Toolkit of Strategies to Improve Learning

Summary for Schools
Spending the Pupil Premium

By

Professor Steve Higgins,

Durham University

Dr Dimitra Kokotsaki
and
Professor Robert Coe,

CEM Centre, Durham University

May 2011

CEM
Centre for Evaluation & Monitoring

Durham University

Toolkit to improve learning

The aim of the Government's new Pupil Premium is to raise achievement among disadvantaged children[1]. It will provide additional funding for disadvantaged pupils to ensure they benefit from the same educational opportunities as pupils from wealthier families. Although the precise mechanism for allocation and evaluation is still being finalised, the amount per pupil has been announced initially as £430 for 2011/12, then perhaps rising to as much as £1750 by 2014/15.

The Pupil Premium has a number of wider aims:

- to increase social mobility;
- to enable more pupils from disadvantaged backgrounds to get to the top Universities;
- to reduce the attainment gap between the highest and lowest achieving pupils nationally.

Simply spending more on children from less affluent backgrounds, however, will not necessarily improve their learning or their aspirations. There is no direct link between spending on schools and outcomes for pupils. Extensive research in this area shows that it is a complex issue, indicating that the way the money is spent is crucial. So if the Pupil Premium is to succeed in achieving its ambitious goals, the choices that schools make in allocating the money will be vital so that the funding can help raise pupils' attainment and aspirations.

The aim of this document is therefore to summarise some of the research evidence on improving learning and attainment to help schools to make more informed choices about how to support their pupils who are eligible for the additional funding. In each area we have identified different approaches to improving learning in schools, and identified the strength of the existing research evidence and then made an estimate of the costs of adopting the approaches. The toolkit also provides guidance on whether the approaches are applicable to primary or secondary school settings, and in which core subjects – English, maths or science.

Applying this knowledge in schools about each approach, to ensure that the impact of any changes benefits pupils' attainment will still be challenging, as there is no simple solution or guaranteed bet. The aim of the toolkit is to encourage schools and teachers to make their own informed choices and adopt a more 'evidence based' approach: they will need to monitor the effect of their chosen approach and evaluate the investment to ensure it is having the desired effect.

We also present a short summary and an outline of what we see as the relative benefit of the different approaches. Two appendices indicate the approximate levels of additional funding a school is likely to receive initially according to its size and the

[1] For the latest Government guidance on the Pupil Premium, see: http://www.education.gov.uk/schools/teachingandlearning/assessmentandachievement/premium/a0076063/pupil-premium-what-you-need-to-know

percentage of pupils receiving free school meals (Appendix 1). £430 is about equivalent to two full days of additional teacher time per pupil per year, or a bit less than a month of a teacher's full and undivided attention over the course of their school career. If the Pupil Premium is increased to about £1750 by 2014/15 this is nearer eight days of teacher time per year or nearer four months total schooling (Appendix 2). The challenge is to use this effectively to support improved educational outcomes for these learners.

Finally, we should point out that it is always challenging to apply the findings of educational research from one context to another. However our aim is to help schools to identify potentially productive strategies and approaches which they will then need to develop and evaluate to suit their own setting and context.

Spending for learning: linking resources and learning outcomes

The challenge to establish a clear link between educational expenditure and pupils' learning is harder than you would think. 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. Much of the early research in this area failed to find a convincing connection. More recent research suggests that there is a link between spending and outcomes, but that it is a complex picture.

Investing in better learning, or 'spending for learning', is therefore not easy, particularly when the aim is to support disadvantaged learners. 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.

Though there is not clear evidence of the link between additional spending and learning, we interpret this to mean that it is difficult to spend additional resource effectively. On average it may be hard to find a link, but there must be some areas which offer a better bet than others, and this is what this toolkit shows.

We are also not suggesting that all educational aims and outcomes are captured in the literature that we have pulled together. Most of the measures are traditional measures of attainment, curriculum tests and examinations or standardised measures. Some studies include assessment of attitudes and beliefs, but even these may have a specific focus in the studies, depending on the research aims. The rationale for inclusion and evaluation is presented in Appendix 3, with full details in the technical report which accompanies the toolkit.

Our solution to the challenge of linking spending with learning is to focus on what the evidence indicates is effective in improving teaching and learning using typical measures, and then working out what additional costs are associated with these approaches, so as to highlight the issues for schools to explore. A number of other, less effective approaches have been included as a comparison or where they are better known to provide a reference point.

The summaries combine evidence from a range of different research studies. In all cases there is a range of effects which have been combined into a single average for each area. We are not claiming that this will necessarily be the impact when schools try them out. 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 this evidence 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.

The toolkit purposefully compares broad approaches against one another, rather than advocating specific schemes, although we hope in time to develop a number of links to specific resources for teachers to explore further, and make their own choices. Whatever approach is chosen, of course, it will still be necessary for each school to evaluate the actual benefits of any changes in a real context to ensure the investment really does help pupils from low-income families achieve their educational potential.

This summary document represents the first stage of the toolkit. The next stages will involve an evaluation of how a number of schools fare in actually implementing the toolkit, and the development of the toolkit into a more interactive online resource for schools and teachers.

Toolkit to improve learning: summary overview

Approach	*Potential gain* [2]	*Cost*	*Applicability*	*Evidence estimate*	*Overall cost benefit*
Effective feedback	+ 9 months	ff	**Pri**, **Sec** **Maths Eng** Sci	☆ ☆ ☆	Very high impact for low cost
Meta-cognition and self-regulation strategies.	+ 8 months	ff	**Pri, Sec**, **Eng Maths Sci**	☆ ☆ ☆ ☆	High impact for low cost
Peer tutoring/ peer-assisted learning	+ 6 months	ff	**Pri**, Sec **Maths Eng**	☆ ☆ ☆ ☆	High impact for low cost
Early intervention	+ 6 months	fffff	**Pri**, **Maths Eng**	☆ ☆ ☆ ☆	High impact for very high cost
One-to-one tutoring	+ 5 months	fffff	**Pri**, Sec Maths Eng	☆ ☆ ☆ ☆	Moderate impact for very high cost
Homework	+ 5 months	f	Pri, **Sec** **Maths Eng** Sci	☆ ☆ ☆	Moderate impact for very low cost
ICT	+ 4 months	ffff	Pri, Sec All subjects	☆ ☆ ☆ ☆	Moderate impact for high cost

[2] Maximum approximate advantage over the course of a school year that an 'average' student might expect if this strategy was adopted – see Appendix 3.

Assessment for learning	+ 3 months	ff	Pri, **Sec** Maths Eng	☆	Moderate impact for moderate cost
Parental involvement	+ 3 months	fff	**Pri**, Sec **Maths Eng** Sci	☆ ☆ ☆	Moderate impact for moderate costs
Sports participation	+ 3 month	fff	Pri, **Sec** Maths Eng Sci	☆ ☆	Moderate impact for moderate cost.
Summer schools	+ 3 months	fff	Pri, Sec Maths Eng	☆ ☆	Moderate impact for moderate cost
Reducing class sizes	+ 3 months	fffff	**Pri**, Sec **Maths Eng**	☆ ☆ ☆	Low impact for very high cost
After school programmes	+ 2 months	ffff	**Pri**, **Sec** **Maths Eng** Sci	☆ ☆	Low impact for moderate cost.
Individualised instruction	+ 2 month	ff	Pri, **Sec** Maths Eng Sci	☆ ☆ ☆	Low impact for low cost.
Learning styles	+ 2 month	f	Pri, **Sec** All subjects	☆ ☆	Low impact, low or no cost
Arts participation	+ 1 month	ff	Pri, **Sec** Maths Eng Sci	☆ ☆ ☆	Very low impact for moderate cost.

Performance pay	+ 0 months	£££	Pri, **Sec** **Maths Eng** Sci	☆	Very low/no impact for moderate cost
Teaching assistants	+ 0 months	££££	**Pri**, **Sec** **Maths Eng** Sci	☆ ☆	Very low/no impact for high cost
Ability grouping	± 1 month	£	**Pri**, **Sec** **Maths Eng** Sci	☆ ☆ ☆	Very low or negative impact for very low or no cost
Block scheduling and timetabling	± 1 month	£	Pri, **Sec** Maths Eng **Sci**	☆ ☆	Very low or negative impact for very low or no cost
School uniforms	± 1 month	£	Pri, Sec Maths Eng Sci	☆	Very low or negative impact for very low or no cost

Toolkit to improve learning: summary overview

Approach	*Potential gain* [2]	*Cost*	*Applicability*	*Evidence estimate*	*Overall cost benefit*
Effective feedback	+ 9 months	ff	**Pri**, **Sec** **Maths Eng** Sci	☆ ☆ ☆	Very high impact for low cost
Meta-cognition and self-regulation strategies.	+ 8 months	ff	**Pri, Sec**, **Eng Maths Sci**	☆ ☆ ☆ ☆	High impact for low cost
Peer tutoring/ peer-assisted learning	+ 6 months	ff	**Pri**, Sec **Maths Eng**	☆ ☆ ☆ ☆	High impact for low cost
Early intervention	+ 6 months	fffff	**Pri**, **Maths Eng**	☆ ☆ ☆ ☆	High impact for very high cost
One-to-one tutoring	+ 5 months	fffff	**Pri**, Sec Maths Eng	☆ ☆ ☆ ☆	Moderate impact for very high cost
Homework	+ 5 months	f	Pri, **Sec** **Maths Eng** Sci	☆ ☆ ☆	Moderate impact for very low cost
ICT	+ 4 months	ffff	Pri, Sec All subjects	☆ ☆ ☆ ☆	Moderate impact for high cost

[2] Maximum approximate advantage over the course of a school year that an 'average' student might expect if this strategy was adopted – see Appendix 3.

Assessment for learning	+ 3 months	ff	Pri, **Sec** Maths Eng	☆	Moderate impact for moderate cost
Parental involvement	+ 3 months	fff	**Pri**, Sec **Maths Eng** Sci	☆ ☆ ☆	Moderate impact for moderate costs
Sports participation	+ 3 month	fff	Pri, **Sec** Maths Eng Sci	☆ ☆	Moderate impact for moderate cost.
Summer schools	+ 3 months	fff	Pri, Sec Maths Eng	☆ ☆	Moderate impact for moderate cost
Reducing class sizes	+ 3 months	fffff	**Pri**, Sec **Maths Eng**	☆ ☆ ☆	Low impact for very high cost
After school programmes	+ 2 months	ffff	**Pri**, **Sec** **Maths Eng** Sci	☆ ☆	Low impact for moderate cost.
Individualised instruction	+ 2 month	ff	Pri, **Sec** Maths Eng Sci	☆ ☆ ☆	Low impact for low cost.
Learning styles	+ 2 month	f	Pri, **Sec** All subjects	☆ ☆	Low impact, low or no cost
Arts participation	+ 1 month	ff	Pri, **Sec** Maths Eng Sci	☆ ☆ ☆	Very low impact for moderate cost.

Performance pay	+ 0 months	£££	Pri, **Sec** **Maths Eng** Sci	☆	Very low/no impact for moderate cost
Teaching assistants	+ 0 months	££££	**Pri**, **Sec** **Maths Eng** Sci	☆ ☆	Very low/no impact for high cost
Ability grouping	± 1 month	£	**Pri**, **Sec** **Maths Eng** Sci	☆ ☆ ☆	Very low or negative impact for very low or no cost
Block scheduling and timetabling	± 1 month	£	Pri, **Sec** Maths Eng **Sci**	☆ ☆	Very low or negative impact for very low or no cost
School uniforms	± 1 month	£	Pri, Sec Maths Eng Sci	☆	Very low or negative impact for very low or no cost

After School Programmes

What is it?

After-school programmes are services offered during term time and at the end of the school day where children or young people are involved in planned activities which are supervised by adults. However, the goals, objectives and approaches of the programmes may vary greatly. Sometimes they 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 findings indicate that participants in after-school programmes score higher on measures of academic achievement. However the gains are modest on tested attainment of reading or mathematics (with an effect size between .13 and .31) but no clear pattern of impact. There is evidence that there are wider benefits for low-income students in terms of behaviour and relationships with peers. For these families, attending a formal after-school programme where children spend more time in academic and enrichment activities reliably but modestly linked with their learning and behaviour, relations with peers and their attitudes to learning.

Impact summary: + 2 months (effect size 0.16)

How secure is the evidence?

There are a number of reviews and a comprehensive meta-analysis, mainly using data from the USA, but with broadly similar findings in the UK.

Strength of the evidence: ☆ ☆

What are the costs?

Analysis suggests that enthusiasm for after-school programs has outpaced the research base indicating the need for more rigorous evaluations with outcome measures that demonstrate effectiveness.

Most of the cost estimates after-school programmes come from the US, with a wide range from $8 to $36 (£5 - £25) per day for each young person involved, with an average at about (£10). The costs of different after-school programmes depend on a number of factors, including decisions about the types of activities provided, the staff-to-young person ratio, and the extent of investment in factors such as fundraising and the future sustainability of the programme. However assuming £10 a day for about half a school year (100 days) comes to about £1000 per pupil.

Cost summary: ££££

How applicable is it?

Programmes with greater structure, a strong link to the school-day curriculum, well qualified and well-trained staff, and opportunities for one-to-one tutoring seem particularly promising and more clearly related to academic benefits. Programmes may not be equally effective with all students. At risk children may benefit more as do younger children (5 – 10 year olds). In terms of subjects, positive effects for reading were highest for younger primary pupils and in secondary schools. For mathematics the gains were higher for older primary and secondary pupils. However it is harder to attract and retain pupils as they get older.

Further information: There is a 2007 review by the Collaborative for Academic, Social and Emotional Learning (CASEL): http://www.casel.org/sel/meta.php .

Ability grouping

What is it?

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

How effective is it?

The evidence is consistent that though there may be some benefits for higher attaining pupils in some circumstances (e.g. gifted and talented programmes), these are largely outweighed by the negative effects on attitudes for middle and lower performing learners (with an average effect size of about 0.12). There is some evidence that effective and flexible grouping for particular tasks can be beneficial, or when effective teachers are assigned to low attaining groups. However, more routine setting arrangements tend to undermine low attainers' confidence and the belief that effort is more important than ability. This is a clear example of what *not* to do if you want low income pupils to benefit as they are more likely to be assigned to lower attaining groups. Some reviews suggests the overall impact on learners is negative (i.e. over time their performance deteriorates – an effect size of -.06).

Impact summary: ± 1 month (effect size .12 / -.06)

How secure is the evidence?

The evidence is robust and has accumulated over at least 30 years of research. Although there is some variation depending on methods and research design, the conclusions are relatively consistent. One of the complexities of the issue is the language of 'ability' which implies a fixed construct rather than approaches which suggest learners can improve and which emphasise success through effort.

Strength of the evidence: ☆ ☆ ☆

What are the costs?

Ability grouping is an organisational strategy which has few, if any, financial costs associated with it. Schools should be aware of the detrimental effects on low attaining pupils, of which those receiving free school meals will be likely to form a large proportion. If schools adopt mixed ability grouping they more likely to use inclusive teaching strategies and to promote higher aspirations for their pupils.

Cost summary: £

How applicable is it?

The evidence suggests that impact of setting is most detrimental to low attaining pupils in mathematics who do better in mixed attainment groups. The effects appear to be less evident in English and Science, though negative effects are more commonly reported for low attaining pupils across the curriculum. The evidence is available particularly for upper primary and lower secondary education.

Further information:

An older, but thought provoking article by Adam Gamoran from 1992 available at: http://gayleturner.net/Gamoran_Is_Ability_Grouping_Equitable.pdf .

Assessment for learning

What is it?

Assessment for learning (AfL) is based on the idea that students need a clear understanding of what it is that they need to learn and evidence about their current level of performance, so they can close this gap. The concept developed from definitions of formative assessment and feedback and has clear links with other approaches such as mastery learning. A number of classroom strategies have been developed to support the approach such as using the traffic lights metaphor to indicate learners' confidence or personalizing the process with characters (such as Walt, Wilf, Tib and Oli) for younger learners.

How effective is it?

The theoretical evidence for the benefits of feedback on learning is strong. Reviews consistently suggest that the potential benefit in terms of effect size is large (about 0.8). It is less easy to make this work in practice in a classroom setting with a large group of learners. The Kings Medway Oxford Formative Assessment Project (KMOFAP), one of the few quantitative studies of Assessment for Learning, found an average effect of 0.32, or about half a GCSE grade.

Impact summary: + 3 months (effect size .32)

How secure is the evidence?

There are a number of robust reviews and meta-analyses of the importance of feedback for learning. There are no meta-analyses of the impact of Assessment for Learning or the wider use of formative assessment in schools. There is some evidence of the importance of supporting teachers' and pupils' understanding of the purposes and processes involved (see also metacognitive approaches).

Strength of the evidence: ☆

What are the costs?

The actual costs of implementing Assessment for Learning are low, however the evidence suggests that professional development is needed. In the KMOFAP project studies, the teachers involved had about six days of professional development and were supported in implementing and evaluating the approach. The costs were estimated at about £2000 per teacher.

Cost summary: ££

How applicable is it?

There is limited research evidence about its applicability, however it is readily adaptable to all phases and subjects of education.

Further information: The national strategies produced support materials for schools in England which are still available at: http://nationalstrategies.standards.dcsf.gov.uk/node/97905 and the GTC's Teacher Learning Academy has a research summary with practical advice: http://www.gtce.org.uk/tla/rft/afl_prac0904/ .

Arts participation

What is it?

Participation in artistic and creative activities, including dance, drama, music, painting, sculpture etc, both in terms of performance and creation. Participation may be organised as regular weekly activities or more intensive programmes such as summer schools or residential courses.

How effective is it?

Overall the impact on academic learning is moderate, though greater effects have been identified for younger learners of primary school age in terms of impact on cognitive tests.

Impact summary: + 1 months (effect size 0.05)

How secure is the evidence?

There are a number of systematic reviews and meta-analyses which have consistenly found benefits, though these vary according to approach and age group studied so the effects are hard to generalise.

Strength of the evidence: ☆ ☆ ☆

What are the costs?

Costs vary considerably from small local junior drama groups with small annual subscription (about £20), through organised dance groups for young people at about £5/ session to high quality music tuition at about £35/hour. Costs are estimated at £150/ year, though it should be noted that some kinds of participation would be considerably more (e.g. nearer £1500 for individual music tuition).

Cost summary: ££

How applicable is it?

There is consistent evidence that participation in artistic and creative activities are beneficial. Effects 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.

Further information:

The Department for Culture, Media and Sport (DCMS) set up the Culture and Sport Evidence (CASE) programme was set up by to collect evidence about participation in culture and sport and their recent review is available at: http://www.culture.gov.uk/images/research/CASE-systematic-review-July10.pdf .

Block scheduling and timetabling

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 the 4x4 block seemed to produce higher cross-subject achievement than traditional schedules. However, 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 (effect sizes between 0.13 to 0.42); 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 (effect sizes between -0.15 to 1.55). There is not therefore sufficient evidence to support the introduction of block scheduling in secondary schools to raise attainment. It is important to be aware that the findings do not indicate that block schedules are likely to produce negative outcomes for pupils across subjects, but the findings on the benefits or positive effects are not strong enough to recommend adoption. It 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.

Impact summary: ± 1 month (effect size -.02 / 0.15)

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.

Strength of the evidence: ☆ ☆

What are the costs?

The costs of altering timetabling is mainly in terms of organisational or administrative time to set up the timetable. Additional costs are low.

Cost summary: £

How applicable is it?

Timetabling issues tend to affect secondary schools, though the time spent on different areas of the curriculum is of importance in primary education. The research has mainly looked at impact on mathematics, English and science.

Further information:

A review undertaken by the EPPI Centre is available at:
http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2476 .

Early intervention

What is it?

Early intervention approaches are where the aim is to ensure that young children have educationally based pre-school or nursery experience to prepare them effectively for school. Many programmes focus on disadvantaged children. Some also offer parental support.

How effective is it?

Overall the evidence suggests that early intervention is beneficial with above average levels of impact (an average effect size of 0.45). There is some evidence that these programmes need to be whole day (rather than half-day) and of longer duration (up to a year) rather than for shorter periods of time. The impact tends to wear off over time, however, though it tends to have a more durable effect on attitudes to school than measures of attainment.

Impact summary: + 6 months (effect size .0.45)

How secure is the evidence?

There are a number of systematic reviews and meta-analyses which have looked at the impact of early intervention. Most of these are from the US however, where children tend to start school at later age. Evaluations of Sure Start in the UK indicate some caution is needed in translating research evidence into practice, particularly at policy level.

Strength of the evidence: ☆ ☆ ☆ ☆

What are the costs?

The costs, understandably, are high as adult/child ratios in pre-school provision tend to be higher than school classes and family interventions have similar high costs. Estimates are around £900 per child per year, the Sure Start average was about £1000 in 2006.

Cost summary: £££££

How applicable is it?

Applicable to early years education. Most of the studies are in the US however, where children tend to start school later. Evaluation of Sure Start in the UK suggest that the benefits may be harder to achieve at larger scale.

Further information: The US Government's 'What Works' Clearinghouse contains details of early childhood intervention programmes with robust evidence of impact: http://ies.ed.gov/ncee/wwc/reports/Topicarea.aspx?tid=13 .

Effective Feedback

What is it?

Feedback is information given to the learner and/or the teacher about the learner's performance relative to the learning goals which then redirects or refocuses either the teachers or the learners actions to achieve the goal. It can be about the learning activity or *task* itself, about the *process* of the task or activity, about the student's management of their own learning or their *self-regulation* or about them as individuals (e.g.*"good girl")* Research suggests that feedback is best directed at the task and process level. Research suggests that it should be:

• about *challenging* tasks or goals (rather than easy ones);
• given *sparingly* (i.e. needs to be meaningful);
• more important to give feedback about what is *right* than what is *wrong*;
• important to be as *specific* as you can and, if possible, compare what they are doing right now with what they have done wrong before; and
• it should *encourage* them, and not threaten their self-esteem.

How effective is it?

Feedback studies have tend to have high effects on learning. However some studies also show that feedback can have negative effects so it is important to understand the potential benefits and limitations. This was part of the rationale for the design of Assessment for Learning. Research-based approaches which provide feedback to learners, such as Bloom's 'mastery learning', also tend to have a positive impact on learning when used in schools.

Impact summary: + 9 months (effect size 0.73)

How secure is the evidence?

There are a substantial number of reviews and meta-analyses of the effects of feedback. However some are theoretical studies, particularly in psychology exploring both positive and negative effects. Educational studies tend to identify positive benefits where the aim is to improve learning. Estimates of effect sizes from meta-analyses ranging from 0.74 to 1.13 are identified in one recent review. The challenge is making feedback work practically in the classroom.

Strength of the evidence: ☆ ☆ ☆

What are the costs?

The costs of providing more effective feedback are not high. One study even estimates that the impact of rapid feedback on learning is 124 time more cost effective that reducing class sizes! However it probably requires sustained professional development to improve classroom practice.

Cost summary: ££

How applicable is it?

Feedback has effects on all types of learning. Evidence in schools has focussed particularly on English, mathematics and to a lesser extent science.

Further information:

There is a good review by Valerie Shute for the Educational Testing Service (ETS) in the USA and a practical summary table of what to do to support learners in the summary: http://www.ets.org/Media/Research/pdf/RR-07-11.pdf .

Homework

What is it?

Homework refers to tasks given to pupils by their teachers to be completed outside of class, with the normal expectation that it will be completed at home. Common homework activities may be reading or preparing for work to be done in class, or practising and completing things already taught or started. It may include extended activities to develop pupils' inquiry skills or work such as revision for tests and 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 looked at homework to explore this issue. There is some evidence that when homework is used as an intervention it is effective in improving students' attainment (an effect size of 0.60). Overall the benefits are likely to be more modest.

Impact summary: + 5 months (effect size 0.36)

How secure is the evidence?

Homework has been extensively studied, both by looking at the connection (correlation) between homework and how well schools do, and by studying what happens when homework is introduced.

Strength of the evidence: ☆ ☆ ☆

What are the costs?

There are few costs associated with homework, though there are implications for staff time.

Cost summary: £

How applicable is it?

The research strongly suggests that it is more valuable at secondary school level and much less effective for children of primary school age. There is also an optimum level of between 1-2 hours per school day (longer for older pupils), but the effects tail off as the time students spend increases. Pupils also benefit from feedback on homework and effective integration with teaching in lessons.

Further information: The Northwest Regional Educational Laboratory in the USA has a useful summary: http://www.netc.org/focus/strategies/home.php .

Individualised instruction

What is it?

Individualised instruction is based on the idea that all learners are different and therefore have different needs, so an individualised or tailored approach to instruction ought to be more effective. Examples of this have been tried over the years in education, particularly in areas like mathematics where pupils have a individual sets of activities which they complete, moving on to the next after successful completion of a task. This has also been tried extensively in computer-based learning and other approaches such as Bloom's 'mastery learning' where, perhaps surprisingly, group approaches are more effective than individual tuition.

How effective is it?

Individualising instruction does not tend to be particularly beneficial. One possible interpretation is that the role of the teacher becomes too managerial in terms of organising and monitoring learning, but not interacting and using formative feedback to refocus effort. Effect sizes tend overall to be low, or even negative.

Impact summary: + 1 month (effect size 0.10)

How secure is the evidence?

There have been a number of meta-analyses which have found broadly similar effects. Confirmation comes from other areas such as learning with technology and Bloom's 'mastery learning' where group effects are higher than individual.

Strength of the evidence: ☆ ☆ ☆

What are the costs?

Costs are usually low, unless the approach uses technology (such as tutoring programs or integrated learning systems).

Cost summary: ££

How applicable is it?

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 and history and geography.

Further information:

A summary of some approaches and issues with individualised instruction can be found at: http://education.stateuniversity.com/pages/2085/Individualized-Instruction.html.

Information and communication technologies (ICT)

What is it?

The use of digital technologies to support learning, either through a teaching programme (computer-assisted learning), or where learners use technology in problem-solving or more open ended learning, or where teachers use technology to support learning, such as with an interactive whiteboard or learning platform (virtual learning environment). The range of approaches, equipment and content is varied, making it hard to draw clear conclusions.

How effective is it?

Overall studies consistently find that ICT is associated with better learning, however there is considerable variation in impact. The gains are usually moderate, and is certainly the case that it is more important to think about the way the technology is used which is important rather than the technology itself. Evidence also suggests that technology should be used to supplement other teaching, rather than replace more traditional approaches. There is some evidence it is more effective with younger learners.

Impact summary: + 4 months (effect size 0.35)

How secure is the evidence?

There is extensive research evidence of the impact of different technologies. It is relatively consistent and tends to show moderate benefits for technology use. However, due to the increasing pace of technological change, it is usually about yesterday's technology rather than today's and certainly makes it difficult to know what to buy for tomorrow.

Strength of the evidence: ☆ ☆ ☆ ☆

What are the costs?

The costs of investing in new technologies is high, but they are already part of the society we live in and most schools are already equipped with computers and interactive whiteboards. A personal netbook costs about £250 today and an interactive whiteboard set-up for class use about £2,000 - £3,000. 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. 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 on how well the technology is used.

Cost summary: £££

How applicable is it?

There is evidence across age groups and for most areas of the curriculum over the last 40 years or so, suggesting that the impact is relatively robust. In particular there is clear evidence that it more beneficial for areas like writing than spelling or mathematics.

Further information: The US based North Central Regional Educational Laboratory has a review of technology and learning which was updated in 2005: http://www.ncrel.org/sdrs/areas/issues/methods/technlgy/te800.htm .

Learning styles

What is it?

The idea is based on the notion that, as individuals, we all have different approaches, or styles of learning and that learning will therefore be more effective or more efficient if we are taught in accordingly. It has proved difficult to identify robust learning 'styles' reliably, however. As individuals our preferences change in different situations and over time. There is some evidence that cognitive style and task type may be connected (visualisation in some areas of mathematics is valuable, for example) and it is certainly helpful to have different representations of ideas when developing understanding, but it is unhelpful to assign learners to groups or categories on the basis of a learning style.

How effective is it?

Studies where targeted learning with activities that match an identified learning style have not shown convincingly that there is any benefit, particularly for low attaining pupils. In fact in some studies the controls did better than the learning styles groups. There may be some benefit in learners believing that they can succeed in a task if they can choose the particular approach they use. The effect sizes in independent meta-analyses are low (e.g. 0.14 or negative (-0.03), suggesting that only one or two pupils in a class of 30 might benefit from being taught in this way.

Impact summary: + 2 month (effect size 0.14)

How secure is the evidence?

The evidence for the lack of impact (and in some cases detrimental effect) of using learning styles approaches has been demonstrated in a number of studies and meta-analyses. The unreliability of learning styles tests and assessments has also been the subject of a number of reviews. Overall the picture is consistent and robust.

Strength of the evidence: ☆ ☆

What are the costs?

The costs are relatively low, though some of the available tests of learning styles require payment.

Cost summary: £

How applicable is it?

The lack of impact has been documented at all stages of education. 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.

Further information: A recent critique, published in *Psychological Science in the Public Interest* entitled "Learning Styles: Concepts and Evidence" by Harold Pashler and colleagues is available at:
http://psychologicalscience.org/journals/pspi/PSPI_9_3.pdf .

Meta-cognitive and self regulation strategies

What is it?

Metacognitive strategies are teaching approaches which make learners' thinking about learning more explicit in the classroom. This is usually through teaching pupils strategies to plan, monitor and evaluate their own learning. It is usually more effective in small groups so learners can support each other and make their thinking explicit through discussion. Self-regulation refers to managing one's own motivation towards learning as well as the more cognitive aspects of thinking and reasoning.

How effective is it?

Meta-cognitive approaches have a consistently high or very high levels of impact with meta-analyses reporting effect sizes between 0.59 and 0.73. These are substantial gains equivalent to moving a class from 50th place in a league table of 100 schools to about 25th. Encouragingly there is also evidence it is particularly helpful for low achieving pupils.

Impact summary: + 8 months (effect size .67)

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.

Strength of the evidence: ☆ ☆ ☆ ☆

What are the costs?

The 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 the use of the strategies as they use them in the classroom.

Cost summary: ££

How applicable is it?

The evidence suggests that it tends to be more effective with older pupils, and is beneficial for lower achieving students. Most of the studies have looked at the effects on English (reading and writing) or mathematics, though there is some evidence from other areas of the curriculum such as science, suggesting the benefits are likely to be more widely applicable.

Further information: There is an Education Resources Information Center (ERIC) digest in the USA which provides a sound basic overview, if a little dated: http://www.ericdigests.org/pre-9218/developing.htm .

Examples of metacognitive strategies can be found at: http://academic.pgcc.edu/~wpeirce/MCCCTR/metacognition.htm .

One-to-one tutoring

What is it?

This is where an individual pupil is removed from their class and given intensive remedial tuition, for short, regular sessions (about 30 mins, 3-5 times a week) over a set period of time (6-12 weeks).

How effective is it?

Evidence indicates that in areas like reading and mathematics it can enable learners to catch up with their peers. Meta-analyses suggest an average effect size of about 0.4, indicating that pupils might improve by about 4 or 5 months during the intensive programme.

Impact summary: + 5 months (effect size 0.40)

How secure is the evidence?

The evidence is consistent, particularly for younger learners who are behind their peers in primary schools and for subjects like reading and mathematics. Overall there is strong evidence for its benefits. Programmes which used experienced teachers who are given training are more effective than those using volunteers or classroom assistants. Evidence also suggest tutoring should be additional or supplemental to normal instruction, rather than replace it.

Strength of the evidence: ☆ ☆ ☆ ☆

What are the costs?

The costs are high as the support is intensive. A single pupil receiving 30 mins, 5 times a day for 12 weeks is about 4 full days of a teacher's time, £800 or so. 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.

Cost summary: ££££

How applicable is it?

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

Further information: The Best Evidence Encyclopaedia (BEE) has information on one-to-one tutoring: http://www.bestevidence.org/reading/strug/summary.htm .

Parent involvement

What is it?

Actively involving parents in supporting their children to improve children's learning at school.

How effective is it?

Although parent involvement is consistently associated with pupils' success at school, the evidence about *increasing* involvement to *improve* attainment is much less conclusive. This is particularly the case for poorer 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 intervention, 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 clearly also important, though again there is insufficient evidence to show that *changing* parents' aspirations for their children will *raise* their children's aspirations and achievement over the longer term. Two recent meta-analyses of studies in the USA suggest that the effects in primary and secondary schools are about 0.27 and 0.25 respectively.

Impact summary: + 3 months (effect size 0.25)

How secure is the evidence?

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

Strength of the evidence: ☆ ☆ ☆

What are the costs?

The costs of the different parent involvement approaches vary enormously, from running parent workshops and improving communications, which need little additional resource, to intensive family support programmes with specially trained staff.

Cost summary: ££

How applicable is it?

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

Further information: A summary of one of the more recent studies into parental involvement by William Jeynes at the Harvard Family Research Project is available at: http://www.hfrp.org/publications-resources/browse-our-publications/parental-involvement-and-student-achievement-a-meta-analysis (viewed 30/3/11).

The GTC also have a summary of a review on the Teacher Learning Academy site: http://www.gtce.org.uk/tla/rft/parent0206/ .

Peer tutoring/ peer-assisted learning strategies

What is it?

These are 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 usually takes the tutoring role and is paired with a younger tutee or tutees. Peer-Assisted Learning Strategies (PALS) is a structured approach for mathematics and reading requiring set periods of time for implementation of about 25-35 minutes 2 or 3 times a week. In the collaborative learning strategy 'Reciprocal Peer Tutoring' learners alternate between the role of tutor and tutee. The common characteristic is that the learners take on responsibility for aspects of teaching and for evaluating the success of the learner.

How effective is it?

The evidence of impact is relatively high (typically an effect size of 0.5 or above, 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.

Impact summary: 6 months gain (an effect size of 0.5)

How secure is the evidence?

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

Strength of the evidence: ☆ ☆ ☆ ☆

What are the costs?

There are few costs directly associated with the approach, though it does need some time to organise and set up, particularly in the early stages. Cross-age tutoring in particular needs some co-ordination as it involves at least two classes from different year groups working together. Training to support staff is usually recommended.

Cost summary: ££

How applicable is it?

Consistent positive effects have been found in different countries, across different age groups (from young children to adults) and in different areas of the curriculum.

Further information:

A summary article is available at: http://www.readingrockets.org/article/22029 . This site is funded by the US Department of Education's Office of Special Education Programs.

Performance pay

What is it?

We know that teachers are the most important part of the education system in terms of improving students' learning. Performance pay or performance-related pay is where there is an attempt to link a teacher's wages or bonus payments directly to their performance in the classroom. In the USA it is sometimes referred to as 'merit pay', and, due to federal government incentives through the Teacher Incentive Fund (TIF), there has been an increase in these approaches to link teachers' remuneration to student attainment by state governments. In India there is evidence of the benefit of performance pay in the private school sector but not the state sector. One key issue is how performance is measured and how closely this is linked to outcomes for learners. In the UK it was one component in the performance threshold assessment introduced in 2000, but very loosely connected and at the discretion of the head teacher.

How effective is it?

Estimates based on cross-national comparisons suggest that the effect size should be about 0.25. One study in the UK estimates the benefit is about half a GCSE grade, which is about the same level of effect. However when looking at more rigorous evaluations within countries such as the USA, the actual impact has been closer to zero.

Impact summary: + 0 months (effect size 0)

How secure is the evidence?

The evidence is not conclusive. Although there has been extensive research, most of this is either from correlational studies linking national levels of pay with general national attainment or from naturally occurring experiments where it is hard to control for other variables which may influence the impact of pay increases.

Strength of the evidence: ☆

What are the costs?

Increases are usually of the order of £2000 per teacher or about the equivalent of £70 per pupil across a class of 30 students.

Cost summary: ££

How applicable is it?

It is hard to draw clear conclusions from the different types of evidence. There is some evidence from the UK suggesting that the impact may be greater on low achieving pupils and possibly worth as much as half a GCSE grade. However, 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, it would not seem like a good investment without further study. Whilst teacher quality is an important aspect of education, it may be more effective to recruit and retain effective teachers, rather than look for improvement based on financial reward.

Further information:

An Australian report from 2007 sets out the issues and evidence succinctly: http://www.dest.gov.au/sectors/school_education/publications_resources/profiles/research_on_performance_pay_for_teachers.htm .

Reducing class sizes

What is it?

Intuitively, it is appealing to reduce the number of pupils in a class to improve the quality of teaching and increase the level of personalisation or the amount of individual feedback a learner receives. As the size of a class or teaching group gets smaller, the range of approaches a teacher can employ increases.

How effective is it?

Overall the benefits are not particularly large or clear, until class size is reduced to under 20 or even below 15. There is little advantage in reducing classes from, say, 30 to 25. The issue is whether the teacher changes their teaching approach when working with a smaller class and whether, as a result, the pupils change their learning behaviours. Having 15 pupils in a class and teaching them in exactly the same way as a class of 30 will not make much difference. However there is evidence that, when it is successful, the benefits can be identified in behaviour and attitudes as well as on attainment, and that they persist for a number of years (from early primary school through to Key Stage 3). Evidence from both the USA and from the UK does not support the use of teaching assistants as an alternative. It appears to be important that a class teacher has responsibility for the learning of a class. Optimistically a school might expect a few months additional gain a year for pupils in smaller classes (an effect size of about 0.21). There is some evidence to support the additional benefit of professional development when class sizes are reduced to enable teachers to capitalise on the potential benefits by developing their teaching skills and approaches.

Impact summary: + 2 months (effect size 0.20)

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 higher attainment (when other factors are controlled for) and when class sizes have been deliberately reduced in more experimental evaluations.

Strength of the evidence: ☆ ☆ ☆

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 30 pupils with 50% of them receiving free school meals would only be allocated an extra £6,450 under the Pupil Premium in 2011/12; this would not be sufficient to appoint an additional teacher.

Cost summary: £££££

How applicable is it?

The strongest evidence comes from research into primary schools in the USA with younger children where the benefits appear to be sustained for 3 – 4 years.

Further information:

A good recent summary and analysis of the research evidence is available at: http://www.brookings.edu/papers/2011/0511_class_size_whitehurst_chingos.aspx .

School Uniforms

What is it?

Schools identify clothing considered appropriate for pupils to wear in school, usually including style and colour. There is a general belief in the UK that such an approach supports the development of a whole school ethos and therefore is supportive of discipline and motivation. It should be noted that in other cultures the opposite view prevails, and school uniforms are associated with regulation and the loss of individuality.

How effective is it?

There is no robust evidence that introducing a school uniform will 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.

Impact summary: ± 1 months (0.03 Eng/ -.06 Ma)

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.

Strength of the evidence: ☆

What are the costs?

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

Cost summary: £

How applicable is it?

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 seems likely not to be applicable in other settings.

Further information:

A good summary of the debate and evidence cam be found in an article on the Education World site at:

http://www.educationworld.com/a_admin/admin/admin130.shtml .

Sports participation

What is it?

Physically engaging in sports as a participant. This might be through organised after school activities or as an organised programme by a local sporting club or association. Sometimes the 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, or ICT or literacy or mathematics lessons.

How effective is it?

The overall impact on academic achievement tends to be low (an effect size around .02), 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 (an effect size of 0.8) when combined with a structured numeracy programme.

Impact summary: + 3 months (effect size 0.19)

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 considerable variation in impact, including some studies which show negative effects.

Strength of the evidence: ☆ ☆

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 £200/pa excluding clothing and equipment. These costs vary according to equipment and venue, sports like rowing and ice hockey tend to have an annual fee (about £50, plus monthly subscription (£30/month).

Cost summary: £££

How applicable is it?

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

Further information:

The Department for Culture, Media and Sport (DCMS) set up the Culture and Sport Evidence (CASE) programme was set up by to collect evidence about participation in culture and sport and their recent review is available at:
http://www.culture.gov.uk/images/research/CASE-systematic-review-July10.pdf .

Summer schools

What is it?

Summer schools are when students attend lessons or classes during their summer holidays, often as catch-up or enrichment lessons. Some summer 'schools' do not have an academic focus and concentrate on sports or other non-academic activities. These approaches are not usually evaluated for, or associated with, learning gains. 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 effect size of about 0.26), though usually more beneficial for higher attaining pupils and less effective for low-SES pupils. Programmes are usually more effective in mathematics, when they are specifically tailored to students needs, and when parents are involved. Other variables seem to make less difference, such as whether the teacher is one of the student's usual teachers.

Impact summary: + 3 months (effect size 0.26)

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 above, a crucial factor is whether the summer school has an academic focus.

Strength of the evidence: ☆ ☆

What are the costs?

The costs involved are employing teachers for the duration of the summer school, with associated venue and resource costs (books, photocopying etc). Residential courses are in the region of £300 per week per student.

Cost summary: £££

How applicable is it?

The impacts vary according to the focus of the summer school, but benefits have been identified in a range of subjects, particularly for secondary school pupils.

Further information:

There is a good summary of the research evidence and helpful advice on running effective summer schools from Child Trends, a non-profit, non-partisan research centre in the US which studies children at all stages of development: http://www.childtrends.org/Files/Child_Trends-2009_09_01_FS_WWSummerLearning.pdf .

Teaching Assistants

What is it?

A teaching assistant or classroom assistant (or sometimes called an educational assistant or paraprofessional) 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, such as literacy or numeracy. They are also often responsible for hearing children read, and helping teachers' preparation by photocopying, or sorting out equipment.

How effective is it?

Most studies have consistently found very small or no effects on attainment, though pupils' perceptions and attitudes may be more positively affected. There are also positive effects in terms of teacher morale and reduced stress of working with a teaching assistant. One clear implication from this is that if teaching assistants are used with the intention of improving the learning of pupils, they should not undertake the tasks they are routinely assigned. There is some evidence that there is greater impact when teaching assistants are given a particular pedagogical role or responsibility in specific curriculum interventions where the effect appears to be greater, particularly with training and support. Even here, however, comparisons with qualified teachers suggest they are consistently less effective (achieving about half the gains compared with qualified teachers).

Impact summary: ± 0 months (effect size 0.0)

How secure is the evidence?

There are a number of systematic reviews of the impact of support staff in schools, though no meta-analyses specifically looking at the impact of teaching assistants 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 low attaining pupils do less well with a teaching assistant.

Strength of the evidence: ☆ ☆

What are the costs?

The average teaching assistant salary is about £16,000 pa or about half of an average teaching salary (including headteachers and deputies).

Cost summary: ££££

How applicable is it?

Evidence suggests that impact is similar across literacy and mathematics and at both primary and secondary level.

Further information: A report on the role and impact of teaching assistants in the UK was commissioned by the Department for Education and Skills and undertaken by a team from the Institute of Education at London University: http://www.education.gov.uk/publications/eOrderingDownload/RR605.pdf .

Appendix 1: Estimates for additional income per school 2011/12

(based on an allocation of £430 for pupils receiving free school meals)

	%fsm								
No on roll	10%	20%	30%	40%	50%	60%	70%	80%	90%
50	£2,150	£4,300	£6,450	£8,600	£10,750	£12,900	£15,050	£17,200	£19,350
100	£4,300	£8,600	£12,900	£17,200	£21,500	£25,800	£30,100	£34,400	£38,700
150	£6,450	£12,900	£19,350	£25,800	£32,250	£38,700	£45,150	£51,600	£58,050
200	£8,600	£17,200	£25,800	£34,400	£43,000	£51,600	£60,200	£68,800	£77,400
250	£10,750	£21,500	£32,250	£43,000	£53,750	£64,500	£75,250	£86,000	£96,750
300	£12,900	£25,800	£38,700	£51,600	£64,500	£77,400	£90,300	£103,200	£116,100
350	£15,050	£30,100	£45,150	£60,200	£75,250	£90,300	£105,350	£120,400	£135,450
400	£17,200	£34,400	£51,600	£68,800	£86,000	£103,200	£120,400	£137,600	£154,800
450	£19,350	£38,700	£58,050	£77,400	£96,750	£116,100	£135,450	£154,800	£174,150
500	£21,500	£43,000	£64,500	£86,000	£107,500	£129,000	£150,500	£172,000	£193,500
550	£23,650	£47,300	£70,950	£94,600	£118,250	£141,900	£165,550	£189,200	£212,850
600	£25,800	£51,600	£77,400	£103,200	£129,000	£154,800	£180,600	£206,400	£232,200
650	£27,950	£55,900	£83,850	£111,800	£139,750	£167,700	£195,650	£223,600	£251,550
700	£30,100	£60,200	£90,300	£120,400	£150,500	£180,600	£210,700	£240,800	£270,900
750	£32,250	£64,500	£96,750	£129,000	£161,250	£193,500	£225,750	£258,000	£290,250
800	£34,400	£68,800	£103,200	£137,600	£172,000	£206,400	£240,800	£275,200	£309,600
850	£36,550	£73,100	£109,650	£146,200	£182,750	£219,300	£255,850	£292,400	£328,950
900	£38,700	£77,400	£116,100	£154,800	£193,500	£232,200	£270,900	£309,600	£348,300
950	£40,850	£81,700	£122,550	£163,400	£204,250	£245,100	£285,950	£326,800	£367,650
1000	£43,000	£86,000	£129,000	£172,000	£215,000	£258,000	£301,000	£344,000	£387,000
1100	£47,300	£94,600	£141,900	£189,200	£236,500	£283,800	£331,100	£378,400	£425,700
1200	£51,600	£103,200	£154,800	£206,400	£258,000	£309,600	£361,200	£412,800	£464,400
1300	£55,900	£111,800	£167,700	£223,600	£279,500	£335,400	£391,300	£447,200	£503,100
1400	£60,200	£120,400	£180,600	£240,800	£301,000	£361,200	£421,400	£481,600	£541,800
1500	£64,500	£129,000	£193,500	£258,000	£322,500	£387,000	£451,500	£516,000	£580,500
1600	£68,800	£137,600	£206,400	£275,200	£344,000	£412,800	£481,600	£550,400	£619,200
1700	£73,100	£146,200	£219,300	£292,400	£365,500	£438,600	£511,700	£584,800	£657,900
1800	£77,400	£154,800	£232,200	£309,600	£387,000	£464,400	£541,800	£619,200	£696,600
1900	£81,700	£163,400	£245,100	£326,800	£408,500	£490,200	£571,900	£653,600	£735,300
2000	£86,000	£172,000	£258,000	£344,000	£430,000	£516,000	£602,000	£688,000	£774,000

Appendix 2: Estimates for additional income per school 2014/15

(based on an allocation of £1750 for pupils receiving free school meals)

	%fsm								
No on roll	10%	20%	30%	40%	50%	60%	70%	80%	90%
50	£8,750	£17,500	£26,250	£35,000	£43,750	£52,500	£61,250	£70,000	£78,750
100	£17,500	£35,000	£52,500	£70,000	£87,500	£105,000	£122,500	£140,000	£157,500
150	£26,250	£52,500	£78,750	£105,000	£131,250	£157,500	£183,750	£210,000	£236,250
200	£35,000	£70,000	£105,000	£140,000	£175,000	£210,000	£245,000	£280,000	£315,000
250	£43,750	£87,500	£131,250	£175,000	£218,750	£262,500	£306,250	£350,000	£393,750
300	£52,500	£105,000	£157,500	£210,000	£262,500	£315,000	£367,500	£420,000	£472,500
350	£61,250	£122,500	£183,750	£245,000	£306,250	£367,500	£428,750	£490,000	£551,250
400	£70,000	£140,000	£210,000	£280,000	£350,000	£420,000	£490,000	£560,000	£630,000
450	£78,750	£157,500	£236,250	£315,000	£393,750	£472,500	£551,250	£630,000	£708,750
500	£87,500	£175,000	£262,500	£350,000	£437,500	£525,000	£612,500	£700,000	£787,500
550	£96,250	£192,500	£288,750	£385,000	£481,250	£577,500	£673,750	£770,000	£866,250
600	£105,000	£210,000	£315,000	£420,000	£525,000	£630,000	£735,000	£840,000	£945,000
650	£113,750	£227,500	£341,250	£455,000	£568,750	£682,500	£796,250	£910,000	£1,023,750
700	£122,500	£245,000	£367,500	£490,000	£612,500	£735,000	£857,500	£980,000	£1,102,500
750	£131,250	£262,500	£393,750	£525,000	£656,250	£787,500	£918,750	£1,050,000	£1,181,250
800	£140,000	£280,000	£420,000	£560,000	£700,000	£840,000	£980,000	£1,120,000	£1,260,000
850	£148,750	£297,500	£446,250	£595,000	£743,750	£892,500	£1,041,250	£1,190,000	£1,338,750
900	£157,500	£315,000	£472,500	£630,000	£787,500	£945,000	£1,102,500	£1,260,000	£1,417,500
950	£166,250	£332,500	£498,750	£665,000	£831,250	£997,500	£1,163,750	£1,330,000	£1,496,250
1000	£175,000	£350,000	£525,000	£700,000	£875,000	£1,050,000	£1,225,000	£1,400,000	£1,575,000
1100	£192,500	£385,000	£577,500	£770,000	£962,500	£1,155,000	£1,347,500	£1,540,000	£1,732,500
1200	£210,000	£420,000	£630,000	£840,000	£1,050,000	£1,260,000	£1,470,000	£1,680,000	£1,890,000
1300	£227,500	£455,000	£682,500	£910,000	£1,137,500	£1,365,000	£1,592,500	£1,820,000	£2,047,500
1400	£245,000	£490,000	£735,000	£980,000	£1,225,000	£1,470,000	£1,715,000	£1,960,000	£2,205,000
1500	£262,500	£525,000	£787,500	£1,050,000	£1,312,500	£1,575,000	£1,837,500	£2,100,000	£2,362,500
1600	£280,000	£560,000	£840,000	£1,120,000	£1,400,000	£1,680,000	£1,960,000	£2,240,000	£2,520,000
1700	£297,500	£595,000	£892,500	£1,190,000	£1,487,500	£1,785,000	£2,082,500	£2,380,000	£2,677,500
1800	£315,000	£630,000	£945,000	£1,260,000	£1,575,000	£1,890,000	£2,205,000	£2,520,000	£2,835,000
1900	£332,500	£665,000	£997,500	£1,330,000	£1,662,500	£1,995,000	£2,327,500	£2,660,000	£2,992,500
2000	£350,000	£700,000	£1,050,000	£1,400,000	£1,750,000	£2,100,000	£2,450,000	£2,800,000	£3,150,000

Appendix 3: Criteria for inclusion and evaluation

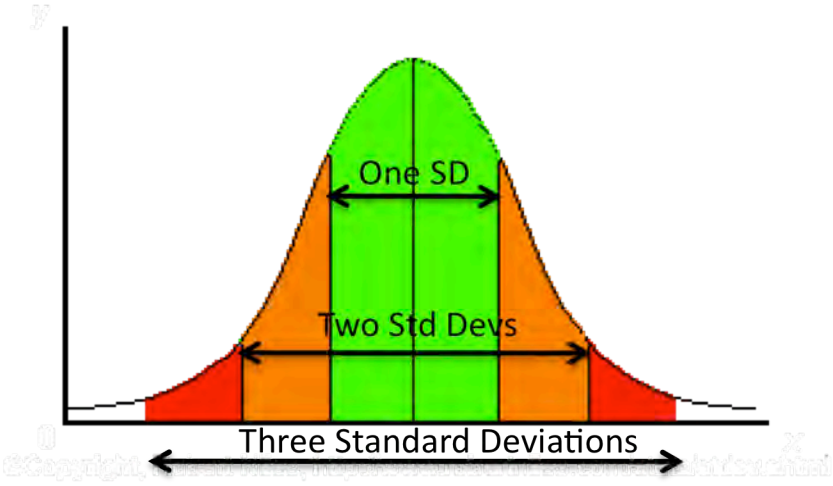
Approach

A range of approaches were selected for analysis and inclusion in the toolkit, based on those commonly mentioned in connection with the policy (such as reducing class sizes or one-to-one tutoring), suggestions identified by schools in informal discussions about how they might spend the Pupil Premium and a number of research-based approaches with a strong evidence of effectiveness (such as feedback and meta-cognitive approaches.

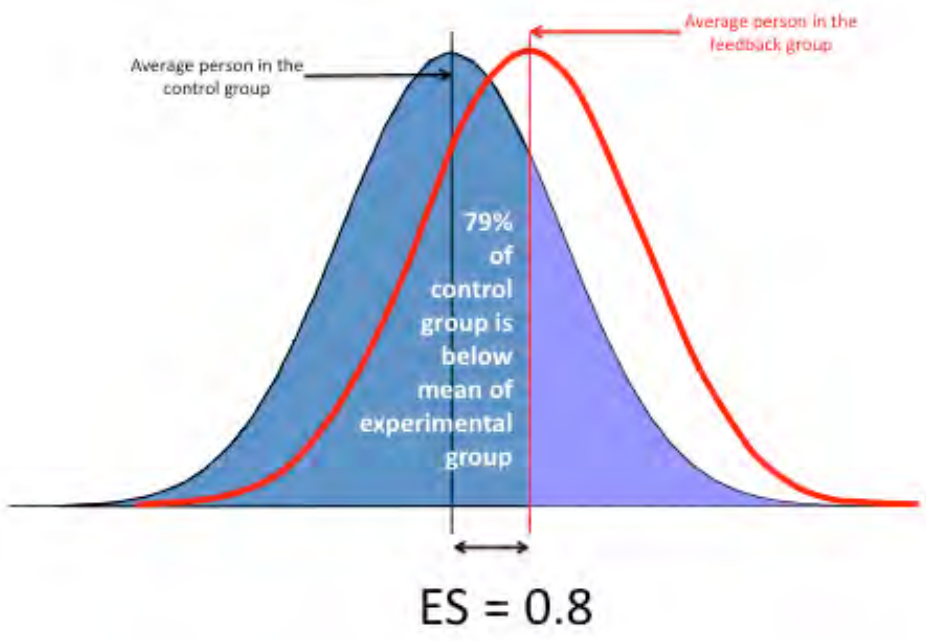
Potential gain

The comparative data is based (where available) on systematic reviews of research and in particular quantitative syntheses of data, such as meta-analyses of experimental studies. These are reviews which have specifically set out to evaluate the impact quantitatively of different approaches on learning using test or examination scores. These therefore offer the best idea of the relative benefit of the different approaches across similar populations of learners. Other types of reviews were included to provide information about applicability and about cost. To be included in the analysis an approach needed to have some quantifiable evidence base for comparison.

> *Effectiveness* – estimated as additional progress in terms of months of schooling you might expect pupils to benefit from as a result of a particular approach being used in school. Average pupil progress over a year is used as a benchmark. This is modelled using national test data from upper Key Stage 2 and Key Stage 3. The progress that an average pupil in a year group of 100 students is equivalent to them moving up from 50th place to 16th place, if all the other students had not made any progress. For national tests this is approximately one standard deviation or one year of progress. From this measure we can estimate the additional progress you might get by adopting different approaches which have been researched using test results by using this common measure or 'effect size' to estimate the additional progress which might be achieved.



Effect size – the extent of the difference between an intervention group and a control group in standardised form. Statistically it is the difference between two groups relative to the distribution of the scores, or the difference between them in standard deviation units. This produces a measure where a comparison can be made across a range of different approaches (based on average relative impact on the students tested). It can also be 'translated' into a measure of improvement, so one GCSE grade improvement is about an effect size of between 0.5 and 0.7, depending on subject.



If a feedback intervention has an effect size of 0.8, it means that, for two classes of pupils which were equivalent before an intervention, afterwards the class receiving the feedback intervention would be outperforming the control class. 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 group. An effect size of 1.6 would be the equivalent of moving the average pupil in an intervention class of 25 pupils to the top of the control class.

Where there were several meta-analyses available for analysis a judgement was made about their comparability, and an indicative value selected as representative. This was related to how realistic the research was or the 'ecological validity', how recent the studies were, and how appropriate outcome measures. So meta-analyses of studies undertaken in schools where students were taught by teachers and where the assessment was made with standardised or national tests undertaken in the last 10 years were given higher weight. Effect size measures are included in the summaries of each approach to provide a more precise figure for comparison. Further details are provided in the technical report in the section on effect sizes.

We do not know to what extent the effects might be additive (i.e. could you get extra benefit by combining different approaches) as there is little evidence in the research literature about this. What information there is suggests it is more likely that there will be a maximum effect from the most beneficial approach and that this is more likely to be reduced if combined with a less effective approach. So a meta-cognitive approach with ICT is more likely will be somewhere between the average metacognitive approach and average ICT approach.

Effect sizes have been estimated in terms of months of progress. These are broad estimates based on national test progress of an effect size of 1.0 over a school year and should not be thought of as guarantees. In the research the analysis is based on there is often considerable variation in impact between different studies and on different pupils in those studies. What is of interest is that there is some consistency in the extent of the effect in different areas of research between the different approaches. So, for example, studies which have looked at approaches improving the quality of feedback in classrooms, for example, tend to have much greater impact on tested learning outcomes than those which have looked at another approach such as matching learning styles. The toolkit is an attempt to look at these patterns and to identify some areas which are likely to be more productive for schools to investigate and to evaluate for themselves.

Months progress	Effect Size from ...	... to	Description
0	-0.07	0.01	Very low or no
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

Approximate costings have been made where possible based on the £430 allocation for 2011/2. Rough equivalents are given to help understand what the Pupil Premium might support. For example at least 60 pupils receiving the pupil premium are needed to employ an additional teacher (assuming Main Pay Scale 3 (£25,168) or Outer London MPS1 (£25,117) in 2010-11. If the Pupil Premium increases to £1750 by 2014/5, this will be reduced to about 15 pupils. The scale used in the costing assumptions is as follows:

£	Very low: up to about £2000 per year per class of 30 pupils, or less than £70 per pupil per year. This is about equivalent to 7 boxes (35 reams) of photocopying paper.
££	Low: £2000 - £5000 per year per class of 30 pupils, or up to about £170 per pupil per year. This would pay for significant CPD for a teacher with in class support and/or replacement teaching. It would also be equivalent to about 3 weeks after school provision per pupil (at about £10/head/day).
£££	Moderate: over £5k to £15k per year per class of 30 pupils, or up to about £500 per pupil per year. This represents the upper limit for the 2011/12 allocation.
££££	High: over £15k up to £30k per year per class of 30 pupils, or up to £1000 per pupil. This would not currently be achievable with the 2011/12 allocation but would represent about 60% of the allocation by 2014.
£££££	Very high: over £30k per year per class of 30 pupils. By 2014/5, a class of 30 pupils where 70% of them are eligible for free school meals would increase a school's budget by about £36,750 per year.

Applicability

This is a summary of where the evidence is drawn from or where the evidence of impact is greatest in terms of impact across primary and secondary schools. Also included is an indication of what the evidence indicates in terms of curriculum subjects. The majority of experimental studies tend to look at impact on test scores in English or literacy, mathematics and science. These include a range of kinds of tests and assessments such as those designed for the research project, teacher designed tests, existing school tests and exams (usually in specific curriculum subjects), and standardised tests. Where are range of measures were available priority was given to existing school measures and standardised tests. Also included in this analysis is a judgement about the challenge associated with adapting or implementing the approach in schools and evidence from the UK is weighted more strongly as is more recent research, particularly in areas where the context changes, such as ICT..

Evidence estimate

This is based on the availability of the evidence (the number of systematic reviews or meta-analyses and the quantity of primary studies which they synthesise) as well as the quality of the primary evidence (from a methodological point of view) combined with the extent of the impact (effect size) and the reliability or consistency of this impact across.

☆	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.
☆ ☆	At least one meta-analysis or systematic review with quantitative evidence of impact on attainment or cognitive or curriculum outcome measures.
☆ ☆ ☆	Two or more rigorous meta-analyses of experimental studies of school age students with cognitive or curriculum outcome measures.
☆ ☆ ☆ ☆	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.
☆ ☆ ☆ ☆ ☆	Consistent[3] high quality evidence from at least five robust[4] and recent[5] meta-analyses where the majority of the included studies have good ecological validity[6] and where the outcome measures include curriculum measures or standardised tests in school subject areas.

[3] Pooled effect sizes are reasonably similar or, where different, similar patterns of effects are found for comparable moderator variables associated with the approach, producing a consistent and coherent picture.
[4] Meta-analysis reported with confidence intervals and heterogeneity. Some checks for bias investigated (e.g. study quality and/or and some moderator exploration.
[5] Within the last 10 years.
[6] Studies conducted in schools with more than one teacher or class.

Toolkit to Improve Learning
Technical Appendices

Overview

The aim of these appendices is to set out some of the assumptions and methods used in the synthesis of effect sizes in the '*Toolkit of Strategies to Improve Learning: Summary for Schools Spending the Pupil Premium'*. The primary aim is to provide schools with evidence from education research which will help them to make informed decisions about spending the Pupil Premium 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. This set of appendices sets out the rationale and sources of evidence for these decisions.

There are of course some limitations and caveats. The quality of the evidence 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 fields that have been summarised for the toolkit. There are examples within each area where interventions have been successful in improving attainment and have been unsuccessful. The most successful approaches have had their failures and the least successful their triumphs. This summarisation, which aims only to provide an overall 'best bet', therefore masks these differences. What we are saying is that the existing evidence so far suggests that some areas are likely to be more productive of success than others and that meta-analysis provides the best evidence for this. What we are not saying is that unsuccessful approaches can *never* work nor that 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 are encouraging for teachers. It 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.

Appendix 1: Resources and pupil learning

It is difficult to establish a clear link between educational expenditure and pupils' learning. Most of the studies have been undertaken at the educational system level, such as countries, states or local authorities, where the relationship between allocation of resources and differences in schools and 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-analysis 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).

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 £2000 *per teacher* per year. A recent evaluation of Every Child a Reader (Tanner et al. 2011) estimates costs of £3100 in the first year and £2600 per year subsequently *per child* with an average reading gain of 13% (non-significant, p 142) (estimated at an effect size of about 0.14: Glass, McGaw & Smith, 1981 p 136). Even with the 2014-15 allocation of about £1,750 per child receiving free school meals, this level of spending will not be possible.

So overall, although there is no clear evidence of a causal link between general additional spending and learning, we interpret this to mean that it is difficult to spend additional resource effectively. It may be hard to find a causal connection, but there must be some areas which offer better prospects than others, and this is what this toolkit seeks to provide. We also think that the evidence shows that if schools want to use the Pupil Premium to benefit disadvantaged learners they should not

assume that the increased allocation alone will improve learning, but they will need to decide specifically and deliberately how it should be spent, and then evaluate the effectiveness of this for themselves. The existing research indicates that this is challenging but achievable task.

References

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

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

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.

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.

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.

Vignoles, A., Levacic, R., Walker, J., Machin, S. & Reynolds, D. (2000) *The Relationship Between Resource Allocation and Pupil Attainment: A Review* London: Department for Education and Employment.

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.

Appendix 2: Cost effectiveness estimates

Approximate costings have been made where possible based on the initial £430 allocation. Rough equivalents are given to help understand what the Pupil Premium might enable during the first year of spending. For example, at least 60 pupils receiving the pupil premium will be needed to employ an additional teacher in 2010-11 (assuming Main Pay Scale 3 (£25,168) or Outer London MPS1 (£25,117). If the Pupil Premium increases to £1,750, this will be reduced to about 15 pupils.

The scale used in the costing assumptions is as follows:

£	Up to about £2000 per year per class of 30 pupils, or less than £70 per pupil per year. This is about equivalent to 7 boxes (35 reams) of photocopying paper.
££	£2000 - £5000 per year per class of 30 pupils, or up to about £170 per pupil per year. This would pay for significant CPD for a teacher with in class support and/or replacement teaching. It would also be equivalent to about 3 weeks after school provision per pupil (at about £10/head/day).
£££	Over £5k to £15k per year per class of 30 pupils, or up to about £500 per pupil per year. This represents the upper limit for the 2011 allocation.
££££	Over £15k up to £30k per year per class of 30 pupils, or up to £1000 per pupil. This would not currently be achievable with the 2011 allocation but would represent about 60% of the allocation by 2014.
£££££	Over £30k per year per class of 30 pupils. By 2014, a class of 30 pupils where 70% of them are eligible for free school meals would increase a school's budget by £36,750 per year.

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

Expenditure	**Rate**	**Cost estimate**
Teacher	£25 - £30k/ per annum (Scale point 3 Eng & Wales – Inner London Scale Point 3)	£27,500 pa
Teaching Assistant	£16-18k per annum	£17,000 pa
Supply cover	£150 - £200/day	£175/day
Computer	Total cost of ownership estimated at £3,000	£600 pa
CPD day course	£60 - £500/ day	£200/day
CPD programme	Training, support and cover for a 5 day programme with classroom development	£2000 pa
Paper	£2 per ream (500 sheets)	£240 pa per class

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.

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

What is it?

Effect size 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.

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?

So, 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 12^{th} or 13^{th} in their class) would now score about the same as the person ranked 6^{th} 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 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.

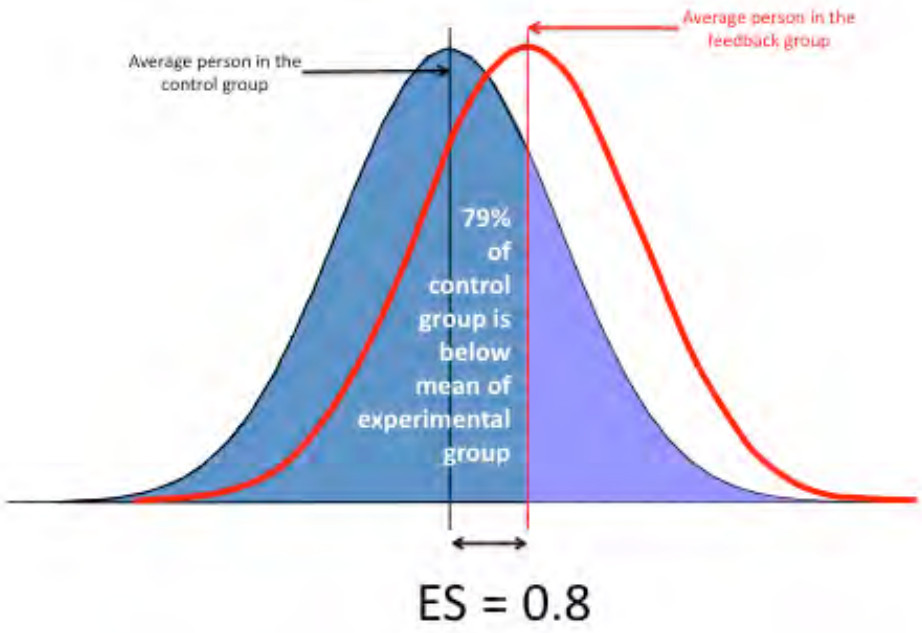


Figure 1: An effect size of 0.8

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 et al. (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, p 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. 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 p 103). However, we should note that the correspondence of one standard deviation to one year’s progress can vary considerably for different ages and types of test.

Months progress	Effect Size from …	… to	Description
0	-0.01	0.01	Very low or no
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

It is also the case that effect size difference reduces with age. Bloom et al. (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[1] from National Curriculum tests (DfES, 2004) indicates annual gains representing and 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 SD 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 month's 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.

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 optimize 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 et al. (1980) calls '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

[1] 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.

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' Δ. 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' g

Hedges' 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' 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' *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. Another key issue is which standard deviation is chosen (Bloom et al. 2008) as this primarily determines the comparability of the effect size. 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, p 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 (p 503), 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.

References

Bloom HS Hill,CJ Black AR and Lipsey, MW (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

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).

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

Cronbach LJ Ambron, SR, Dornbusch, SM, Hess, RO, Hornik, RC., Phillips, DC., 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).

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

Glass GV, McGaw, B & Smith, ML (1981) *Meta-analysis in social research* Beverly Hills, Ca: Sage.

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

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 CJ Bloom HS Black AR & Lipsey MW (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).

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.

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.

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).

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.

Appendix 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 an answer to question 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.

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. 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 where 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. 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 ICT 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 help 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 metacognitive 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 metacognitive 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. It must be acknowledged, however, that this kind of super-synthesis or meta-meta-analysis remains distinctly *controversial* as a research approach.

Search and inclusion criteria

The main source of studies for the toolkit was a database of meta-analyses of educational interventions developed for an ESRC Researcher Development Initiative[2]. Additionally a search was undertaken for systematic reviews with quantitative data (where effect sizes were reported but not pooled) and meta-analyses (where effect sizes are combined to provide a pooled estimated of effect) of intervention research in education using 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). References and sources in existing super-syntheses (e.g. Sipe & Curlette, 1997; Marzano, 1998; Hattie, 2009) were reviewed and obtained where possible. Other reviews and studies were consulted in each area to provide additional contextual information.

A number of areas were specifically included at the request of teachers who were consulted at different stages in the development of the toolkit. Thanks in particular go to ARK and teachers from the TeachFirst Future Leaders programme and a group of Hammersmith and Ealing deputy headteachers as well as 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. Assessment for Learning, 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 have a lower overall quality assessment in terms of the overall warrant from the research evidence.

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

Weight of evidence and quality assessment

The weight of evidence in each field was assessed according to the criteria 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 seemed to be important criteria.

Quality assessment categories

☆	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.
☆ ☆	At least one meta-analysis or systematic review with quantitative evidence of impact on attainment or cognitive or curriculum outcome measures.
☆ ☆ ☆	Two or more rigorous systematic reviews or meta-analyses of experimental studies of school age students with cognitive or curriculum outcome measures.
☆ ☆ ☆ ☆	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.
☆ ☆ ☆ ☆ ☆	Consistent[3] high quality evidence from at least five robust[4] and recent[5] meta-analyses where the majority of the included studies have good ecological validity[6] and where the outcome measures include curriculum measures or standardised tests in school subject areas.

References

Bloom, HS Hill C.J Black AR & Lipsey, MW (2008) 'Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions', *Journal of Research on Educational Effectiveness*, 1: 4, 289-328.

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

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

Hattie JA (1992) Measuring the effects of schooling. *Australian Journal of Education*, 36, 5–13.

Hattie JA (2008) *Visible Learning* London: Routledge.

Hill CJ Bloom HS Black AR & Lipsey MW (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).

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

[3] Pooled effect sizes are reasonably similar or, where different, similar patterns of effects are found for comparable moderator variables associated with the approach, producing a consistent and coherent picture.

[4] Meta-analysis reported with confidence intervals and heterogeneity. Some checks for bias investigated (e.g. study quality and/or and some moderator exploration).

[5] Within the last 10 years.

[6] Studies conducted in schools with more than one teacher or class.

Marzano, RJ (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) .
Sipe T & Curlette WL (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.
Smith NL (1982) Evaluative Applications of Meta- and Mega- Analysis *American Journal of Evaluation* 3(4) pp 43.

Appendix 5: Notes on summaries and additional references

This section contains information additional to that presented in the 'Summary for Schools', and in particular details the sources used for the overview of each area, with any additional information used to supplement the quantitative analysis. Bibliographic details for the meta-analyses and other sources used for the quantitative estimates can be found in Appendix 7.

Ability grouping

As Kulik (1992) observed, the key distinction in ability grouping is between approaches where all ability groups follow the same curriculum, between approaches where groups follow different curricula adjusted to their ability and between approaches which make curricular and other adjustments for particular groups such as the particular needs of highly talented or disadvantaged learners. Overall there is substantial and robust evidence in this area, with a reasonably consistent picture of effects, particularly on low attaining pupils, that grouping by ability is detrimental to these learners' progress (Ireson et al. 1999) and perceptions of themselves as learners (Ireson et al. 2001). One of the first meta-analyses in this field (Kulik & Kulik, 1982) focussed on secondary schools and found that studies where high-attaining students received enriched instruction produced especially clear positive effects while studies of average and below average students produced near-zero effects; the pattern has changed little since then. Boaler's work (2008) shows that it is possible to achieve high attainment even in subjects like mathematics in mixed ability groups in secondary schools.

Ability grouping summary	**± 1 month** ☆☆☆
Kulik & Kukik 1982 (secondary)	0.1
Kulik & Kulik 1984 (elementary)	0.1
Lou et al 1996 (on low attainers)	-0.12
Slavin 1990 (on low attainers)	-0.06

Additional references

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

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 JA (1992) *An Analysis of the Research on Ability Grouping: Historical and Contemporary Perspectives* The National Research Center On The Gifted And Talented.

After school programmes

The evidence in this area is not substantial or particularly robust (Fashola 1998), what evidence there is suggests that the impact of after school programmes is

variable, but positive. Participants in after-school programmes score higher on measures of academic achievement. The average effect size for the benefit for students in after-school programmes over comparison students was 0.21 in reading and 0.16 in mathematics in Scott-Little et al.'s (2002) study. Lauer et al. (2006) found small but statistically significant positive effects of such programmes on reading and mathematics achievement for at-risk students (overall effect size on reading: 0.13 and mathematics: 0.17) and larger positive effect sizes for programmes with specific characteristics such as tutoring in reading. Attending a formal after-school programme where low-income children spend more time in academic and enrichment activities with peers and adults was also correlated in Posner and Vandell's (1994) study with their academic and conduct grades, peer relations and emotional adjustment. Similarly, other studies have shown that participation in school-based, after-school programmes is associated with behaviour that could help youth stay out of trouble and with positive effects on school attitudes and behaviours (Grossman et al., 2002; Woodland, 2008). A recent meta-analysis of after-school programmes (Durlak & Weissberg, 2007) that seek to enhance the personal and social development of children and adolescents indicated that there was improvement in three general areas: feelings and attitudes, indicators of behavioural adjustment, and school performance. More specifically, significant increases occurred in the young people's self-perceptions and views of school, their positive social behaviours, and in their school grades and level of academic achievement. Among programmes intended to increase academic achievement, those that provide greater structure, a stronger link to the school curriculum, with well-qualified and well-trained staff, and opportunities for one-to-one tutoring seem particularly promising (Fashola, 1998). Programmes may not be equally effective with all students however. The emerging pattern seems to suggest that older primary pupils do not show the same gains as the younger children enrolled in after-school programmes. In addition, at-risk children may benefit more from participating in such programmes (Scott-Little et al., 2002), but may be harder to retain or keep engaged (Grossman et al. 2002).

After school programmes summary	**+ 2 months** ☆☆
Durlak & Weissberg 2007	0.16
Fashola 1998	*NPE*
Lauer, Akiba & Wilkerson 2006	0.16
Scott-Little et al 2002	*NPE*

Additional references

Durlak JA (2007) *The Impact of After-School Programs that Promote Personal and Social Skills*. Chicago: CASEL.

Grossman JB Price ML Fellerath V Jucovy LZ Kotloff LJ Raley R & Walker KE (2002) *Multiple Choices After School: Findings from the Extended-Service Schools Initiative.* Philadelphia: Public/Private Ventures.

Posner, JK & Vandell, DL (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, MS & Jurs, SG (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 MH (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.

Arts participation

The challenge in this area is the breadth of areas of activity and intervention, from music (Standley, 2008) to creative and performing arts (Lewis, 2004). Overall the evidence is not conclusive (Winner & Cooper, 2000) with many interventions lacking robust evaluation (Newman et al. 2010) and a range of varying effects identified. There is some evidence of positive benefits, however it is hard to identify factors clearly associated with this.

Arts participation summary	+ 1 month	☆☆☆
Lewis 2004 (performing arts on academic outcomes)	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	

Additional references

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, pp 11-75.

Assessment for learning

One of the challenges is clearly defining 'Assessment for Learning' (Black & Wiliam, 2009; Bennett, 2011). In Black and Wiliam's (1998) early work it is equated with effective formative feedback, drawing on a tradition going back to Bloom et al. (1971), so one might expect effect sizes to be more similar to feedback studies or approaches like mastery learning (0.52: Kulik, Kulik & Bangert Drowns, 1990). Preliminary research in schools (summarised in Wiliam, 2002) indicated benefits were achievable, if not as large as found in experimental studies. Smith and Gorard (2005) show what can go wrong when schools misunderstand or misinterpret the intentions behind the practices associated with a policy version of a research-based intervention (see also Black and Wiliam (2009) and the section on 'Effective Feedback' below).

Assessment for learning summary	**+ 3 months**	☆
[No meta-analyses of school interventions]		
Wiliam 2002 (KMOFAP synthesis)	0.32	

Additional references

Bennett RE (2011) Formative assessment: a critical review, *Assessment in Education: Principles, Policy & Practice,* 18: 1, 5 — 25

Black PJ & Wiliam D (1998) Assessment and classroom learning, *Assessment in Education,* 5, pp. 7–73.

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.
Bloom BS Hastings JT & Madaus GF (Eds) (1971) Handbook on the Formative and Summative Evaluation of Student Learning New York: McGraw-Hill.
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.
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.

Block scheduling and timetabling

The influence of altering timetables to create longer blocks of time or a more intensive series of lessons at secondary level has been investigated, though the evidence is not particularly robust (Dickson et al. 2010) and in particular the impact on teachers' practices has not been studied systematically (Zepeda & Mayers, 2006). The effect sizes identified tend to be low or even negative, which suggests schools should be cautious about making changes without a clear idea of how they will use the changed pattern of lessons effectively (Gruber & Onwuegbuzie, 2001). There is some evidence that such changes are more successful in science, perhaps as longer lessons enable more focused or more complete investigative work to be undertaken. Veal and Flinders (2001) found that block scheduling was perceived by teachers to provide increased variety of instruction and an overall improvement in classroom climate through improved student-teacher relationships. Our interpretation is that timetabling and lesson length changes will not make a difference unless teachers (or pupils) change aspects of teaching and learning interactions to take advantage of the differences.

Block scheduling and timetabling	**± 1 month** ☆☆
Dickson et al. 2010 (achievement)	0.11
Dickson et al. 2010 (mathematics)	-0.02
Dickson et al. 2010 (science)	0.20
Lewis et al. 2005 (mathematics)	-0.10
Lewis et al. 2005 (English)	-0.17
Lewis et al. 2005 (science)	-0.12

Additional references

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.

Early intervention

There is consistent evidence that early intervention is beneficial for children's learning with typical effect sizes around 0.35 to 0.52 in meta-analyses. There is some evidence that these programmes need to be whole day (rather than half-day) and of longer duration (up to a year) rather than for shorter periods of time. The impact tends to wear off over time, however (Lewis & Vosburgh, 1988; Giliam and Zigler, 2000); though such intervention tends to have a more durable effect on attitudes to school than measures of attainment (Nelson et al. 2003). Some studies have also investigated interventions designed to improve the home environment (reviewed by Bakerman-Kranenburg et al., 2005) which suggest impact is harder to achieve with younger parents and with poorer families. Their findings also suggest that the immediate impact of effective interventions is associated with shorter programmes which were home-based. By contrast Campbell and Ramey (1994) found greater impact associated with children's learning in school in the longer term was associated with more lasting programmes (from early infancy to age 8) which influenced both home and school environments; this is consistent with Nelson et al. (2003). Lewis and Vosburgh (1988) found that more durable effects were associated with greater parental involvement. Some caution is needed, however, in generalising these findings about early intervention. In the UK the EPPE study suggested an effect of about 0.18 for pre-school attendance with performance in Reception classes, and a recent evaluation of Sure Start Local Programmes (NESS Team, 2010) did not find any differences in Foundation Profile scores for Sure Start children at the start of school, though quality of provision was linked with better language and communication outcomes (Melhuish et al. 2011).

Early intervention	**+ 6 months** ☆☆☆☆
Anderson et al. 2003	0.35
Gilliam & Zigler 2000	NPE
LaParo & Pianta 2000	0.51
Lewis & Vosburgh 1988	0.41
Nelson et al. 2003	0.52

Additional references

Bakermans-Kranenburg MJ van IJzendoorn MH Bradley RH (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.

Campbell FA & Ramey CT (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.

Melhuish E Belsky J Leyland AH Barnes J et al. (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.

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.

Sylva, K. and MacPherson, K. (2002), *Parents and Parenting in the Early Years: Research evidence*. Oxford: Department of Educational Studies, University of Oxford.

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.

Effective feedback

The challenge in this area is to relate the feedback research literature with classroom practice and effective pedagogical approaches and techniques. Many of the studies in Kluger and de Nisi (1996), for example, are theoretically driven studies where the implications for classroom practice are unclear. Both Black and Wiliam (1998) and Hattie and Timplerley (2007) have summarised the implications for schools, but with slightly different emphases. Hattie and Timperley (2007) suggest that feedback should focus on challenging tasks or goals (rather than easy ones); that is even more important for teachers to give feedback about what is *right* rather than what is wrong. In addition feedback should be as specific as possible and, ideally, compare what students are doing right now with what they have done wrong before; and finally that it should encourage students, and not threaten their self-esteem. Black and Wiliam (1998) emphasise the use of feedback to close the gap on current performance relative to a desired goal or outcome, and highlight the importance of the student in identifying the gap and acting on the information (see also 'Assessment for Learning').

Effective feedback	**+ 9 months** ☆☆☆
Fuchs and Fuchs 1985	0.72
Kluger & De Nisi, 1996	0.41
Lysakowski & Walberg 1982	0.97
Tenebaum & Goldring 1989	0.72
Walberg 1982	0.81

Additional references

Black, PJ & Wiliam, D (1998) Assessment and classroom learning, *Assessment in Education,* 5, pp. 7–73.

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

Homework

A clear distinction needs to be made between correlational studies of homework and intervention studies. More affluent families are more likely send their children to schools which set regular homework, and these will be schools where students complete their homework and are likely to be more successful schools. The part that homework plays in this success is much less obvious. Intervention studies where homework is used as a means to improve learning outcomes are therefore rather different from associational studies and tend to report higher effect sizes (Cooper et al. 2006). There are also clear differences between primary and secondary schools with the evidence indicating that homework is less effective for younger children (Paschal et al., 1984; Cooper et al. 2006). Farrow, Tymms and Henderson's (1999)

correlational analysis relating to homework in the final year of primary school suggests that highest test scores were achieved by pupils who reported doing homework 'once a month'. Homework reported more frequently than this was generally associated with lower attainment. Multilevel models that controlled for a range of important variables did not lend support to a 'more is better' view of homework for primary school pupils. Denvir et al. 1999 also found no association between teachers' reports of frequency of homework and mathematics learning at primary school level. Overall it appears that the quality of homework is more important than the quantity. At secondary school level factors associated with increased learning were receiving feedback on homework and effective integration with teaching in lessons.

Homework summary	**+ 5 months** ☆☆☆
Cooper, Robinson & Patal 2006	0.60
Paschal, Weinsten & Walberg 1984	0.36

Additional references

Denvir H Rhodes V Brown M Askew M Wiliam D & Ranson E (1999) An Investigation On The Effect Of Homework On Pupil Gains In An Assessment Of Numeracy In The First Year Of The Leverhulme Numeracy Research Programme in L. Bills (Ed) *Proceedings of the British Society for Research into Learning Mathematics* 19.3.

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

ICT

There is extensive research evidence of the impact of different technologies. It is relatively consistent and tends to show moderate benefits for technology use (e.g. Tamim et al., 2011). However, due to the increasing pace of technological change, it is usually about yesterday's technology rather than today's and certainly makes it difficult for schools to know what to buy for tomorrow.

The challenge with digital technologies is to tease apart the relationship between different technologies and different teaching approaches and contexts (Crook et al. 2010). Whilst it is unlikely that particular technologies bring about changes in learning directly, different technologies may be more likely to support or enable changes to take place 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. The question should perhaps rather be where is there evidence that technology *can* be used effectively and *how* has it been used to support learning (Higgins, 2003). The other challenge is to evaluate the range of technologies in relation to the range of ways that they can be used in schools to support or improve learning to see if there are some technologies which are more promising than others. Some areas, such as writing, for example (Waxman et al., 2002; Torgerson and Zhu, 2003), appear to show particular promise.

ICT summary	+ 4 months ☆☆☆☆
Liao, 2007	0.55
Niemiec & Walberg, 1985	0.32
Pearson, 2005	0.49
Tamim et al. 2011	0.35
Torgerson & Elbourne 2002	0.37
Torgerson & Zhu 2003 (on reading)	-0.05
Torgerson & Zhu 2003 (on spelling)	0.02
Torgerson & Zhu 2003 (on writing)	0.89
Waxman et al. 2002	0.30
Waxman, Lin, Michko 2003	0.45

Additional references

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.

Individualised instruction

Individualising instruction does not tend to be particularly beneficial (Slavin & Karweit, 1985). One possible interpretation is that the role of the teacher becomes too managerial in terms of organising and monitoring learning, but that this is not supportive of improved interaction or using formative feedback to refocus effort. Effect sizes tend overall to be low, or even negative. There have been a number of meta-analyses which have found broadly similar effects. Confirmation comes from other areas such as learning with technology and Bloom's 'mastery learning' (Kulik, Kulik and Bangert Drowns, 1990) where group effects are higher than individual approaches.

Individualised instruction summary	+ 2 months ☆☆☆
Bangert, Kulik & Kulik, 1983	0.10
Horak, 1981	-0.07
Willett, Yamashita & Anderson, 1983	0.17

Additional references

Slavin RE & Karweit NL (1985) Effects of Whole Class, Ability Grouped, and Individualized Instruction on Mathematics *American Educational Research Journal* 22.3 pp. 351-367.

Learning styles

Studies targeting learning with activities that match an individual's identified learning style have not shown convincingly that there is any benefit, particularly for low attaining pupils (Kavale & Forness, 1987). The evidence of lack of effectiveness of approaches such as VAK (visual, auditory, kinaesthetic) has been available for decades (e.g. Arnold, 1968), yet the idea perennially reappears in both research and practice. In some studies controls outperform the learning styles groups, relatively

unusual in educational research, where most interventions show positive effects. There may be some benefit in learners believing that they can succeed in a task if they can choose the particular approach they use. The effect sizes in independent meta-analyses are low (e.g. Kavale & Forness, 1987) 0.14 or negative (Garlinger and Frank, 1986: -0.03), suggesting that only one or two pupils in a class of 30 might benefit from being taught in this way. The evidence for the lack of impact (and in some cases detrimental effect) of using learning styles approaches has been demonstrated in a number of studies and meta-analyses. Positive effects are more likely to be reported by enthusiasts and in areas other than impact on learning outcomes, or where impact may be due to other factors, such as the use of technology (Slemmer, 2002). The unreliability of learning styles tests and assessments has also been the subject of a number of reviews (e.g. Coffield et al. 2004; Pashler et al. 2008). Overall the picture is consistent and robust (Mayer, 2011) that the evidence to support teaching to students' learning styles does not justify the practice.

Learning styles summary	**+ 2 months** ☆☆☆
Kavale & Forness, 1987	0.14
Tamir 1985	0.02
Garlinger & Frank 1986	-0.03
Slemmer 2002	0.27

Additional references

Arnold RD (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.

Mayer RE (2011) Does styles research have useful implications for educational practice? *Learning and Individual Differences* 21 pp 319–320.

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.

Meta-cognition and self-regulation strategies

Metacognitive strategies are teaching approaches which make learners' thinking about learning more explicit in the classroom (Higgins et al. 2005). This is usually through teaching pupils various strategies to plan, monitor and evaluate their own learning (Haller et al. 1988). It is usually more effective in small groups so learners can support each other and make their thinking explicit through discussion (Higgins et al. 2005). Self-regulation (Dignath et al. 2008) refers to managing one's own motivation towards learning as well as the more cognitive aspects of thinking and reasoning. These approaches tend to have a consistent beneficial impact on learning outcomes both in terms of cognitive measures as well as curriculum outcomes (Higgins et al. 2005; Klauer & Phye, 2008). Unusually, such approaches also appear to benefit low attaining pupils more than high achievers (Chiu, 1998), though this may be because the focus of the programme or approach did not extend high achievers' existing learning strategies.

Meta-cognition and self-regulation	+ 8 months ☆☆☆☆
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

Additional references

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

One-to-one tutoring

The evidence from research studies is reasonably consistent, particularly for younger learners who are behind their peers in primary schools and for subjects like reading and mathematics (Wasik & Slavin, 1995). Overall there is good evidence for the benefits of intensive tutoring. Some caution is required in interpreting recent studies such as Chappell et al. (2010) and Tanner et al. (2011) where attempts have been made to generalise these findings, through out of school provision or larger scale policy interventions. The findings from these evaluations suggest that the impact will not necessarily be achieved on a larger scale and that the intensity or quality of interaction in research studies may need to be understood more effectively. This interpretation is supported by evidence that programmes which used experienced teachers and who are given training are more effective than those using volunteers or classroom assistants (Elbaum et al. 2000; Ritter et al. 2009). Evidence also suggest tutoring should be additional or supplemental to normal instruction, rather than replace it. The evidence does not support one-to-one tutoring over pairs or intensive small group work in terms of greater impact on learning (e.g. Torgerson et al., 2011), suggesting that paired or small group tutoring may be a better investment.

One-to-one tutoring summary	+ 5 months ☆☆☆☆
Cohen, Kulik & Kulik 1982 (on tutees)	0.40
Elbaum et al. 2000	0.41
Ritter et al. 2009	0.30
Tanner et al. 2011	*0.14*
Wasik & Slavin 1993	*NPE*

Additional 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.

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.

Parental involvement

Although the involvement of parents is consistently associated with pupils' success at school, the evidence about increasing involvement to improve attainment is much less conclusive. This is particularly the case for poorer families. There is some

evidence that supporting parents with their first child will have benefits for siblings (Seitz & Apfel, 1994). However there is also conflicting evidence which suggests that, at least in terms of early intervention, the involvement of parents does not increase the benefits for learning in schools. 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 clearly also important, though again there is insufficient evidence to show that changing parents' aspirations for their children will raise their children's aspirations and achievement over the longer term. Two recent meta-analyses of studies in the USA suggest that the effects in primary and secondary schools are about 0.27 and 0.25 respectively. Although there is a long history of research into parent involvement programmes, there is surprisingly little robust evidence of the impact of programmes which have tried to increase involvement to improve children's learning (Mattingly et al. 2002). The association between parent involvement and their children's success at school is well established, but rigorous evaluation of approaches to improve children's learning and achievement through parental involvement is more sparse.

Parental involvement summary	**+ 3 months** ☆☆☆
Jeynes 2005	0.27
Jeynes 2007	0.25
van Steensel et al 2011 (family literacy)	0.18

Additional references

Mattingly DJ Prislin R McKenzie TL Rodriguez JL Kayzar B (2002) Evaluating Evaluations: The Case of Parent Involvement Programs *Review of Educational Research*, 72.4 pp 549-576.
Pomerantz EM & Moorman EA Litwack SD (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 NH (1994) Parent-Focused Intervention: Diffusion Effects on Siblings *Child Development* 65.2 pp 677-683.

Peer tutoring/ peer-assisted learning

These are a range of approaches in which learners work in pairs or small groups to provide each other with explicit teaching support (Topping, 2005). In Cross-Age Tutoring an older learner usually takes the tutoring role and is paired with a younger tutee or tutees. Peer-Assisted Learning Strategies (PALS) is a structured approach for mathematics and reading requiring set periods of time for implementation of about 25-35 minutes 2 or 3 times a week. In the collaborative learning strategy 'Reciprocal Peer Tutoring' learners alternate between the role of tutor and tutee. The common characteristic is that the learners take on responsibility for aspects of teaching and for evaluating the success of the learner. The evidence is reasonably consistent and positive especially for mathematics and reading and at both primary and secondary school levels, though there is some evidence that cross-age tutoring is more beneficial (e.g. long term follow up ES 0.22: Topping et al. 2011).

Peer tutoring/ peer-assisted learning summary	**+ 6 months**	☆ ☆ ☆ ☆

Cohen, Kulik & Kulik 1982 (on tutees)	0.40
Cohen, Kulik & Kulik 1982 (on tutors)	0.33
Ginsburg-Block et al. 2006	0.48
Rohrbeck et al. 2003	0.59

Additional references

Topping KJ (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.

Performance pay

Pay incentives for teachers – sometimes referred to a 'pay for performance' – attempt to tie a teacher's remuneration to performance in the classroom. Most teacher pay scales systems use salary schedules that pay teachers based on their qualifications and years of service. Critics point to research showing that there is little correlation between either teachers' years of experience or their holding an advanced degree and a student's achievement level arguing that teachers should be compensated, at least in part, according to the results they produce in their classroom (Education Commission of the States, Teacher Merit Pay, 2010). There is evidence from correlational studies (e.g. Woessman, 2010) that there is a link between teacher pay and pupil performance. The idea behind the Teacher Advancement Program (TAP - Glazerman & Seifullah, 2010) in Chicago is that performance incentives, combined with tools for teachers to track performance and improve instruction, should help schools attract and retain talented teachers and help all teachers produce greater student achievement. However a school-based randomized trial in over 200 New York City public schools found no evidence that teacher incentives increase student achievement or that they change student or teacher behaviour. Instead, teacher incentives may decrease student achievement, especially in larger schools (Fryer, 2011). This is also in line with Martins' (2009) study. Similarly, no evidence was found that the TAP program raised student test scores in maths and reading. Springer, Balou et al. (2010) reported in their three-year Project on Incentives in Teaching (POINT) that, even though the general trend in middle school mathematics performance was upward over the period of the project, students of teachers randomly assigned to the treatment group (eligible for bonuses) did not generally outperform students whose teachers were assigned to the control group (not eligible for bonuses).

By contrast, Lavy (2002) found that monetary performance incentives to teachers and schools caused significant gains in many dimensions of students' outcomes and these were more cost effective than providing them with additional conventional resources. The UK scheme operated in its first year as a general pay increase for almost all teachers at the eligible point of the scale rather than as an individual

performance related pay (PRP) scheme. Such a general pay increase may have little impact on pupil attainment, though it may help retention rates (Burgess et al., 2001).

In terms of costs, under the UK national scheme for individual based performance related pay (Burgess et al., 2001), successful teachers receive an annual bonus of £2000, which they will continue to receive until the end of their career, without needing to reapply. They also move on to a new, upper pay scale where they will be eligible for further performance-related increments. In POINT (Springer, Balou et al., 2010), the maximum bonus an eligible teacher might earn was $15,000—a considerable increase over base pay in this system.

Performance pay summary	**+ 0 months** ☆
Martins 2009	-0.09
Woessman 2010 (correlational study)	*0.25*

Additional references

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, July 2001.

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.

Reducing class sizes

Overall the benefits are not particularly large or clear, until class size is reduced to under 20 or even below 15. There is little advantage in reducing classes from, say, 30 to 25. The issue is whether the teacher changes their teaching approach when working with a smaller class and whether, as a result, the pupils change their learning behaviours (Glass & Smith 1978). Having 15 pupils in a class and teaching them in exactly the same way as a class of 30 will not make much difference. However there is evidence that, when it is successful, the benefits can be identified in behaviour and attitudes as well as on attainment (McGiverin et al., 1989), and that they persist for a number of years (from early primary school through to Key Stage 3). Evidence from both the USA (Finn & Achilles, 1999)) and from the UK does not support the use of teaching assistants as an alternative (see below). It appears to be

important that a class teacher has responsibility for the learning of a class. Optimistically a school might expect a few months additional gain a year for pupils in smaller classes (an effect size of about 0.21) and that this gain will be sustained. There is some evidence to support the additional benefit of professional development when class sizes are reduced to enable teachers to capitalise on the potential benefits by developing their teaching skills and approaches (McGiverin et al., 1989). In addition disadvantaged students may benefit more (Nye et al. 2004).

Reducing class sizes summary	**+ 3 months** ☆☆☆
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

Additional references

Finn JD & Achilles CM (1999) Tennessee's class size study: findings implications, misconceptions *Educational Evaluation and Policy Analysis* 21.2 pp 97-109.

Nye B Hedges LV Konstantopoulos (2004) Do Minorities Experience Larger Lasting Benefits from Small Classes? *Journal of Educational Research*, 98. 2 pp. 94-100.

School uniforms

There is no robust evidence that introducing a school uniform will improve academic performance, behaviour or attendance (Brunsma & Rockquemore , 1998; 2003). 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. 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 meta-analyses of well-controlled interventions of a school uniform or dress code 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 (e.g. Reynolds, 2004), but speculate that adoption of a uniform policy may provide a symbolic and public commitment to school improvement (Reynolds, 2004; Samuels 2002).

School uniforms summary	**± 1 month** ☆
[No meta-analyses of school interventions]	
Samuels 2002 - language arts	0.03
Samuels 2002 - mathematics	-0.06

Additional references

Brunsma DL & 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 DL & Rockquemore K (2003) Statistics, sound bites and school uniforms: a reply to Bodine *Journal of Educational Research* 97.2 pp 72-77.
Reynolds BA (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).

Sports participation

The overall impact on academic achievement tends to be low (e.g. Lewis, 2004; Newman et al. 2010: an effect size around 0.1 to 0.02), 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 (an effect size of 0.8) when combined with a structured numeracy programme. In this circumstance the 'participation' acts as an enticement to undertake additional instruction (Newman et al. 2010). There have been a number of reviews linking the benefits of participation in sport with academic benefits, including a recent systematic review (Newman et al. 2010) for the Department for Culture, Media and Sport (DCMS). There is considerable variation in impact, including some studies which show negative effects. The most promising approaches include direct teaching of academic skills combined with sports participation, rather than sporting activity alone, though the role of sport in supporting initiatives in disadvantaged communities has also been identified (Coalter, 2005; Foster et al. 2005)

Sports participation summary	**+ 3 month** ☆☆
Newman et al. 2010 (academic outcomes)	0.19
Newman et al. 2010 (mathematics)	0.80
Lewis 2004	0.10

Additional references

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.

Summer schools

The effects are reasonably consistent (with an average effect size of about 0.16-0.26: Cooper et al. 2000; Lauer et al. 2006), though usually more beneficial for higher attaining pupils and less effective for low-SES pupils. Programmes are usually more effective in mathematics, when they are specifically tailored to students needs, and when parents are involved (Cooper et al. 2000), and when the summer school uses tutoring and small group work (Lauer et al. 2006). Other variables seem to make less difference, such as whether the teacher is one of the student's usual teachers. Other approaches include summer work placements and youth employment programmes (McClanahan et al. 2004) which appear to have a beneficial impact on students' aspirations.

Summer schools summary	+ 3 months ☆☆
Lauer, Akiba & Wilkerson 2006	0.16
Cooper et al 2000	0.26

Additional references

McClanahan WS Sipe CL & Smith TJ (2004) Enriching Summer Work: An Evaluation of the Summer Career Exploration Program Philadelphia, Pa: Public Private Ventures.

Teaching assistants

Most studies have consistently found very small or no effects on attainment (e.g. Muijs & Reynolds 2003), though pupils' perceptions and attitudes may be more positively affected (Gerber et al. 2001; Blatchford et al. 2009). There are also positive effects in terms of teacher morale and reduced stress of working with a teaching assistant. One clear implication from this is that if teaching assistants are used with the intention of improving the learning of pupils, they should not undertake the tasks they are routinely assigned. There is some evidence that there is greater impact when teaching assistants are given a particular pedagogical role or responsibility in specific curriculum interventions where the effect appears to be greater, particularly with training and support (Alborz et al. 2009). Even here, however, comparisons with qualified teachers suggest they are consistently less effective (achieving about half the gains compared with qualified teachers). There are a number of systematic reviews of the impact of support staff in schools (Farrell et al. 2010), though there are no meta-analyses specifically looking at the impact of teaching assistants on pupils' learning. However, there have been a number of reviews internationally which have consistently found broadly similar effects. The most recent study in the UK (Blatchford et al. 2009) suggests low attaining pupils do less well with a teaching assistant.

Teaching assistants summary	+ 0 months ☆☆
[No meta-analyses of school interventions]	
Gerber et al. 2001 (compared with regular classes)	NPE (0.0 est)
Gerber et al. 2001 (compared with small classes)	NPE (-0.15 est)
Blatchford et al. 2009	NPE (0.00 est)

Additional 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.

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.

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.

Appendix 6: 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. stds	No. Effts	Mod. anls	Pub bias
Ability grouping												
Slavin 1990 (on low attainers)	-0.06	Δ							29	15		
Lou et al 1996 (on low attainers)	-0.12	g	-0.06		-0.01	-0.24	-1.96	1.52			Yes	
After school programmes												
Durlak & Weissberg 2007	0.16	g	0.08		0.01	0.14	-0.16	0.67	55	66		
Scott-Little et al 2002	*NPE*	*d*					*0.38*	*0.50*				
Fashola 1998	*NPE*	*d*					*0.11*	*0.90*				
Lauer, Akiba & Wilkerson 2006	0.16	g	0.03		0.05	0.27			15		Yes	Yes
Arts participation												
Standley 2008	0.32	d	0.05		0.23	0.41	-0.23	1.70	30		Yes	Yes
Winner & Cooper 2000 (Maths)	0.04	d		0.14					15	15	Yes	Yes
Newman et al. 2010												
Newman et al. 2010 (sec maths)	0.03	gc	0.02		0.00	0.06	-0.01	0.05	1	3	Yes	
Newman et al. 2010 (sec Eng)	0.05	gc	0.02		0.01	0.09	-0.01	0.08	1	3	Yes	
Newman et al. 2010 (sec sci)	0.06	gc	0.01		0.05	0.07	0.05	0.06	1	3	Yes	
Newman et al. 2010 (pri/EY cognitive)	0.45	gc	0.09		0.28	0.62	-0.06	1.13	5	10	Yes	
Lewis 2004	0.20	d		0.15					5			
Assessment for learning												
[No meta-analyses of school interventions]												
Wiliam 2002 (KMOFAP)	0.32	g	0.08		0.16	0.48			1	19		
Block scheduling and timetabling												
Dickson et al. 2010 (achievement)	0.11	gc	0.06		-0.01	0.22	-0.14	0.48		7	Yes	No
Dickson et al. 2010 (maths)	-0.02	gc	0.07		-0.16	0.11	-0.14	0.10		6	Yes	No
Dickson et al. 2010 (sci)	0.20	gc	0.07		0.06	0.33	0.13	0.42		4	Yes	No
Lewis et al. 2005 (maths)	-0.10	g	0.01		-0.11	-0.08	-0.15	-0.09		5		
Lewis et al. 2005 (Eng)	-0.17	g	0.01		-0.18	-0.15	-0.25	-0.05		3		
Lewis et al. 2005 (sci)	-0.12	g	0.01		-0.13	-0.10	-0.16	0.11		2		
Early intervention												
Anderson et al. 2003	0.35	Δ					-0.61	0.89	12	29	No	
Lewis & Vosburgh 1988	0.41		0.04	0.47	0.33	0.73	0.21	0.96	65		No	
LaParo & Pianta 2000	0.51								70		No	
Nelson et al. 2003	0.52	g					0.01	1.25	34		Yes	No
Gilliam & Zigler 2001	NPE	Δ					0.07	0.50	13		No	No
Effective feedback												
Kluger & De Nisi, 1996	0.41	d	0.09		0.23	0.59			131	607	Yes	Yes
Lysakowski & Walberg 1982	0.97	d		1.53			-1.09	4.99	54	94	Yes	Yes
Walberg 1982	0.81	d							19	19		
Tenebaum & Goldring 1989	0.72								15	16		
Fuchs & Fuchs 1985	0.71	d	0.09	0.88	0.53	0.89			21	95	Yes	Yes
Homework												
Cooper, Robinson & Patal 2006	0.60	d	0.26	0.64	0.38	0.82	0.39	0.97	6	9		
Paschal, Weinsten & Walberg 1984	0.36	Δ		0.24			-0.60	1.96	15	81		
ICT												

Niemiec & Walberg, 1985	0.32	d		0.4						102	Yes	Yes
Waxman, Lin, Michko 2003	0.45	Δ	0.14	0.72	0.17	0.72			42	29		
Torgerson & Elbourne 2002	0.37	gc	0.20		-0.02	0.77	-0.11	1.15	7	6		
Torgerson & Zhu 2003 (on spelling)	0.02	gc	0.10		-0.17	0.58				4	Yes	Yes
Torgerson & Zhu 2003 (on reading)	-0.05	gc	0.14		-0.33	0.24				4	Yes	Yes
Torgerson & Zhu 2003 (on writing)	0.89	gc	0.33		0.25	1.54				2	Yes	Yes
Waxman et al. 2002	0.30	Δ	0.18	0.71	-0.05	0.83			13	13		
Tamim et al. 2011	0.35	d	0.04		0.27	0.41			25m		Yes	
Liao, 2007	0.55	Δ		0.72			-1.36	2.54	52		Yes	Yes
Pearson et al. 2005	0.49	gc	0.11	0.74	0.27	0.71	-.20	2.68	20	89	Yes	Yes
Individualised instruction												
Horak, 1981	-0.07	Δ					-1.49	0.53	60	129		
Bangert, Kulik & Kulik, 1983	0.10	Δ	0.05		0.00	0.20	-0.84	1.24	49	49		
Willett, Yamashita & Anderson, 1983	0.17	Δ		0.41			-0.87	1.74	130	341		
Learning styles												
Kavale & Forness, 1987	0.14	d	0.06	0.28	0.02	0.27			39			
Tamir 1985	0.02	d							54	13		
Garlinger & Frank 1986	-0.03	d							7	7		
Slemmer 2002	0.27	g							48	51		
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	Yes	Yes
Haller et al. 1988	0.71	d		0.81			0.25	3.80	20	8	No	
Klauer & Phye 2008	0.69	gc	0.05		0.59	0.79	0.59	0.94	17		Yes	
Higgins et al. 2005	0.62	gc	0.09		0.45	0.80	-0.17	1.61	19	19	No	Yes
Chiu 1998	0.67	gc		0.68			-1.25	2.75	43	123		
Dignath et al. 2008	0.62	d*	0.05		0.52	0.72	0.44	1.00	48		Yes	
One-to-one tutoring												
Cohen, Kulik & Kulik 1982 (on tutees)	0.40	Δ	0.07		0.26	0.54			52		Yes	Yes
Cohen, Kulik & Kulik 1982 (on tutors)	0.33	Δ	0.09		0.15	0.51					Yes	Yes
Elbaum et al. 2000	0.41	Δ	0.05		0.32	0.49	-1.32	3.34	29		Yes	Yes
Wasik & Slavin 1993	NPE	Δ					0.20	1.16	16			
Tanner et al. 2011	0.14	d										
Parental involvement												
Jeynes 2005	0.27	gc					0.00	1.78	41	17		
Jeynes 2007	0.25	gc	0.07		0.11	0.35	0.01	0.83	52	20		
van Steensel at al. 2011	0.18	gc	0.06		0.11	0.24			30	47		
Peer tutoring/ peer-assisted learning												
Ritter et al. 2009	0.30	gc	0.06		0.18	0.42	0.26	0.45	28			Yes
Ginsburg-Block et al. 2006	0.48	g		0.39			0.38	0.78	36	36	Yes	
Rohrbeck et al. 2003	0.59	gc	0.10	0.90	0.40	0.78	0.21	0.62	90		Yes	Yes
Cohen, Kulik & Kulik 1982	0.40	Δ	0.07		0.26	0.54	-1.00	2.30	52		Yes	Yes
Performance pay												
Woessman 2010 (correl)	*0.25*											
Martins 2009	-0.09											
Reducing class sizes												
Goldstein, Yang, Omar, Turner & Thompson, 2000	0.20	d					-0.07	0.60	9	36		
Glass & Smith 1978	0.01	Δ							77	725		
McGiverin, Gilman & Tillitski 1989	0.34	d	0.13		0.09	0.59	-0.74	2.24	10	24		

Slavin 1989	0.17	Δ										
School uniforms												
[No meta-analyses of school interventions]												
Samuels 2002 - language arts	0.03	gc	0.11		-0.18	0.23	-0.06	0.03	1	2	No	No
Samuels 2002 - mathematics	-0.06	gc	0.11		-0.26	0.15	-0.06	0.03	1	2	No	No
Sports participation												
Newman et al. 2010 (academic outcomes)	0.19	gc	0.08		0.03	0.35	0.15	0.34	2	2	No	Yes
Newman et al. 2010 (mathematics)	0.80	gc	0.11		0.58	1.02	0.66	0.98	1	2	No	Yes
Lewis 2004	0.10			0.13					5		Yes	No
Summer schools												
Lauer, Akiba & Wilkerson 2006	0.16	g	0.01		-0.20	0.52			14		Yes	Yes
Cooper et al 2000	0.26	d	0.01		0.24	0.28	-0.20	2.70	30			
Lewis 2004	0.10			0.13					5		Yes	Yes
Teaching assistants												
[No meta-analyses of school interventions]												
Gerber et al. 2001 (with regular classes)	NPE (0.0 est)	d					ns	ns				
Gerber et al. 2001 (with small classes)	NPE (-.15 est)	d					-0.13	-0.26				
Blatchford et al. 2009	0.00											

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

No pooled effect (mean – 1.96*se, mean + 1.96*se)	NPE CI to SE

Appendix 7: Bibliography of meta-analyses and other studies used to estimate effect sizes

Abrami, PC Bernard, RM Borokhovski, E Wade, A Surkes, MA 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.

Anderson, LM, Shinn, C Fullilove, MT, Scrimshaw, SC Fielding, JE Normand, J Carande-Kulis, VG (2003) The Effectiveness of Early Childhood Development Programs: A Systematic Review *American Journal of Preventative Medicine* 24: 32-46.

Bangert RL Kulik JA Kulik CC (1983) Individualized Systems of Instruction in Secondary Schools *Review of Educational Research*, 53.2. pp. 143-158.

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

Chiu, CWT (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.

Cohen PA Kulik JA Kulik CC (1982) Educational Outcomes of Tutoring: A Meta-Analysis of Findings *American Educational Research Journal*, 19.2. pp. 237-248.

Cooper, H Charlton, Valentine, JC Muhlenbruck, M Borman GD (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 Robinson, JC Patall EA (2006) Does Homework Improve Academic Achievement? A Synthesis of Research, 1987-2003 *Review of Educational Research*, 76. 1 pp. 1-62

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.

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

Elbaum B Vaughn SM Hughes MT & Moody SM (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.

Fashola OS (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 LS & Fuchs D (1986) Effects of systematic formative evaluation A meta-analysis. *Exceptional Children*, 53.3 pp 199–208.

Garlinger DK & Frank BM (1986) Teacher-student cognitive style and academic achievement: a review and a mini-meta analysis *Journal of Classroom Interaction* 21.2, 2-8.

Gerber SB Finn JD Achilles CM and Boyd-Zacharias J (2001) Teacher aides and students' academic achievement. *Educational Evaluation and Policy Analysis* 23.2 pp 123-143.

Gilliam W & Zigler EF (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.

Glass GV & 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.

Haller EP Child DA & Walberg HJ (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 VM (1981) A Meta-Analysis of Research Findings on Individualized Instruction in Mathematics , *Journal of Educational Research*, 74:4 p.249

Jeynes WH (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 WH (2007). The Relationship Between Parental Involvement and Urban Secondary School Student Academic Achievement: A Meta-Analysis. *Urban Education*, 42.1 pp 82-110.

Kavale KA & Forness SR (1987) Substance Over Style: Assessing the Efficacy of Modality Testing and Teaching *Exceptional Children*, 54.3 pp. 228-39.

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

Kluger AN & 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.

La Paro KM & Pianta RC (2000) Predicting Children's Competence in the Early School Years: A Meta-Analytic Review *Review of Educational Research*, 70.4 pp 443-484.

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 CP (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 RJ & Vosburgh WT (1988). Effectiveness of Kindergarten Intervention Programs: A Meta-Analysis. *School Psychology International*, 9(4), 265-275.

Lewis CW Winokur MA Cobb RB Gliner GS 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 YC (2007) Effects of Computer-Assisted Instruction on Students' Achievement in Taiwan: A Meta-Analysis *Computers and Education*, 48.2 pp 216-233.

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

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.

Martins, PS., 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.

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.

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.

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.

Rohrbeck, C. a, 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.

Samuels, KS (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)

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.

Slavin, RE (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.

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.

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).

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.

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.

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.

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, John B., J. J. Yamashita, and R. D. Anderson. 1983. "A Meta-Analysis of Instructional Systems Applied in Science Teaching." *Journal of Research in Science Teaching* 20(5):405-17.

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.