimates of the effects of interventions come from research studies that may

[2] http://www.education.gov.uk/rsgateway/DB/SBU/b000481/b02-2004v2.pdf, with thanks in particular to Michelle Weatherburn and Helen Evans at the Department for Education for identifying this data and providing support with the interpretation of National Test data.

optimise rather than typify their effects. For example, research is often conducted by advocates of a particular approach; considerable care is often taken to ensure that the intervention is implemented faithfully in the research setting; outcome measures used in research studies may be better aligned with the aims and focus of the intervention than other more general measures. For these reasons it may be unrealistic to expect schools to achieve the gains reported in research whose impact may be inflated (this is what Cronbach and colleagues (1980) called 'super-realisation bias'). Other evidence suggests that effect sizes will also be smaller as interventions are scaled up or rolled out (Slavin & Smith, 2008). A further problem is that part of the learning gain typically achieved in a year of schooling may be a result of maturational gains that are entirely independent of any learning experiences that are, or could be, provided by the school. For example, Luyten (e.g. 2006; Luyten et al., 2006) has shown that a substantial part (sometimes more than half) of the difference between the attainments of pupils in successive school grades is accounted for by differences in the ages of pupils who have experienced exactly the same schooling. The implication seems to be (though this is somewhat speculative) that any potential accelerating effect of using the kinds of strategies we have discussed in this report may be limited to changing the part of the year's gain that is due to schooling, while the growth that is due to pure maturation may be harder to affect. For these reasons we have selected what we see as a more conservative estimate, based on effect size estimates for younger learners, which can be improved or refined as more data becomes available about effect size transfer from research studies to practice.

Methods of calculation

Over the years there have been a number of methods proposed to calculate the appropriate standard deviation for an effect size. The main approaches are listed below.

Glass's Δ

Gene V. Glass (1977) proposed an estimator of the effect size that uses only the standard deviation of the control group, this is commonly referred to as Glass's Δ (delta). He argued that if several interventions or treatments were compared with the control group it would be better to use just the standard deviation from the control group, so that effect sizes would not differ under equal means and different variances.

Cohen's d

Cohen's *d* is defined as the difference between two means divided by an unspecified standard deviation for the data. This definition of Cohen's d is termed the 'maximum likelihood estimator' by Hedges and Olkin (1985).

Hedges's g

Hedges's *g*, suggested by Larry Hedges (1981) is based on a standardized mean difference, like the other measures, but the pooled standard deviation is computed slightly differently from Cohen's *d*.

d or g (corrected)?

Hedges's *g* is biased for small sample sizes. However, this bias can be adjusted (*g* (corrected)). Hedges and Olkin (1985) refer to this unbiased estimate as *d*, but it is not the same as Cohen's *d*. In most recent meta-analyses when an effect size is referred to as Hedges's *g* it is the bias-corrected formula which has been used, though some studies also refer to this as *d*.

Final issues

There are some notes of caution in comparing effect sizes across different kinds of interventions. Effect size as a measure assumes a normal distribution of scores. If this is not the case then an effect size might provide a misleading comparison. If the standard deviation of a sample is decreased (for example, if the sample does not contain the full range of a population) or inflated (for example, if an unreliable test is used), the effect size is affected. A smaller standard deviation will increase the effect size, a larger standard deviation will reduce it. Another key issue is which standard deviation is chosen (Bloom et al., 2008) as this primarily determines the comparability of the effect size (Coe, 2004). This explains the variation in methods advocated above.

There is also evidence that there is some systematic variation in effect sizes in education. One factor, for example, is the age of the pupils, where studies with younger learners tend to have higher effect sizes. One reason for this is likely to be the narrower distribution of scores producing a smaller standard deviation and therefore a larger effect size, though there is also a relationship with the subject (e.g. mathematics or English) being researched (Hill, Bloom & Lipsey, 2009). In England the standard deviations of National Test scores[1] increase from 3.9 at age 7, to 4.3 at age 11, and 6.8 at 14 as the distribution of scores widens and flattens (DfES, 2004).

There is also some variation associated with the type of outcome measure with larger effect sizes typically reported in mathematics and science compared with English (e.g. Higgins et al., 2005) and for researcher designed tests and teacher assessments compared with standardised tests and examinations (e.g. Hill et al., 2007: 7).

Slavin and Smith (2009) also report that there is a relationship between sample size and effect size in education research, with smaller studies tending to have larger effect sizes. The correlation found was -0.28 (p503), suggesting that is explains about 8% of the variation between large and small studies. The issue is important in terms of comparing effects between different kinds of interventions which tend to be small scale (such as areas of research looking at interventions to address special needs for example) and others which tend to have larger samples (class size interventions for example).

Other systematic factors may also affect such comparisons. Studies reporting effect sizes with groups from either end of the distribution (high attaining or low attaining learners) are likely to be affected by regression to the mean if they don't compare like with like (Shagen & Hogden, 2009). This would inflate effect sizes for low attaining pupils (who are more likely to get higher marks on re-test) and depress effect sizes for high performing students when they are compared with 'average' pupils. If the correlation between pre-test and post-test is 0.8, regression to the mean may account for as much as 20% of the variation in the difference between test and retest scores when comparing low and average students.

The aim of the *Toolkit* is not to provide definitive claims as to what *will* work to bring about improvement in a new context. Rather it is an attempt to provide the best possible estimate of what is likely to be beneficial based on existing evidence. In effect it summarises what *has worked* as a 'best bet' for what might work in the future. The applicability of this information to a new context is always likely to need active enquiry and evaluation to ensure it helps to achieve the desired effects.

Section 4: Meta-analysis and 'super-synthesis' of intervention research in education

Meta-analysis is a method of combining the findings of similar studies to provide a combined quantitative synthesis or overall 'pooled estimate of effect'. The results of, say, interventions seeking to improve low attaining students' learning in mathematics can be combined so as to identify clearer conclusions about which interventions work and what factors are associated with more effective approaches. The advantages of meta-analysis over other approaches to reviewing are that it combines or 'pools' estimates from a range of studies and should therefore produce more widely applicable or more generalisable results.

In addition, it can show whether the findings from similar studies vary more that would be predicted from their samples so that the causes of this variation can be investigated (moderator analysis). In education research this is particularly valuable as the results from small studies can be combined to provide answers to questions without being so dependent on the statistical significance of each of the individual studies which relates closely to sample size. Many small studies with moderate or low effects may not reach statistical significance and if you review the field by simply counting how may were statistically significant, you may be misled into thinking that the evidence is less conclusive than if you combine these studies into one combined study or meta-analysis. The statistical techniques to undertake meta-analysis form a set of transparent and replicable rules which are open to scrutiny (Aguinis et al. 2010).

Another key advantage of meta-analysis is that it helps to deal with the quantity of information in education research which can overwhelm other approaches (Chan and Arvey, 2012). This is particularly important when trying to draw relative inferences across different areas of education research. The number of studies available to review in any area of education is extensive, so techniques to aggregate and build up knowledge to propose further research and test theories and ideas are invaluable. In fields like psychology and medicine meta-analysis is relatively uncontroversial as a synthesis technique with nearly 40 years development of the principles and methods involved.

'Super-synthesis'

It is also tempting to look at results across different kinds of studies with a common population, so to provide more general or comparative inferences. This approach is, of course, vulnerable to the classic 'apples and oranges' criticism which argues that you can't really make a sensible comparison between different kinds of things. However as Gene Glass (2000) said, "Of course it mixes apples and oranges; in the study of fruit nothing else is sensible; comparing apples and oranges is the only endeavor worthy of true scientists; comparing apples to apples is trivial."

A number of publications have attempted to take meta-analysis this stage further, by synthesising the results from a number of existing meta-analyses – producing what has been called a 'meta-meta-analysis' (Kazrin, Durac & Agteros, 1979), a 'mega-analysis' (Smith 1982), 'super-analysis' (Dillon, 1982) or 'super-synthesis' (e.g. Sipe & Curlette, 1997). However, one can make a clear separation of types within these studies. Some use the meta-analyses as the unit of analysis in order to say something about the process of conducting a meta-analysis and identifying statistical commonalities which may be of importance (e.g. Ioannidis & Trikalinos, 2007; Lipsey and Wilson, 1993). Others, however, attempt to combine different meta-analyses into a single message about a more general topic than each individual meta-analysis can achieve (e.g. Bloom, 1984; Walberg, 1984; Hattie, 1992; Sipe & Curlette, 1997). Even here, there appears to be a qualitative difference – some retain a clear focus, either by using meta-analyses as the source for identifying original studies with an overarching theoretical focus (e.g. Marzano, 1998) in effect producing something might best be considered as a series of larger meta-analyses rather than a meta-meta-analysis. Others, though, make claims about broad and quite distinct educational areas by directly combining results from identified meta-analyses (e.g. Hattie, 1992; Sipe & Curlette, 1997). In terms of the apples and oranges analogy, this is a little like asking which fruit is best for you, as a lot depends on what you mean by 'best' and how this is measured.

Hattie (2009) synthesized more than 800 meta-analyses and came up with some interesting findings. First of all, he concluded that most things in education 'work' as the average effect size is about 0.4. He then uses this to provide a benchmark for what works above this 'hinge' point. There are, of course, some reservations about this 'hinge' as small effects may be valuable if they are either cheap or easy to obtain, or tackle an otherwise intractable problem. Similarly, large effect sizes may be less important if they are unrealistic and if they cannot be replicated easily in classrooms by teachers. Further reservations about combining effect sizes of different kinds suggest that intervention effects should be distinguished from maturational differences or correlational effects sizes. The underlying distributions may be of different kinds, so that unlike comparing fruit, it is more like comparing an apple with a chair (Higgins & Simpson, 2011).

Although there are clearly limitations to the application of quantitative synthesis in this way, the data from meta-analysis offers the best source of information to try to answer comparative questions between different areas of educational research. It is hard to compare areas without some kind of benchmark. If you have two narrative reviews, one arguing that, say, parental involvement works and another arguing that digital technology is effective, and both cite studies with statistically significant findings showing they each improve reading comprehension, it is hard to choose between them in terms of which is likely to offer the most benefit. Meta-analysis certainly helps to identify which researched approaches have made, on average, the most difference, in terms of effect size, on tested attainment of pupils in reading comprehension or other areas of attainment. We suggest that this comparative information should be treated cautiously, but taken seriously. If effect sizes from a series of meta-analysis in one area, such as meta-cognitive interventions for example, all tend to be between 0.6 and 0.8, and all of those in another area, such as individualised instruction, are all between -0.1 and 0.2, then this is persuasive evidence that schools should investigate meta-cognitive approaches to improve learning, rather than focus on individualised instruction. Some underlying assumptions are that the research approaches are sufficiently similar (in terms of design for example), that they compared sufficiently similar samples or populations (of school pupils) with sufficiently similar kinds of interventions (undertaken in schools) and similar outcome measures (standardised tests and curriculum assessments). So, if you think that a meta-analysis of intervention research into improving reading comprehension has a set of broadly similar set of studies, on average, to a meta-analysis investigating the development of understanding in science, then you might be tempted to see if any approaches work well in both fields (such as reciprocal teaching) or, indeed, don't work well in both fields (such as individualised instruction). Our argument is that so long as you are aware of the limits of the inferences drawn, then the approach has value. We suggest that this provides the best evidence we have so far, particularly where we have no studies providing direct comparisons.

Toolkit themes

The initial themes for the Toolkit were based on expectations of how schools seemed likely to spend the Pupil Premium when it was first announced. A number of areas were specifically included at the request of teachers who have been consulted at different stages in the development of the Toolkit. Thanks in particular go to ARK and teachers from the TeachFirst Future Leaders programme, a group of Hammersmith and Ealing deputy headteachers and a number of teachers in the North-East of England who were generous with their time in attending conference or workshop presentations about earlier drafts of the Toolkit. Some of these areas (e.g. School Uniforms, Performance Pay) did not have any quantitative systematic reviews or meta-analyses to support a pooled estimate of effect. Inferences drawn from single studies or projects are limited, so these topics have a lower overall quality assessment in terms of the overall warrant from the research evidence. Feedback from schools and teachers forms an important part of the development of the Toolkit.

Search and inclusion criteria

The initial source of studies for the Toolkit was a database of meta-analyses of educational interventions developed for an ESRC Researcher Development Initiative.[3] Additionally repeated systematic searches have been undertaken for systematic reviews with quantitative data (where effect sizes are reported but not

[3] ESRC Grant RES-035-25-0037: 'Training in the Quantitative synthesis of Intervention Research Findings in Education and Social Sciences'.

pooled) and meta-analyses (where effect sizes are combined to provide a pooled estimated of effect) of intervention research in education in each of the areas of the Toolkit. These searches have been applied to a number of information gateways including Web of Knowledge, FirstSearch, JSTOR, ERIC, Google Scholar and ProQuest Dissertations. In addition a number of journals were hand searched (e.g. Review of Educational Research and Education Research Review). Journal publishers' websites offering full-text searching (Elsevier, Sage, Wiley-Blackwell) were also searched for meta-analyses. Relevant references and sources in existing super-syntheses (e.g. Sipe & Curlette, 1997; Marzano, 1998; Hattie, 2009) were identified and obtained where possible. A record of the search strategy used and studies found are kept for each of the Toolkit themes. Other studies found during the search process are also consulted in each area to provide additional contextual information.

Estimating overall impact

In each area of the Toolkit an overall estimate of the effects is identified. Where the data is available a weighted mean is used. This is based on calculating a weight for each meta-analysis according to its variance, based on the reciprocal of the square of the standard error (Borenstein et al. 2010). Where the data is not available for this an estimate is given based on the available evidence and a judgement made about the most applicable estimate to use (such as the impact on disadvantaged pupils, or the most rigorous of the available meta-analyses). Where no meta-analyses of educational interventions in a given area could be found an effect size is estimated from correlational studies or large scale studies investigating the relationship under review. If there is no information in this area, then individual studies are identified which can provide a broad estimate of effect..

Weight of evidence and quality assessment

The weight of evidence in each field was assessed according to the criteria in Table 2 below and a judgement made about how well the descriptors matched each area included in the Toolkit. These criteria are weighted to identify consistency in terms of the findings (both the overall pooled effect the pattern of effects relating to moderator variables) and to give weight to ecological validity (where studies took place in schools with interventions managed by teachers rather than researchers). The focus of the Toolkit is on providing advice to schools about how to spend additional resource to benefit disadvantaged learners, so these were judged to to be important criteria.

TABLE 2: QUALITY ASSESSMENT CRITERIA

Rating	Description
★	*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.
★★	*Limited:* At least one meta-analysis or systematic review with quantitative evidence of impact on attainment or cognitive or curriculum outcome measures.
★★★	*Moderate:* Two or more rigorous meta-analyses of experimental studies of school age students with cognitive or curriculum outcome measures.
★★★★	*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.
★★★★★	*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.

Section 5: Data table of meta-analyses and other studies used to estimate effect sizes

Meta-analysis	Pooled effect	ES	SE	SD	CI lower	CI upper	Min ES	Max ES	No. studies	No. Effects	Number of pupils	Mod. analysis	Pub bias
Ability grouping													
Kulik & Kulik 1982 (secondary - all)	0.10	Δ	0.05	0.32	0.01	0.19	-1.25	1.50	52	36		Yes	
Kulik & Kulik 1984 (elementary - all)	0.07	Δ	0.04						28			Yes	
Lou et al 1996 (on low attainers)	-0.12	g	-0.06		-0.01	-0.24	-1.96	1.52	103	51	16073	Yes	
Slavin 1990 b (secondary low attainers)	-0.06	Δ	-0.03	-0.12					29	15		Yes	
Indicative effect size (weighted mean)	**-0.007**												
Ability Grouping: Gifted and Talented													
Kulik & Kulik, 1987 (within class grouping)	0.62											Yes	
Kulik & Kulik, 1987 (between class grouping)	0.33											Yes	
Kulik & Kulik, 1992 (accelerated classes)	0.87	d							23			No	
Kulik & Kulik, 1992 (enriched classes)	0.41	d							25			No	
Rogers, 2007 (promotion)	1.00								32			No	
Rogers, 2007 (starting school early)	0.49								68			No	
Rogers, 2007 (ability grouping)	0.49								32			No	
Rogers, 2007 (pull-out groups)	0.65								7			No	
Rogers, 2007 (subject acceleration)	0.59								21			No	
Rogers, 2007 (G&T collaborative groups)	0.26								3			No	
Steenbergen-Hu & Moon, 2010 (all studies)	0.18	g	0.128		-0.072	0.431			28	274		Yes	No
Steenbergen-Hu & Moon, 2010 (school pupils with peers)	0.40	g	0.187		0.029	0.762			13			Yes	No
Vaughn, Feldhunsen & Asher, 1991 (pull-out)	0.65		0.19						3			No	
Indicative effect size (median)	**0.49**												
Adventure Education													
Cason & Gillis, 1994 (all)	0.31	d		0.62			-1.48	4.26	43	147	11238	Yes	
Cason & Gillis, 1994 (school grades)	0.61	d		1.527			-1.48	4.26	43	147		Yes	
Gillis & Speelman, 2008 (overall)	0.43	d					-0.24	2.83	44	390	2796	Yes	
Gillis & Speelman, 2008 (academic achievement)	0.26	d					-0.24	2.83	44	390		Yes	
Hattie et al. 1997 (all)	0.34	d	0.09						96	1728	12057	Yes	

Hattie et al. 1997 (academic outcomes)	0.45	d	0.23						96	1728		Yes	
Laidlaw, 2000	0.17			0.39			-0.43	1.38	48	389	3,550	Yes	
Indicative effect size (weighted mean)	**0.23**												
After school programmes													
Crawford, 2011	0.40	d	0.05		0.30	0.50	0.02	1.70		23		Yes	Yes
Durlak & Weissberg 2007	0.16	g	0.08	0.57	0.01	0.14	-0.16	0.67	55	66			No
Fashola 1998	*NPE*	*d*					*0.11*	*0.90*					
Lauer, Akiba & Wilkerson 2006	0.07	g	0.03		0.01	0.11			15	21		Yes	Yes
Scott-Little et al 2002	*NPE*	*d*					*0.38*	*0.50*	*23*				*No*
Tanner et al. 2011	0.14		0.11										
Zief et al. 2006 (GPA)	0.08	d			-0.03	0.20			5			No	No
Zief et al. 2006 (reading)	0.03	d			-0.10	0.16			2			No	No
Indicative effect size (median)	**0.10**												
Arts participation													
Lewis, 2004	0.20	d	0.15	0.09					5				
Newman et al. 2010 (pri/EY cognitive)	0.45	g(c)	0.09	0.19	0.28	0.62	-0.06	1.13	5	10			
Newman et al. 2010 (sec Eng)	0.05	g(c)	0.02	0.04	0.01	0.09	-0.01	0.08	1	3			
Newman et al. 2010 (sec maths)	0.03	g(c)	0.02	0.03	0.00	0.06	-0.01	0.05	1	3			
Newman et al. 2010 (sec sci)	0.06	g(c)	0.01	0.01	0.05	0.07	0.05	0.06	1	3			
Standley 2008	0.32	d	0.05	0.25	0.23	0.41	-0.23	1.70	30			Yes	Yes
Winner & Cooper 2000 (Literacy)	0.02	d	0.01	0.05			-0.25	0.66	24	24	19277	Yes	Yes
Winner & Cooper 2000 (Maths)	0.04	d	0.02	0.14			-0.14	0.34	15	15	18736	Yes	Yes
Indicative effect size (weighted mean)	**0.15**												
Aspiration interventions													
Cummings et al. 2012(parental interventions-aspirations)	0.24-0.66	d							60				
Cummings et al. 2012(mentoring-aspirations)	0.11-0.24	d							60				
Cummings et al. 2012 (extra curricular-aspirations)	0.043-0.155	d							60				
Cummings et al. 2012 (parental aspirations-attainment)	0.17-0.45	d							60				
Cummings et al. 2012 (extra curricular-attainment)	0.032-0.092	d							60				
Cummings et al. 2012 (mentoring-attainment)	0.09-0.22	d							60				
Petscher, 2010 (reading)	0.32	Zr			0.28	0.36			32	118	214615	Yes	No
Gollwitzer & Sheeran, 2006	0.65	d			0.6	0.7			63	94	8461		

Ma & Kishor, 2007 (maths)	0.12	r			0.12	0.13			113	108	82941	Yes	No
Indicative effect size	**0.00**												
Behaviour interventions													
Chitiyo et al. 2012 (students with disabilities)	0.87	r							5		25	No	
Gansle, 2005	-0.11	Q		0.28					9			Yes	
Gonzalez et al. 2004	0.49	Zr			0.43	0.55			19	7		Yes	
Quinn et al. 1999(emotional disorder)	0.05	M	0.14	0.6					17			Yes	No
Reddy et al. 2009 intervention-emotional disturbance	1.78	d							29	18	1405	No	Yes
Reddy et al. 2009 prevention-emotional disturbance	0.28	d							29	118		No	Yes
Sander et al. 2012	0.02	d			-0.18	0.22			15	134		Yes	No
Wilson & Lipsey, 2007	0.22				0.15	0.25			249			Yes	
Indicative effect size	**0.32**												
Block scheduling and timetabling													
Dickson et al. 2010 (achievement)	0.11	g (c)	0.06		-0.01	0.22	-0.14	0.48	12	7		Yes	No
Dickson et al. 2010 (maths)	-0.02	g (c)	0.07		-0.16	0.11	-0.14	0.10	12	6		Yes	No
Dickson et al. 2010 (science)	0.20	g (c)	0.07		0.06	0.33	0.13	0.42	12	4		Yes	No
Lewis et al. 2005 (maths)	-0.10	g	0.01		-0.11	-0.08	-0.15	-0.09	7	5	82463	Yes	82463
Lewis et al. 2005 (English)	-0.17	g	0.01		-0.18	-0.15	-0.25	-0.05	7	3		Yes	
Lewis et al. 2005 (science)	-0.12	g	0.01		-0.13	-0.10	-0.16	0.11	7	2		Yes	
Indicative effect size	**0.00**												
Collaborative Learning													
Igel, 2010	0.44	g			0.22	0.66	-0.08	2.45	20		2412	Yes	No
Johnson et al. 1981	0.78	d		0.99					16		70	Yes	
Johnson et al. 2000 (academic controversy)	0.86	d	0.1									Yes	
Johnson et al. 2000 (cooperative integrated read & composition)	0.18	d	0									Yes	
Johnson et al. 2000 (group investigation)	0.62	d	0.44									Yes	
Johnson et al. 2000 (jigsaw)	0.09	d	0.11									Yes	
Johnson et al. 2000 (learning together)	0.91	d	0.04						164	194		Yes	No
Johnson et al. 2000 (team assisted individualization)	0.19	d	0.04									Yes	
Johnson et al.2000 (student-team achievement)	0.28	d	0.07									Yes	
Romero, 2009	0.4	g			0.255	0.574			30			Yes	No

Indicative effect size (weighted mean)	**0.42**												
Digital technology													
Bayraktar 2001 (science)	0.27	g	0.02	0.11	0.24	0.31	-0.69	1.295	42	108		Yes	
Camnalbur & Erdoğan 2008 (in Turkey)	1.05	d	0.07	0.07	0.91	1.19			78			No	
Cheung & Slavin, 2011 (maths)	0.15	d			0.12	0.21			85		60000	Yes	No
Christmann & Badgett, 2003	0.34	d											
Li & Ma 2010 (maths)	0.28	d	0.08	0.28	0.13	0.43	-0.66	3.76	46	85			No
Liao 2005	0.55	d	0.06	0.73	0.43	0.67	-0.768	1.914	52	134			Yes
Lou, Abrami, d'Apollonia 2001	0.16	gc	0.02	0.20	0.12	0.20	-1.14	3.37	100	178			Yes
Means et al 2009	0.16	g	0.10	0.69	-0.04	0.59	-0.04	0.356	46	51			No
Moran, et al. 2008	0.49	gc	0.11	0.74	0.27	0.71	-0.204	2.679	7	7			Yes
Morphy & Graham, 2012	0.52	d*	0.10		0.33	0.71							
Onuoha 2007	0.26	d	0.05	0.28	0.17	0.35	-0.38	1.12	35	67			No
Oostdam, Otter, Overmaat 2002	0.19	d	0.06	0.40	0.08	0.30	0.13	0.25	42	50			Yes
Pearson, et al. 2005	0.49	g	0.11	0.74	0.27	0.71			30	89		Yes	
Rosen & Salomon, 2007	0.46	d	0.01	0.62	0.44	0.48	-1.152	2.003	32				Yes
Sandy-Hanson 2006 (gen. academic)	0.25	d	0.02	0.47	0.22	0.28	-1.04	1.33	31	31		Yes	
Seo & Bryant 2009	0.37	d	0.01	0.03	0.43	0.67			11				No
Sitzman et al. 2006	0.15	gc	0.02	0.17	0.11	0.19			71				
Strong, Torgerson, Torgerson, Hulme 2011	0.08	d	0.09	0.21	-0.09	0.25	-1	0.17	6	8			No
Tamim et al. 2009	0.35		0.04		0.27	0.43						Yes	
Tamim et al. 2011	0.33	d	0.04	0.20	0.25	0.41	-0.09	0.55	25	574			No
Tokpah, 2008	0.38	d	0.03	0.14	0.34	0.43	-0.47	1.23	31	102			Yes
Torgerson & Elbourne 2002	0.37	g	0.20	0.53	-0.02	0.77	-0.11	1.15	7	6			
Torgerson & Zhu 2003	0.02	g	0.19	0.38	-0.17	0.58				4		Yes	Yes
Torgerson & Zhu 2003	-0.05	g	0.14	0.29	-0.33	0.24				4		Yes	Yes
Torgerson & Zhu 2003	0.89	gc	0.33	0.47	0.25	1.54				2		Yes	Yes
Vogel et al 2006	0.07		0.01	0.06	0.05	0.09			32				
Waxman, Connell, and Gray, 2002	0.30	*Δ*	0.15	0.63	0.00	0.60	-0.154	0.39	20	138			Yes
Waxman, Lin, Michko 2003	0.45	*Δ*	0.14	0.72	0.17	0.72			42	29			
Zhao 2003	0.81	d	0.13	0.72	0.55	1.07	0.28	2.82	9	29			Yes
Indicative effect size (weighted mean)	**0.28**												

Early years intervention													
Anderson et al. 2003	0.35	*Δ*	0.18	0.62			-0.61	0.89	12	29		No	
Camilli et al. 2010	0.23								123				
Gilliam & Zigler 2001	NPE	*Δ*					0.07	0.50	13			No	No
Gorey, 2010 (esti.on long-term impact)	0.55	*U3*							35	80	18000	Yes	Yes
Karoly et al. 2005	0.28												
LaParo & Pianta 2000	0.51		0.26	2.18					70			No	
Lewis & Vosburgh 1988	0.41		0.04	0.39	0.33	0.73	0.21	0.96	65	46	3194	No	
Manning et al. 2010 (adolescent education)	0.53	d			0.40	0.68			23			Yes	Yes
Nelson et al. 2003	0.52	g	0.27	1.55			0.01	1.25	34			Yes	No
Indicative effect size (weighted mean)	**0.45**												
Extended School Time													
Cooper et al. 2003 (district level comparison) with comparison group	0.06	d			-0.02	0.02			13	39	44000	Yes	No
Cooper et al. 2003 (with matched controls)	0.11	d							13			Yes	No
Baker et al. 2004 (international comparison -maths in UK)	*0.12*	*r*											
Indicative effect size	**0.11**												
Feedback													
Bangert-Drowns et.al. 1991	0.26		0.06	0.38			-0.83	1.42	40	58			
Fuchs & Fuchs 1985	0.72	U3	0.09	0.88					21	95		Yes	
Kingston & Nash, 2011 (AfL)	0.20	Q	0.08		0.19	0.21			13	42		Yes	
Kluger & De Nisi, 1996	0.41	d	0.09	1.03	0.23	0.59			131	607	23663	Yes	Yes
Lysakowski & Walberg 1982	0.97	d	0.49	1.53			-1.09	4.99	54	94	14689	Yes	Yes
Tenenbaum & Goldring 1989	0.72		0.37	1.42					15	16			
Walberg 1982	0.81	d	0.41	1.80					19	19			
Indicative effect size (weighted mean)	**0.62**												
Homework													
Cooper, Robinson & Patal 2006	0.60	d	0.26	0.64	0.38	0.82	0.39	0.97	6	9		Yes	
Paschal, Weinsten & Walberg 1984	0.36	*Δ*	0.18	0.24			-0.60	1.96	15	81			
Indicative effect size (Primary)	**0.07**												
Indicative effect size (weighted mean - Secondary)	**0.44**												

Individualised instruction													
Aiello & Wolfe, 1980 (science)	0.35	*d*							115	182		Yes	
Bangert, Kulik & Kulik, 1983	0.10	*Δ*	0.05	0.37	0.00	0.20	-0.84	1.24	49	49		Yes	
Horak, 1981	-0.07	*Δ*	-0.04	-0.28			-1.49	0.53	60	129			
Willett, Yamashita & Anderson, 1983	0.17	*Δ*	0.09	0.41			-0.87	1.74	130	341			
Indicative effect size	**0.10**												
Learning styles													
Garlinger & Frank, 1986	-0.03	d							7	7			
Kavale & Forness, 1987	0.14	d	0.06	0.28	0.02	0.27			39		3087		
Lovelace, 2005 (Dunn & Dunn Model)	0.67	d					0.67	0.80	76	168	7196	Yes	Yes
Slemmer, 2002 (ICT context)	0.13	d	0.03		0.08	0.19			48	51		Yes	Yes
Tamir, 1985	0.02	d							54	13			
Indicative effect size	**0.10**												
Mastery Learning													
Bangert, Kulik & Kulik, 1983	0.10	*Δ*	0.053						49			Yes	No
Guskey & Piggott, 1998 (language)	0.60	gc			0.18	1.02	0.02	1.70	46	78		Yes	
Guskey & Piggott, 1998 (maths)	0.70	gc			0.50	0.90	0.02	1.70	46	78		Yes	
Kulik, Kulik & Bangert-Drowns, 1990	0.52	g	0.033						108			Yes	
Waxman et.al. 1985	0.39	*Δ*					-1.18	1.73	38	309	7200	No	No
Indicative effect size (weighted mean)	**0.40**												
Mentoring													
Bernstein et al. 2009 (maths)	-0.05	d							32		2573	Yes	
Bernstein et al. 2009 (reading)	-0.04	d							32			Yes	
Bernstein et al. 2009 (science)	-0.03	d							32			Yes	
DuBois et al. 2002 (academic)	0.11	d			-0.08	0.08			55	575		Yes	No
Wheeler, Keller & DuBois, 2010 (maths)	-0.02	d			-0.02	0.03			3			No	
Wheeler, Keller & DuBois, 2010 (reading)	-0.01	d			-0.05	0.04			3			No	
Wood & Mayo-Wilson, 2012 (academic performance)	-0.01	g			-0.11	0.08			6		4769	Yes	
Indicative effect size	**0.05**												

Meta-cognition and self-regulation strategies													
Abrami et al. 2008	0.34	gc	0.01	0.61	0.31	0.37	-1.00	2.75	117	161	20698	Yes	Yes
Chiu 1998	0.67	gc	0.34	0.68			-1.25	2.75	43	123	3475	Yes	
Dignath et al. 2008	0.62	d*	0.05		0.52	0.72	0.44	1.00	48	263		Yes	
Haller et al. 1988	0.71	d	0.36	0.81			0.25	3.80	20	115	1553	No	
Higgins et al. 2005	0.62	gc	0.09	0.38	0.45	0.80	-0.17	1.61	19	19		No	Yes
Klauer & Phye 2008	0.69	gc	0.05		0.59	0.79	0.59	0.94	73		3600	Yes	
Indicative effect size (weighted mean)	**0.62**												
One-to-one tutoring													
Cohen, Kulik & Kulik 1982 (on tutees)	0.40	*Δ*	0.07	0.50	0.26	0.54			52			Yes	Yes
Cohen, Kulik & Kulik 1982 (on tutors)	0.33	*Δ*	0.09		0.15	0.51			38			Yes	Yes
Elbaum et al. 2000	0.41	*Δ*	0.05	0.25	0.32	0.49	-1.32	3.34	29		1539	Yes	Yes
Jun, Ramirez & Cumming, 2010 (by adults)	0.70	*d*			0.48	0.93			12			Yes	No
Ritter et al. 2009	0.30	*g*	0.06	0.32	0.18	0.42	0.26	0.45	21		1676	Yes	No
Slavin et al. 2011 (1-1 phonics tutoring)	0.62	*d*							10			Yes	
Tanner et al. 2011	0.14	*d*											
Wasik & Slavin 1993	NPE	*Δ*					0.20	1.16	16				
Indicative effect size (weighted mean)	**0.44**												
Parental involvement													
Bus et al. 1995 (joint book reading)	0.59	d							16	33	3410	Yes	Yes
Jeynes 2005	0.27	gc	0.14	0.57			0.00	1.78	41	17	20000	Yes	
Jeynes 2007	0.25	gc	0.07		0.11	0.39	0.01	0.83	52	20	30000	Yes	
Layzer et al. 2001 (across school age)	0.27	d							562	11.112		Yes	No
Layzer et al. 2001 (preschool)	0.37	d							562	11.112		Yes	No
Nye, Turner & Schwartz, 2006	0.43	g			0.30	0.56			19			Yes	No
Senechal & Young, 2008 (family literacy)	0.65	d			0.53	0.76			16	1.340		Yes	
Van-Steensel et al. 2011 (family literacy)	0.18	d	0.06						30	152		Yes	No
Indicative effect size (weighted mean)	**0.26**												
Peer tutoring/ peer-assisted learning													
Cohen, Kulik & Kulik 1982 (on tutees)	0.40	*Δ*	0.07	0.50	0.26	0.54	-1.00	2.30	52			Yes	Yes
Cohen, Kulik & Kulik 1982 (on tutors)	0.33	*Δ*	0.09						11			Yes	Yes
Ginsburg-Block et al. 2006	0.48	g	0.24	0.39			0.38	0.78	36	36		Yes	

Jun, Ramirez & Cumming, 2010 (cross-age)	1.05	*d*			0.45	1.44			12			Yes	
Ritter et al. 2009	0.30	gc	0.06	0.32	0.18	0.42	0.26	0.45	28				Yes
Rohrbeck et al. 2003	0.59	gc	0.10	0.90	0.40	0.78	0.21	0.62	90			Yes	Yes
Indicative effect size	**0.48**												
Performance pay													
No meta-analyses or systematic reviews. ES estimated from:													
Woessman 2010 (correl)	0.25		0.13										
Martins 2009	-0.09		-0.05										
Indicative effect size	**0.00**												
Phonics													
Camilli, Vargas & Yurecko, 2003	0.24	d							40			Yes	No
Ehri et al. 2001	0.41	d	0.03		0.36	0.47	-0.47	3.71	66	65		Yes	Yes
Jeynes, 2008	0.30	gc	0.10		0.10	0.50	-1.21	2.02	22		5000	Yes	Yes
Slavin et al. 2011 (1-1 phonics tutoring)	0.62	d							10			No	
Slavin et al. 2011 (small groups)	0.35	d							22			No	
Torgerson, Brooks & Hall, 2006	0.27	d	0.09		0.10	0.45	-0.19	2.69	14			Yes	Yes
Indicative effect size (weighted mean)	**0.35**												
Physical Environment													
No meta-analyses or systematic reviews to estimate ES.													
Indicative effect size	**0.00**												
Reducing class sizes													
Goldstein, Yang, Omar, Turner & Thompson, 2000	0.20	d	0.10	0.31			-0.07	0.60	9	36			
Glass & Smith 1978	0.01	*Δ*	0.00	0.04					77	725	9000		No
McGiverin, Gilman & Tillitski 1989	0.34	d	0.13		0.09	0.59	-0.74	2.24	10	24			
Slavin 1990 (a)	0.17	*Δ*	0.09										
Indicative effect size	**0.20**												
Repeating Years													
Allen et al. 2009 (low quality studies)	-0.30	d							22	207		No	No
Allen et al. 2009 (medium & high quality studies)	0.04	d							22	207		No	No
Bright, 2011	-0.50	d					0.11	1.17	26	245		No	
Jimerson, 2001	-0.31	d							20	246			

Holmes & Matthews, 1984	-0.34	d	0.036						44	575	11132	Yes	
Yoshiba, 1989	-0.60	Δ		0.61			-1.98	0.75	34	242		Yes	
Indicative effect size (weighted mean)	**-0.32**												
School uniforms													
No meta-analyses or systematic reviews. ES estimated from:													
Samuels 2003 - language arts	0.03	g(c)	0.11	0.16	-0.18	0.23	-0.06	0.03	1	2	9585	No	No
Samuels 2003 - mathematics	-0.06	g(c)	0.11	0.16	-0.26	0.15	-0.06	0.03	1	2		No	No
Sowell, 2012 mathematics	0.02										1152		
Indicative effect size	**0.00**												
Small Group Tuition													
Elbaum et al. 2000 (pairs)	0.4	Δ			0.24	0.56			19	116		Yes	
Elbaum et al. 2000 (small group-*NB only 1 study*)	1.61	Δ			0.75	2.48						Yes	
Lou et al. 2001 (with ICT), individual	0.16	d			0.12	0.2			122	486		Yes	
Lou et al. 2001 (with ICT), small group	0.31	d			0.2	0.43			122	486		Yes	
Lou et al. 2001 (with ICT), pairs compared with groups of 3-5	0.08	d							122	486	11317	Yes	
Slavin et al. 2011	0.31								20				
Indicative effect size (weighted mean)	**0.34**												
Social & Emotional Learning													
Durlak et al. 2011	0.27	g			0.15	0.39			213		270034	Yes	No
Multon et al. 1991	0.26	r							36	38	4998	Yes	No
Payton et al. 2008	0.28	d			0.14	0.41			29			Yes	
Indicative effect size (weighted mean)	**0.27**												
Sports participation													
Newman et al. 2010 (academic outcomes)	0.19	gc	0.08	0.12	0.03	0.35	0.15	0.34	2	2		No	Yes
Newman et al. 2010 (mathematics)	0.80	gc	0.11	0.16	0.58	1.02	0.66	0.98	1	2		No	Yes
Lewis, 2004	0.10	d		0.13					5			Yes	No
Shulruf, 2010 GPA	0.15	d			0.07	0.23			29	15	59804	Yes	
Indicative effect size (weighted mean)	**0.18**												
Summer schools													
Lauer, Akiba & Wilkerson 2006	0.05	g	0.01	0.04	0.01	0.11			14			Yes	Yes
Cooper et al 2000	0.26	d	0.01	0.06	0.24	0.28	-0.20	2.70	30				

Lewis, 2004	0.10	d	0.13			5		Yes	Yes
Indicative effect size (weighted mean)	**0.19**								
Teaching assistants									
No meta-analyses or systematic reviews. ES estimated from:									
Gerber et al. 2001 (with regular classes)	NPE (0.0 est)	d		ns	ns		1985		
Gerber et al. 2001 (with small classes)	NPE (-.15 est)	d		-0.13	-0.26		1985		
Blatchford et al. 2009	0.00						7578		

Total studies	**8147**	**Total pupils**	**1132463**

KEY		
Single studies with ES reported in italics		
Types of effect size		
Control group SD	Glass	*Δ*
SD unspecified	Cohen's d	d
Pooled SD	Hedges g	g
Pooled SD corrected for small sample bias	Hedges g corrected	gc
gc is also sometimes confusingly referred to as an 'unbiased estimator' or d		d*
Values in red calculated		
Distribution overlap, percentage of scores in the lower-mean group exceeded by the average score in the high-mean group.		U3
Fisher z transformed correlation coefficient.		Zr
No pooled effect	NPE	

Section 6: Bibliography of meta-analyses and other studies used to estimate effect sizes

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.

Aiello, N.C. (1981). *A meta-analysis comparing alternative methods of individualized and traditional instruction in science.* Virginia Polytechnic Institute and State University). ProQuest Dissertations and Theses, p 142. Retrieved from http://search.proquest.com/docview/303140546 (30/6/12).

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.

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.

Baker, D. P., Fabrega, R., Galindo, C., & Mishook, J. (2004). Instructional time and national achievement: Cross-national evidence. *Prospects*, 34:3, 311-334.

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.

Bangert-Drowns, R.L.Chen-Lin, C., Kulik, J.A., & Kulik, C.C. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research, 61*:2, 213-238.

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.

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.

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.

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.

Camilli, G., Vargas, S., & Yurecko, M. (2003). Teaching Children to Read. The Fragile Link between Science and Federal Education Policy. *Educational Policy Analysis Archives, 11*:15, ISSN 1068-2341.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-Analysis of the effects of early education interventions on cognitive and social development. Teachers College Record, 112:3, pp. 579–620.

Camnalbur, M., & Erdogan, Y. (2010). A Meta-Analysis on the Effectiveness of Computer-Assisted Instruction: Turkey Sample. *Educational Sciences: Theory & Practice, 8*:2, 497-505.

Cheung, A.C.K. & Slavin, R.E. (2011). The Effectiveness of Education Technology for Enhancing Reading Achievement: A Meta-Analysis. *From the Center for Research and Reform in Education, John Hopkins University: http://www. bestevidence. org/reading/tech/tech. html.*

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.

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.

Christmann, E. P., & Badgett, J. L. (2003). A meta-analytic comparison of the effects of computer-assisted instruction on elementary students' academic achievement. *Information Technology in Childhood Education Annual*, *2003*(1), 91-104.

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: pp. 237-248.

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

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.

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.

Crawford, S.T. (2011). Meta-analysis of the impact of after-school programs on students reading and mathematics performance. University of North Texas). ProQuest Dissertations and Theses, Retrieved from http://search.proquest.com/docview/919077210 (20/6/12).

Cummings, C., Laing, K., Law, J., McLaughlin, J., Papps, I., Todd, L., et.al. (2012). *Can Changing Aspirations and Attitudes Impact on Educational Attainment? A Review of Interventions.* Joseph Rowntree Foundation Report.

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.

Dignath, C, Buettner,G and 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.

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.

Durlak, J.A. & Weissberg, R.P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.

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.

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.

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.

Elbaum, B., Vaughn, S., Hughes, M. T., Moody, S. W., & Schumm, J. S. (2000) (b). 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.

Fashola. O.S. (1998). *Review Of Extended-Day And After-School Programs And Their Effectiveness Report No. 24* Baltimore, Center for Research on the Education of Students Placed At Risk (CRESPAR), Johns Hopkins University.

Fuchs, L.S. & Fuchs, D. (1985). Effects of systematic formative evaluation A meta-analysis. *Exceptional Children*, 53.3 pp 199-208.

Gansle, K.A. (2005). The Effectiveness of School-Based Anger Interventions and Programs: A Meta-Analysis*. Journal of School Psychology, 43,* 321-341.

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.

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.

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.

Ginsburg-Block, M.D., Rohrbeck, C.A., & Fantuzoo, J.W. (2006). A Meta-Analytic Review of Social, Self-Concept and Behavioral Outcomes of Peer-Assisted Learning. *Journal of Educational Psychology, 98* (4), 732-749.

Glass, G.V. & Smith, M.L. (1978). *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 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.

Gollwitzer, P.M., & Sheeran, P. (2006). Implementation Intentions and Goal Achievement: A Meta-Analysis of Effects and Processes. *Advances in Experimental Social Psychology, 38*, 70-119.

Gonzalez, J.E., Nelson, R.J., Gutkin, T.B., Saunders, A., Galloway, A. & Shwery, C.S. (2004). Rational Emotive Therapy with Children and Adolescents: A Meta-Analysis. *Journal of Emotional and Behavioral Disorders, 12*:4, 222-235.

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

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.

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.

Horak, V.M. (1981). A Meta-Analysis of Research Findings on Individualized Instruction in Mathematics. *Journal of Educational Research*, 74:4 p 249.

Igel, C.C., 2010. *The Effect of Cooperative Learning Instruction on K-12 Student Learning: A Meta-Analysis of Quantitative Studies from 1988 to 2009.* University of Virginia, Dissertation retrieved from Pro-Quest Dissertations.

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.

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.

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.

Johnson, D.W., Johnson, R.T. & Beth-Stanne, M. (2000*). Cooperative Learning Methods: A Meta-Analysis*. Minneapolis, MN: University of Minnesota.

Jun, S.W., Ramirez, G., & Cumming, A. (2010). Tutoruing Adolescents in Literacy: A Meta-Analysis. *Journal of Education, 45 (2),* 219-238.

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.

Karoly, L., Kilburn, R., Cannon, J.S. (2005). *Early Childhood Interventions: Proven Results, Future Promise*. Rand Corporation. Available at: http://www.rand.org/pubs/monographs/2005/RAND_MG341.pdf (30/6/12).

Kingston, N. & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice, 30*:4, 28-37.

Klauer, K.J. & Phye, G.D. (2008). Inductive Reasoning: A Training Approach. *Review of Educational Research*, 78.1 pp 85-123.

Kluger, A.N. & De Nisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119:2 pp 254-284.

Kulik, C.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 pp. 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., & Kulik, C. L. C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly*, 36:2, 73-77.

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 pp 443-484.

Layzer, J.I., Goodson, B.D., Bernstein, L. & Price, C. (2001). *National Evaluation of Family Support Programs. Final Report Volume A: A Meta-Analysis*. Administration for Children, Youth and Families (DHHS), Washington DC.

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.

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)

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

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* MPR Associates, Inc., Berkeley, CA/ Office of Vocational and Adult Education, U.S. Department of Education.

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, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, *22*:3, 215-243.

Lou Y, Abrami P.C., Spence J.C., Poulsen, C., Chambers B., &D'Apollonia S. (1996). Within-Class Grouping: A Meta-Analysis. *American Educational Research Association, 66* (4), 423-458.

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.

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, pp 176-183.

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.

Ma, X., & Kishor, N. (1997). Assessing the Relationship between Attitude towards Mathematics and Achievement in Mathematics: A Meta-Analysis. *Journal for Research in Mathematics Education. 28:1,* 26-47.

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 Children and Youth Services Review 32, 506–519

Martins, P.S. (2009). *"Individual Teacher Incentives, Student Achievement and Grade Inflation,"* IZA Discussion Papers 4051, Institute for the Study of Labor (IZA).

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.

Morphy P. & Graham S. (2012). Word Processing Programs and Weaker Writers/Readers: A Meta-Analysis of Research Findings, *Reading and Writing, 25*, 641-678.

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.

Nelson, G., Westhues, A. & MacLeod, J. (2003). A Meta-analysis of Longitudinal Research on Preschool Prevention Programs for Children. *Prevention and Treatment*, 6(2), 46-50.

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.

Niemiec, R.P. & Walberg, H.J. (1985). Computers and Achievement in the Elementary Schools. *Journal of Educational Computing Research*, 1:4, 435-40.

Nye, C, Turner, J., & Schartz, C (2006). *Approaches to Parent Involvement for Improving the Academic Performance of Elementary School Children*. Campbell Collaboration, Campbell Library of Systematic Reviews.

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.

Patall, E.A., Cooper, H., & Allen, A.B. (2010). Extending the School Day or School Year: A Systematic Review of Research (1985-2009). *Review of Educational Research, 80*:3, 401-436.

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).

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.

Petscher, Y. (2009). A Meta-Analysis of the Relationship between Student's Attitudes towards Reading and Achievement in Reading. *Journal of Research in Reading, 33*:4, 335-355.

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.

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.

Rogers, K. B. (2007). Lessons Learned About Educating the Gifted and Talented A Synthesis of the Research on Educational Practice. *Gifted Child Quarterly*, 51:4, 382-396.

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.

Romero C.C. 2009. *Cooperative learning instruction & science achievement for secondary and early post-secondary students: A systematic Review*. Dissertation, Colorado State University.

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)

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.

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*

Scott-Little, C. (2002). Evaluations of After-School Programs: A Meta-Evaluation of Methodologies and Narrative Synthesis of Findings. *American Journal of Evaluation*, 23:4, 387-419.

Senechal, M. & Young, L. (2008). The Effect of Family Literacy Interventions on Children's Acquisition of Reading from Kindergarten to Grade 3: A Meta-Analytic Review. *Review of Educational Research, 78*:4, 880-907.

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.Sitzmann, T., Kraiger, K., Stewart, D. & Wisher, R. (2006). The Comparative Effectiveness Of Web-Based And Classroom Instruction: A Meta-Analysis. *Personnel Psychology* 59, 623–664.

Slavin, R.E. (1989). Class Size and Student Achievement: Small Effects of Small Classes. *Educational Psychologist*, 24. Pp 25-77

Slavin, R.E. (1990). Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis. *Review of Educational Research*, 60:3, 471.

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.

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).

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.

Steenbergen-Hu, S., & Moon, S. M. (2011). The effects of acceleration on high-ability learners: A meta-analysis. *Gifted Child Quarterly*, 55:1, 39-53.

Sowell, R. E. (2012). The relationship of school uniforms to student attendance, achievement, and discipline (Doctoral dissertation, Liberty University: ProQuest Dissertations and Theses).

Tamim, R.M. (2009). *Effects of Technology on Students' Achievement: A Second-Order Meta-Analysis.* Unpublished PhD dissertation Concordia University, Montreal Canada (ProQuest Dissertations and Theses).

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.

Tamir, P. (1985). Meta-analysis of cognitive preferences and learning. *Journal of Research in Science Teaching* 22:1 pp 1-17.

Tanner, E., Brown, A., Day, N. Kotecha, M., Low, N., Morrell, G., Brown, V., Collingwood, A., Chowdry, H., Greaves, E., Harrison, C., Johnson, G., & Purdon, S. (2011). *Evaluation of Every Child a Reader (ECar):* Research Report DFE-RR114 London: Department for Education.

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.

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.

Torgerson C.J., Brooks, G. & Hall, J. (2006). A Systematic Review of the Research Literature on the use of Phonics in the Teaching of Reading and Spelling. *Research Report, No.711*

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.

Vaughn, V. L., Feldhusen, J. F., & Asher, J. W. (1991). Meta-analyses and review of research on pull-out programs in gifted education. *Gifted Child Quarterly,* 35:2, 92-98.

Walberg, H.J. (1982). What makes schooling effective? *Contemporary Education Review*, 1, 1–34.

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.

Waxman, H.C., Connell, M.L., & Gray, J. (2002). *A Quantitative Synthesis of Recent Research on the Effects of Teaching and Learning With Technology on Student Outcomes*. Naperville, Illinios.

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.

Wheeler, M. E., Keller, T. E., & DuBois, D. L. (2010). Review of Three Recent Randomized Trials of School-Based Mentoring. *Social Policy Report*, 24:3,1-27

Wiliam, D. (2002). '*Linking Research and Practice: knowledge transfer or knowledge creation?*' Plenary presentation at the annual conference of the North American Chapter of the International Group for the Psychology of Mathematics Education, Athens, Georgia, USA, October 2002.

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.

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.

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.

Woessmann, L. (2010). *Cross-Country Evidence on Teacher Performance Pay.* CESifo Working Paper No. 3151 Category 5: Economics Of Education Munich: CESifo.

Wood, S., & Mayo-Wilson, E. (2012). School-Based Mentoring for Adolescents A Systematic Review and Meta-Analysis. *Research on Social Work Practice*, 22:3, 257-269.

Zief, S.G., Lauver, S., Maynard, R.A. (2006). *Impacts of After-School Programs on Student Outcomes* Oslo: The Campbell Collaboration

Zvoch K., & Stevenes J.J, (2013). Summer School Effects in a Randomized Field Trial. *Early Childhood Research Quarterly, 28*, 24-32.

Technical Appendices References

Aguinis, H., Pierce, C. A., Bosco, F. A., Dalton, D. R., & Dalton, C. M. (2011). Debunking myths and urban legends about meta-analysis. *Organizational Research Methods*, 14(2), 306-331.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher,* 13(6), 4-16.

Bloom, H.S., Hill, C.J., Black, A.R. and Lipsey, M.W. (2008). Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions. *Journal of Research on Educational Effectiveness*, 1.4 pp 289-328.

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, 1(2), 97-111.

Burtless, G. (ed.) (1996). *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success.* Washington, D.C.: Brookings Institute Press.

Chan, M. E., & Arvey, R. D. (2012). Meta-Analysis and the Development of Knowledge. *Perspectives on Psychological Science*, 7(1), 79-92.

Coe R, (2002). *It's the Effect Size, Stupid: What effect size is and why it is important* Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England, 12-14 September 2002. Available at: http://www.leeds.ac.uk/educol/documents/00002182.htm (accessed 2/3/11).

Coe, R. (2004) 'Issues arising from the use of effect sizes in analysing and reporting research' in I. Schagen and K. Elliot (Eds) *But what does it mean? The use of effect sizes in educational research.* Slough, UK: National Foundation for Educational Research. Available at: http://www.nfer.ac.uk/nfer/publications/SEF01/SEF01.pdf

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.O., Hornik, R.C., Phillips, D.C., Walker, D.F. & Weiner, S.S. (1980). *Toward reform of program evaluation: Aims, methods, and institutional arrangements*. San Francisco, Ca.: Jossey-Bass.

Department for Education and Skills (DfES) (2004). *Statistics of Education: Variation in Pupil Progress* 2003 London: Department for Education and Skills. Available at: http://www.education.gov.uk/rsgateway/DB/SBU/b000481/b02-2004v2.pdf (accessed 15/06/11).

Department for Children, Schools and Families (DCSF). *Departmental Report* 2009. London: Department for Children Schools and Families. Available at: http://media.education.gov.uk/assets/files/pdf/d/departmental%20report%202009.PDF (accessed 18/05/13).

Dillon, J.T., (1982). Superanalysis. *American Journal of Evaluation* 3(4) pp 35-43.

Glass G.V., McGaw, B. & Smith, M.L. (1981). *Meta-analysis in social research* Beverly Hills, Ca: Sage.

Glass, G.V. (1977). Integrating Findings: The Meta-Analysis of Research. *Review of Research in Education*, 5: 351-379.

Greenwald, R, Hedges, LV, and Laine, RD (1996). 'The Effect of School Resources on Student Achievement'. *Review of Educational Research*, 66(3), 361-396.

Hanushek, E., and Woessman, L. (2010). The economics of international differences in educational achievement. *NBER Working Paper 15949*, April 2010, p18

Hattie, J.A. (1992). Measuring the effects of schooling. *Australian Journal of Education*, 36, 5-13.

Hattie, J.A. (2008). *Visible Learning.* London: Routledge.

Hedges, L.V. & Olkin, I. (1985). *Statistical Methods for Meta-Analysis.* Orland, Florida: Academic Press.

Higgins, S. & Simpson, A. (2011). Book Review: Visible Learning. *British Journal of Educational Studies* 59.2 pp 90-92.

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.

Hill, C.J., Bloom, H.S., Black, A.R. & Lipsey, M.W. (2007). *Empirical Benchmarks for Interpreting Effect Sizes in Research* MDRC Working Papers on Research Methodology New York: MDRC Avaliable at: www.mdrc.org/publications/459/full.pdf (accessed 17/5/11).

Holmlund, H. McNally, S. and Viarengo, M. (2008) : *Does money matter for schools?,* IZA discussion papers, No. 3769, http://nbn-resolving.de/ urn:nbn:de:101:1-20081126249

Ioannidis, J.P.A. & Trikalinos, T.A. (2007). The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *Canadian Medical Association Journal* 176 p 8.

Levačić, R., & Vignoles, A. (2002). Researching the links between school resources and student outcomes in the UK: a review of issues and evidence. *Education Economics* 10 (3), 312-331.
Luyten, H. (2006). 'An empirical assessment of the absolute effect of schooling: regression/ discontinuity applied to TIMSS-95'. *Oxford Review of Education*, 32: 3, 397-429.
Luyten, H., Peschar, H., Coe, R. (2008). Effects of Schooling on Reading Performance, Reading Engagement, and Reading Activities of 15-Year-Olds in England. *American Educational Research Journal* 45 (2) pp 319-342.
Pugh, G., Mangan, J., & Gray, J. (2011). Do increased resources increase educational attainment during a period of rising expenditure? Evidence from English secondary schools using a dynamic panel analysis. *British Educational Research Journal*, 37(1), 163-189.
Marzano, R.J. (1998). *A Theory-Based Meta-Analysis of Research on Instruction*. Aurora, Colorado, Mid-continent Regional Educational Laboratory. Available at: http://www.mcrel.org:80/topics/products/83/ (viewed 31/05/11).
National Foundation of Educational Research (2011). *Appendix A: Progress in International Reading Literacy Study (PIRLS) 2011 Overview.* Slough: NFER. Available at http://www.nfer.ac.uk/publications/PRTZ01/PRTZ01App.pdf (accessed 18/05/13).
OECD (2009). *PISA 2009 Results: Learning Trends.* Paris: OECD. Available at http://www.oecd.org/pisa/pisaproducts/48852742.pdf (accessed 18/05/13).
Schagen, I. & Hodgen, E. (2009). *How Much Difference Does It Make? Notes on Understanding, Using, and Calculating Effect Sizes for Schools* Wellington NZ: Research Division, Ministry of Education Available at: http://www.educationcounts.govt.nz/publications/schooling/36097/2 (accessed 5/6/11).
Sipe, T. & Curlette, W.L. (1997). A Meta-Synthesis Of Factors Related To Educational Achievement: A Methodological Approach To Summarizing And Synthesizing Meta-Analyses. *International Journal of Educational Research* 25.7. pp 583-698.
Slavin, R. & Smith, D. (2009). The Relationship Between Sample Sizes and Effect Sizes in Systematic Reviews in Education. *Educational Evaluation and Policy Analysis* 31.4: 500-506.
Smith, N.L. (1982). Evaluative Applications of Meta- and Mega-Analysis. *American Journal of Evaluation* 3(4) pp 43.
Steele, F., Vignoles, A., & Jenkins, A. (2007). The effect of school resources on pupil attainment: a multilevel simultaneous equation modeling approach. *Journal of Royal Statistical Society* Series A,.170(3), 801-824.
Tanner, E., Brown, A., Day, N., Kotecha, M., Low, N., Morrell, G., Brown, V., Collingwood, A., Chowdry, H., Greaves, E., Harrison, C., Johnson, G., & Purdon, S. (2011). *Evaluation of Every Child a Reader (ECar): Research Report DFE-RR114* London: Department for Education.
Tymms, P. (2004). Are standards rising in English primary schools? *British Educational Research Jounrnal*, 30(4), 477-494.
Tymms, P. & Merrell, C. (2007). *Standards and quality in English primary schools over time: the national evidence.* Primary Review Research Survey 4/1. Cambridge: The Primary Review.
Vignoles, A., Levačić, R., Walker, J., Machin, S. & Reynolds, D. (2000). *The Relationship Between Resource Allocation and Pupil Attainment: A Review* London: Department for Education and Employment.
Walberg, H. J. (1984). Improving the productivity of America's schools. *Educational Leadership*, 41(8), 19-27.
Wiliam, D. (2002). '*Linking Research and Practice: knowledge transfer or knowledge creation?*' Plenary presentation at the annual conference of the North American Chapter of the International Group for the Psychology of Mathematics Education, Athens, Georgia, USA, October 2002.
Wiliam, D. (2010). 'Standardized Testing and School Accountability'. *Educational Psychologist*, 45: 2, 107-122.
Wiliam, D., Lee, C., Harrison, C. & Black P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education*, 11, 1 49-65.