Draft paper for IACEP conference, Durham, July 2005.
Please do not quote without permission from the authors

How do you solve a problem like Maria? What meta-analysis can tell us about effective educational innovations and the teacher effect.

Elaine Hall
Steve Higgins

Centre for Learning and Teaching
School of Education, Communication and Language Science
University of Newcastle upon Tyne

When engaged upon a meta analysis of interventions in education, whether that be thinking skills programmes, as undertaken by our team (Higgins et al, 2004, EPPI review 2, 2005 in press, EPPI review 3, under review), or some other approach to raising standards, one inevitably refers to the *Ur*-analysis conducted by John Hattie (1999, 2004), which leads us to expect a range of positive effect sizes from 0.1 – the improvement from normal development without any teaching, through 0.25 – the average effect of having a teacher, *any* teacher, through to 0.42 – the mean effect size for all educational interventions, up to around 0.8 – what Hattie describes as 'significant' effects. Of course, some interventions have minimal, or even negative effects but we tend to focus our attentions upon the cluster in the centre, straining our powers of deduction to explain why some things are a little better than others. Inevitably, having trawled the literature and used exclusion criteria to refine the sample, there will be a couple of studies in your final group which have larger effect sizes – perhaps even greater than 1.5. But what if you have a study with effect sizes of 5.93 and 6.56 for maths achievement or 8.55 for attitude towards maths? Do you just designate it the mother of all outliers and move on? Or do you wonder what you'd have to do in order to get that kind of a result, what kinds of pre-conditions might be associated with it, what it might tell us about innovation and teacher effects?

Thinking Skills Interventions

The teaching of thinking skills is an explicit part of the National Curriculum in England and Wales and contributes directly to the DfES's current initiative 'Teaching and Learning in the Foundation Subjects' at Key Stage 3. The descriptive review by Carol McGuinness (1999) provided an overview of current research into the teaching of thinking skills and built on the work of earlier reviews in this area. Nisbet and Davies (1990) listed 30 specific programmes and indicated that there were then over 100 on the market in America. Hamers and Van Luit (1999) show that this is not an English speaking phenomenon and that interest in teaching thinking is evident amongst practitioners and educational researchers in many other European countries.

Thinking skills initiatives have been used in schools in the UK since the early 1980s and have been in existence for somewhat longer, but the term itself is ambiguous and there is disagreement about how it relates to aspects of pedagogy more broadly. Our working definition for the purposes of this series of reviews is that thinking skills interventions can be identified as approaches or programmes which identify for learners translatable mental processes and/or which require learners to plan, describe and evaluate their thinking and learning. These can therefore be characterized as approaches or programmes which require learners to articulate and evaluate specific learning approaches and which identify specific cognitive, affective or conative processes that are amenable to instruction.

A thinking skills approach therefore not only specifies the content of what is to be taught (often framed in terms of thinking processes such as understanding, analysing or evaluating) but also the pedagogy of how it is taught (usually with an explicit role for discussion and articulation of both the content as well as the process of learning or metacognition). Implicit in the use of the term is an emphasis on so-called 'higher-order' thinking, drawing on Bloom and colleagues' taxonomy (Bloom et al., 1956). This consisted of six major categories arranged in the following order *Knowledge, Comprehension, Application, Analysis, Synthesis* and *Evaluation*.

Examples of programmes and approaches commonly used in schools are Instrumental Enrichment (Feuerstein, Rand, Hoffman, and Miller, 1980), Philosophy For Children (Lipman, Sharp and Oscanyan, 1980) Cognitive Acceleration Through Science Education (Adey, Shayer and Yates, 1995), or Somerset Thinking Skills (Blagg, Ballinger and Gardner, 1988). Considerable interest has also been shown in how these more formal programmes can be integrated effectively or 'infused' into teaching approaches and adopted more widely by teachers (McGuinness, Wylie, Greer,& Sheehy,1995; McGuinness, 1999; Leat and Higgins, 2002). Our interpretation of the literature is that an in-depth analysis is needed to evaluate the claims of any impact of such approaches on teaching and learning in classrooms. A further reasonable aim is therefore to try to identify any common features of the impact of implementing thinking skills approaches and to consider how well these relate to wider findings about teaching and learning more broadly, such as formative assessment and feedback in classrooms (Torrance and Pryor, 1998; Black and Wiliam, 1998) or classroom talk and interaction (Mercer, 1995; Galton, 1999). In the reviews carried out by our group, it has been our intention to maintain this focus on the relationship between the general characteristics of thinking skills approaches and interventions and any impact on teaching and learning.

Thinking skills approaches are generally popular with teachers and there is evidence that they seem to support changing patterns of interaction in classrooms (Baumfield and Oberski, 1998; Higgins and Leat, 1997; Leat and Higgins, 2002). This understanding is influenced by concepts and ideas derived from cognitive acceleration (Adey and Shayer, 1994), Instrumental Enrichment (Feuerstein et al., 1980), Philosophy for Children (Lipman, 1994), 'probes' for understanding (White and Gunstone, 1992) reciprocal teaching (Palincsar and Brown, 1984), scaffolding and social constructivism (Wood and Wood, 1996), research on classroom talk (Edwards and Westgate, 1987, Mercer 1995), self-theories (Dweck, 1999) and collaborative group work (Webb and Farrivar, 1994; Galton et al., 1999).

The meta-analysis: what we did and what we meant by it.

The review was conducted as one of the EPPI-Centre[1] co-ordinated reviews to develop an evidence base to inform policy and practice funded by the Department for Education and Skills. The review group set out to answer a number of questions over the three years of funding for the reviews. These questions were clustered around our

[1] http://eppi.ioe.ac.uk/EPPIWeb/home.aspx

central question about what studies were available and relevant to answering the broad question 'What is the impact of the implementation of thinking skills interventions on teaching and learning?' The meta-analysis is the third phase of this systematic review. The first phase addressed the question of evidence for impact on learners and the second turned to the evidence for impact on teachers.

In our first review we identified and described 191 studies up until 2002. We used narrative synthesis methods to address the question 'what is the impact of thinking skills interventions on pupils?'. Twenty-three studies were included and reviewed in depth (Higgins *et al.* 2004). This review identified a number of positive findings about thinking skills approaches but concluded that the selection and implementation of thinking skills approaches as means to improve teaching and learning needed to be based on more precise information on their effectiveness and efficiency. Meta-analysis is a method for pooling the quantitative estimates of effects of interventions from multiple studies to give a more reliable and precise estimate of their benefits (or potential harm). Comparing these estimates across different types of interventions can also pinpoint which aspects of interventions offer the most potential in the classroom. Meta-analysis is proving to be a useful approach to addressing the key question of practitioners interested in thinking skills in terms of "What works?" (e.g. Marzano *et al.* 2001, Hattie *et al*. 1996).

In our meta-analysis we tested the hypothesis that thinking skills interventions have a positive impact by addressing the following questions:

1. What is the magnitude of the quantitative impact of thinking skills interventions on pupils' cognitive achievement?
2. What is the magnitude of the quantitative impact of thinking skills interventions on pupils' curriculum attainment?
3. What is the magnitude of the quantitative impact of thinking skills interventions have on pupils' affective states?

It was our expectation that intervention effects would be positive, since that vast majority of educational innovations have a positive effect but we were interested in whether they would achieve above the 0.4 level cited by Hattie as the average intervention effect size from his meta-analysis of 200,000 effect sizes. Indeed, given that there are interventions with effects ranging from 0.6 to over 1, he considers 0.5 as a minimum for an intervention to be considered 'educationally significant' (Hattie, 2004).

Method

For the first review the team had conducted a broad sweep of the education literature by applying a range of relevant keywords to 16 databases (such as British Education Index, ERIC, PsychINFO, Education Abstracts, etc.) using the online gateways of BIDS, Web of Science and First Search. This yielded references to almost 6,500 books, chapters, dissertations and academic articles which formed the basis of the field. An analysis of 1500 of these items indicated that the majority of references (61%) were concerned with pupils' thinking in school settings as opposed to further and higher education or the use of thinking skills approaches in the workplace. The focus of the review required sources which, implicitly or explicitly, sought to evaluate the implementation of thinking skills programmes in the classroom and so the team identified a subset of 896 by screening abstracts and titles which were subjected to a detailed analysis to fit the review requirements. Criticisms relating to the quality of educational research are not new (e.g. Hillage *et al.*, 1998; Tooley and Darby, 1998) but we feel it is necessary to comment on the quality of abstracting we encountered. Descriptions were often misleading, frequently over-played the amount or types of data actually reported and displayed a tendency to report findings which were inferences rather than conclusions from the data. The screening of abstracts often therefore proved inconclusive and we needed the full paper to determine whether or not the study met the criteria.

At this stage studies were included which:
1. Were set in a school or schools and are concerned with any section of the school population (including pupils with special needs).
2. Evaluated the impact of the implementation of thinking skills interventions on teaching and/or learning.
3. Were concerned with the phases of compulsory schooling (5 –16 year olds).
4. Contained empirical classroom research with data or evidence (pupil outcomes, classroom processes, teacher's views).

Thinking skills interventions were defined as approaches or programmes which require learners to articulate and evaluate learning strategies and/or which identify specific thinking processes that are amenable to instruction in order to improve learning and/or teaching. These interventions could be taught as separate programmes or infused into curriculum teaching. Impact included, for example, pupil and teacher motivation and engagement; patterns of classroom interaction; self-regulation and metacognitive monitoring as well as learners' attainment. 191 studies met all these specifications and were used by the team to form a descriptive 'map' of the field. Of these around 100 appeared to contain quantitative data about impact on pupil attainment.

For the meta-analysis additional inclusion criteria needed to be applied to these studies which examined the impact of thinking skills interventions on pupils. To be considered, studies needed to contain quantitative data of the impact of thinking skills approaches on pupils. This quantitative data had to include either a reported effect size or sufficient detail to calculate an effect size for at least one of the impact measures used. This included standardised tests of cognitive ability (such as Raven's Progressive Matrices); standardised tests of curriculum attainment (such as in reading or mathematics); criterion referenced tests (such as a school science test); tests of impact on affect or conation (such as attitude or self-efficacy scales). The studies also needed to use a control or comparison group which had not received the intervention. We also rejected studies with fewer than 10 research subjects. The initial search was also repeated to identify any studies which met both sets of criteria that had been published since 2002). When these, as well as a small number which focused on children with severe or complex special educational needs, were removed there were 30 studies with evidence about the impact of thinking skills approaches on which to perform the meta-analysis (see Table 1 below). All 30 of these studies were published, 27 as peer-reviewed journal articles, three as chapters in edited collections.

These 30 studies have been reviewed in depth according to the data extraction guidelines for the EPPI-Centre reviewing process (EPPI, 2001), two studies contained effect sizes but insufficient information to calculate the confidence interval and therefore combine the findings with the other studies; and one study contained effect sizes but we have been unable to confirm the calculations using other data presented in the report (we have asked one of the authors of the paper to check the figures from their original data). The data from the studies was initially converted to Cohen's *d* using Coe's (2003) Effect Size calculator.

Table 1: Characteristics of the studies in the meta-analysis

Study	Thinking skills programme or approach	School phase	Intervention effects measured	Effect sizes *d*
Adey (1990)	CASE	Secondary	Cognitive (PRT)	0.21
Adey (2002)	CASE	Primary	Cognitive (PRT)	0.43, 0.47
Cardelle- Elawar (1992)	Metacognitive	Primary	Cognitive (Ravens) Curriculum (Maths) Curriculum (Basic Skills) Affective (attitude to maths)	1.11, 1.42 5.93, 6.56 0.2, 0.74 8.55, 5.36
Chang (1999)	Search, Solve, Create and Share	Secondary	Cognitive Curriculum (Science)	0.7 0.3
Collings (1994)	Formal operations, metacognition	Secondary	Cognitive Curriculum (Science)	1.09, 1.13, 1.53 0.36, 0.68, 0.92
Csapo (1992)	Operational abilities	Primary and Secondary	Cognitive (logical operations) (combinative operations) (systematizing operations)	P= 0.4, 0.64, 0.48, S=0, 0.05, 0.01 P=0.91, 0.8, 0.39, S= 0.4, 0.32, 0.09 P=0.48, -0.29, 0.08, S=0.16, 0.09, -0.1
Cunningham (2002)	Skills for Positive thinking	Primary	Affective	0.23; -0.22; 0.01; 0.38; 0.2
De Koning (1999)	Inductive reasoning	Primary	Cognitive (Ravens) Curriculum (Reading)	0.84 1.24
Donegan (1998)	Vernon	Primary	Affective	0.65
Greenberg (2000)	COGNET/ CEA	Secondary	Curriculum (Maths) Curriculum (Reading)	1, 1.5, 1.1 0.55, 1, 0.2, 0.9, 0.3, 0.9
Haywood (1998)	Feuerstein Enrichment	Secondary	Cognitive (Ravens) Cognitive (PMA) Curriculum (Maths) Curriculum (Reading)	0.7 1.21, 0.77 -.46, -.54, -.33, -.18, -0.62 -0.18
Hoek (1999)	Metacognitive	Secondary	Curriculum	0.2, 0, 0.38
Iqbal (2000)	CASE	Secondary	Curriculum (Maths) Curriculum (Science)	0.004, 0.3, -0.35, 0.43 0.07, -0.7, 0.16, 0.16, 0.67, 0.87
Kaniel (1992)	Feuerstein's Instrumental Enrichment	Secondary	Cognitive (Ravens) Cognitive (Analogies and organiser)	0.67 0.41, 0.83, 1.53
Kramarski (1997)	Metacognitive	Secondary	Cognitive Curriculum (Maths)	0.19, 0.47, 1.08, 2.90, 3.28, 4.44 0.39
Maqsud (1998)	Metacognitive	Secondary	Cognitive Curriculum (Maths) Affective	0.98, 0.83 0.86, 1.19, 1.80, 1.22, 1.21 0.73
Martin (1984)	Feuerstein' Instrumental Enrichment	Secondary	Cognitive (Ravens) Curriculum (Maths)	0.89 0.38
Mercer (1999)	Talk, Reasoning and Computers	Primary	Cognitive (Ravens)	0.32
Muttart (1984)	Feuerstein's Instrumental Enrichment	Secondary	Cognitive (PMA) Curriculum (achievement) Affective	0.89, 0.97 -0.5, 0.41, 0.02, 1.17
Naval-Severino (1993)	Creative thinking	Primary	Cognitive (Torrance TCT)	1.6, 1.57, 1.38
Oladunni (1998)	Metacognitive/ Heuristic	Secondary	Cognitive (creative maths)	1.04
Riding (1987)	Critical Thinking	Primary	Cognitive (Ravens) Curriculum (Maths) Curriculum (Reading)	-0.05, 0.79, 0.68,1.42 -0.82, -0.77, -0.49, 0.69 -0.17, 0.06, 1.10, 1.32
Ritchie (1996)	CoRT (De Bono)	Secondary	Cognitive (CoRT)	0, 0.19, 0.32
Scheinin (1999)	Cognitive education	Secondary	Cognitive Affective	-0.04; 0.3; 0.13; 0.28; 0.43, 0.42; 0.57 0.02; 0.6; 0.11; 0.78, -0.08
Schmid (1990)	Concept mapping	Secondary	Curriculum (Reading)	0.23, 0.40, 0.43
Shayer (1987)	Feuerstein'	Secondary	Cognitive	1.07-1.22

	Instrumental Enrichment		Curriculum (Maths)	0.21; 0.37; 0; 0.23; 0.98; -0.26; -0.35; 0.07; 0.36; 0.26; 0.47; 0.21; 0.57; 0.46
Strang (1993)	Feuerstein Enrichment	Secondary	Curriculum (Science)	1.15
Tenenbaum (1986)	CPR+FB/C	Secondary	Curriculum (Maths) Curriculum (Science)	1.71 1.75
Tzuriel (1994)	Feuerstein's Instrumental Enrichment	Secondary	Cognitive (FE)	0.21
Ward (1993)	Think-Aloud	Primary	Curriculum (Reading)	1.02, 1.61

Sample sizes were in the range 10->900 and because of this we conducted tests which confirmed that sample size and effect size[2] were not significantly related using a 'funnel' plot (Lipsey and Wilson, 2001): see figure 1 below.

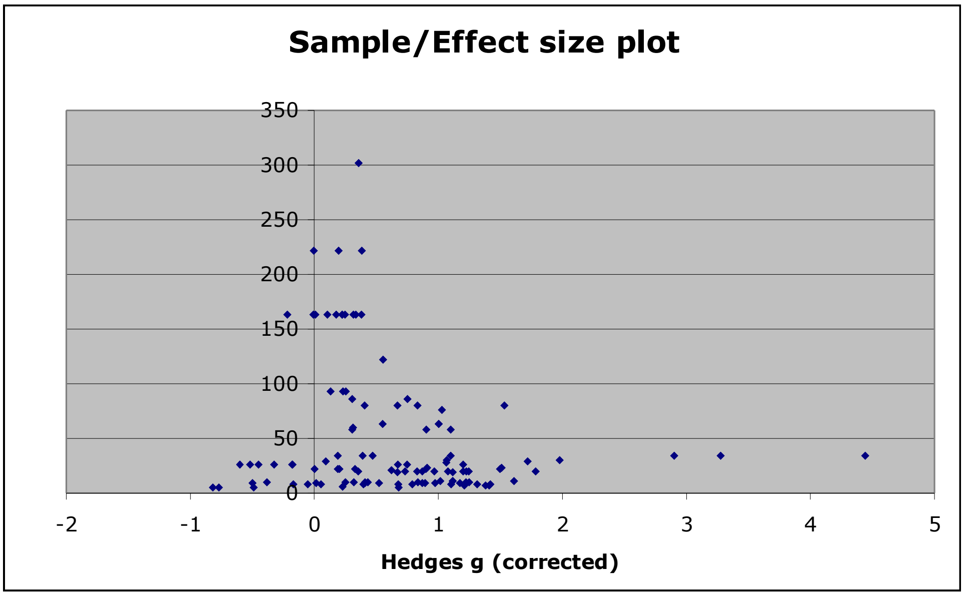

Figure 1 Effect size plotted against sample size

Results of the meta analysis

Analysis of these studies indicate that thinking skills approaches are effective in improving pupils' learning. A meta-analysis of this impact found an overall effect size of 0.71 on cognitive measures (such as tests of reasoning or non-verbal measures such as Ravens Progressive Matrices) and an effect size of 0.66 for curriculum outcomes (such as mathematics or science tests). These effect sizes indicate that an 'average' class of pupils who received such interventions would move from 50th place in a rank of 100 similar classes to about 26th on curriculum tests and to about 24th place on cognitive measures.

However some caution is required in interpreting this meta-analysis as there are considerable differences in the thinking skills approaches and programmes and included in the analysis.

A breakdown of the effect sizes and types of outcomes is presented in Table 2. What is interesting is that although the average impact of thinking skills interventions on all outcome measures was similar to the overall average effect size of all interventions

[2] Effect sizes were entered into the EPPI-Centre database and analysed using the on-line meta-analysis tools. This database uses Hedges g as the standard effect size.

reported by Hattie (2004), the impact on cognitive measures (the key outcome for such interventions) was considerably higher (0.71). Furthermore attempts to improve pupils' thinking do seem to have a positive effect on their attainment in curriculum subjects (0.66). Study heterogeneity was calculated by the meta-analysis software in the EPPI-Reviewer database.

Table 2: Effect sizes and outcome measures

Type of outcome	ES	CI	No of effects/No of studies
All outcomes	0.47	0.43 ⌧0.50	119 effects from 27 studies[3]
Cognitive outcomes	0.71	0.63 ⌧0.78	35 effects from 15 studies[4]
Curricular outcomes	0.66	0.59 ⌧0.74	55 effects from 17 studies[5]
Affective outcomes	0.26	0.20 ⌧0.33	22 effects from 5 studies[6]
Creative outcomes	0.51	0.36 ⌧0.65	8 effects from 4 studies[7]

Meta-analysis does appear to be finding consistent messages in the educational research literature. Our study found an overall mean effect of 0.47, similar to that of Hattie's vast database of meta-analyses (Hattie, 2004). Looking at a smaller part of our study, the impact of Feuerstein's Instrumental Enrichment which is one of the most extensively researched of the thinking skills programmes, our results broadly concur with that of Romney and Samuels' (2001) study which found moderate overall effects and an effect size of 0.43 on reasoning ability (p.28) – our findings were similar with an overall effect size of 0.32 (26 effects from 5 studies) and an effect size of 0.49 on tests of reasoning (6 effects from 4 studies). This suggest to us that the findings from meta-analysis are worth considering as a part of the story of 'what works' in education by offering comparative information about *how well* different interventions work. This echoes Hattie's (1999) plea that:

- "We need to make <u>relative</u> statements about what impacts on student work.
- We need estimates of magnitude as well as statistical significance – it is not good enough to say that this works because lots of people use it etc., but that this works because of the magnitude of impact.
- We need to be building a model based on these relative magnitudes of effects."

Dealing with variation and context: the 'problem like Maria'

One of the studies in our final group reports the impact of a series of metacognitive – focused initiatives to improve maths performance and attitude towards maths conducted in an elementary school in the United States by Maria Cardelle- Elawar (Cardelle- Elawar, 1992). The students in the study were sixth graders (mean age 11 years) who were identified by having a lower than average score on the Iowa Test of

---

[3] Heterogeneity statistic Q = 1.1E+03 df = 120 p = 0 Test statistic (combined effect) z = 23.6 p < 0.001 Inverse Variance (fixed effects model)

[4] Heterogeneity statistic Q = 195 df = 34 p = 0 Test statistic (combined effect) z = 17.9 p < 0.001 Inverse Variance (fixed effects model)

[5] Heterogeneity statistic Q = 443 df = 55 p = 0 Test statistic (combined effect) z = 17.6 p < 0.001 Inverse Variance (fixed effects model)

[6] Heterogeneity statistic Q = 364 df = 21 p = 0 Test statistic (combined effect) z = 7.73 p < 0.001 Inverse Variance (fixed effects model)

[7] Heterogeneity statistic Q = 21.7 df = 7 p = 0.00282 Test statistic (combined effect) z = 6.77 p < 0.001 Inverse Variance (fixed effects model)

Basic Skills (ITBS) and they came from a public school in a predominantly Hispanic, deprived socio-economic area. Over two waves of experimentation 122 students took part. The intervention used existing classes and the experimental teaching took place within the school. The researcher herself worked with the students in the first wave and she trained the teachers in the school to administer the intervention in the second wave, in order to explore the possibility that her own presence had skewed the results. Video recording of the teachers and independent assessment ensured fidelity of implementation (.79). Both waves produced gains for the experimental groups in cognitive scores (Raven's matrices), basic skills (ITBS), maths attainment (20 mathematics problems) and attitude to maths (as calculated by Aitken E and V attitude scales, see table 3 below), most of which could be considered significant.

Table 3: Results from Cardelle-Elawar (1992)

Measure	Experiment 1 (effect size)	Experiment 2(effect size)
Raven's matrices	1.42	1.11
Iowa Test of Basic Skills	0.2*	0.74
Mathematics problems	5.93	6.56
Aitken E and V attitude scales	8.55	5.36

*This result would not be considered significant.

The 'cultural norms' of meta-analysis and systematic review suggest that we should not take too much interest in this study: the number-crunching will even out the impact measures and the study weightings will highlight the important elements that this study has in common with others. We have argued elsewhere (Hall and Higgins, 2004) that the synthesising process has a 'blanding' as well as a blending effect: producing 'virtual smoothies' which cannot tell us much about the original ingredients. We therefore feel it is important to explore some of the possible reasons why Cardelle-Elawar achieved such good results. The 'problem like Maria' seems to generate at least two possible types of explanation: the first is that excessive improvement follows from an intervention superimposed upon severe underachievement or poor teaching – which we have come to refer to as 'parasitic pedagogy'; the second is that there is a 'virtuous cycle' of well-targeted intervention, authentic teacher belief and skill and sustained feedback loops for both teacher and students which generates an unusually high 'teacher effect'. We will explore both of these possibilities in turn.

Parasitic pedagogies

If you have an enormous effect size, it may reflect several things: the excellence of the intervention; the timeliness in terms of the pupils' needs (ZPD or Piagetian developmental framework) or the previous under-achievement of the pupils in this area. There is an important distinction to be drawn between *low achievement* related to intrinsic ability in the given area and *under-achievement* related to previous lack of experience or poor teaching. There is a degree of evidence to suggest that successful innovations may be parasitic upon previous underachievement and that significant short term gains may reflect 'catch-up' to normal levels of performance. This may explain the particular success of interventions in maths and science, where standards of teaching vary more widely than in literacy or humanities.

Unpack: what are the implications of this?

Virtuous cycles

In our review, we used fidelity of implementation as one of our quality criteria but there is a further element, one of authenticity. While it is important for a systematic review to know that Instrumental Enrichment, for example, has been correctly taught, it is equally important, in our view, that those teacher who delivered it genuinely believed that IE might benefit their teaching and learning. Mere compliance is less likely to engage and enthuse pupils the level of belief and commitment felt by the teacher and (instantly) picked up upon by the children helps us to understand the greater success of 'early adopters' and the low impacts of 'roll-out'. Authenticity also relates to fidelity not only of implementation but also fidelity to context: a reflective understanding of an innovation not as 'A Good Thing' but relevant to the needs of learners in particular places. This is sometimes linked to ownership in terms of discovery and sometimes in terms of 'conversion'.

Moreover, the teachers' faith in the value of the innovation is needed to balance the risk of innovation, the personal exposure and the risk to the pupils. It is important to emphasise teachers' commitment to not messing up each child's one chance of Year 7 through 'change for change's sake', which is often interpreted as conservatism or resistance to change but might more helpfully be seen as Hippocratic – 'first do no harm' – in nature. Innovation is risky, exposing, normally more work than usual and necessarily disruptive. The psychological effects of unsuccessful or insufficiently successful innovation are severe, as resistance to further risk and change becomes defensively entrenched (Anna Freud, 1938). Risk taking is nested for us within an understanding of the social capital relationships (Bourdieu, 1999) of the particular learning context: social capital is not a universal currency and schools themselves have differing 'rates of exchange' which are often opaque/ unexplored, historically shaped or otherwise 'taken for granted'. This engagement with capital and field has to coexist with an engagement with the individual personal and professional self-concept: the 'defended self' can make only limited use of even abundant capital in a propitious field.

The evidence suggests that successful innovations tend to be underpinned by feedback loops: Hattie's significant innovations group is packed with things that tell the teacher "How am I doing?" (not just "How are the pupils doing?") this focuses the teacher's attention on what is in their sphere of influence. Feedback loops are ipsative formative assessment for teachers as critical learners – their usefulness may be dependent on their frequency and on their resonance – whether as comforting supporting vibrations or cognitive dissonance. The self-concept and tolerance of ambiguity of individual teachers will shape how they can make use of feedback loops. Teacher autonomy has been challenged by top-down curriculum change and prescriptive inspection regimes, so it is important to concentrate on what can be done within current frameworks and limitations, rather than to adopt passive or helpless behaviour. This is why action research as continuing professional development is motivating and enabling: the cyclical process, the rooting in real problems and questions, the support for change, the relatively small scale of that change (risk/stakes).

Innovation effects wear off fast – even programmes which have evidence for long term difference and transfer (eg CASE, CAME) experience some tail off – this may be because of automation by the teacher and the less targeted nature of each repetition, the changes in each cohort of students which could lead to greater divergence over time or simple boredom. It is vital that movement and change be considered central elements in the culture of teaching, from ITT onwards and that teachers and teacher educators recognise the overwhelming need to counter the initiative/ toolkit paradigm of 'solutions' in policy discourse. For these reasons we have found it useful, in our work with teachers, to think less in terms of single innovation-evaluation relationships but more helpfully as a continuing professional development model of practitioner enquiry/ action research which uses a cyclical model of question generation, and change which should support continuing teacher engagement (Hall et al in press). This requires a support network: the tail off of effect, 'clean windows syndrome' and other psychological challenge require both peer and 'supportive other' support. Codification of results and a wider critical engagement are also key features of these effective networks, since challenge has to complement support..

'Good practice' and 'evidence based practice' have increasingly been delineated and described in a homogenous and reductive way and this is a weakness, as every change and development looks like a retraction or admission of failure. Rather than portraying teachers in research-led practice as searching for definitive solutions we should emphasise that the goal is better questions. The action research model of context-specific innovations will build up a pool of reflective, evidence-based practice with thematic links rather than cookie-cutter applicability to other schools. The question should not be "Will that work in my school?" but "What is working well/ not so well in my school?".

References

Adey, P. & Shayer, M. (1993) An Exploration of Long-Term Far-Transfer Effects Following an Extended Intervention Program in the High School Science Curriculum, *Cognition & Instruction*, 11(1), pp. 1-29.
Adey, P.S., Shayer, M. and Yates, C. (1995) *Thinking Science: The Curriculum Materials of the CASE Project* London: Thomas Nelson and Sons.
Bassey, M. (2000) *Fuzzy Generalisations And Best Estimates Of Trustworthiness: a step towards transforming research knowledge about learning into effective teaching practice.* TLRP Annual Conference 2000.
http://www.tlrp.org/acadpub/Bassey2000.pdf
Baumfield , V.M. and Oberski, I.O (1998) What do Teachers Think about Thinking Skills? *Quality Assurance in Education*, 6:1, pp 44-51.
Black, P. and Wiliam, D. (1998) Assessment and Classroom Learning *Assessment in Education* 5.1. pp 7-71
Blagg, N., Ballinger, M. and Gardner, R. (1988) *Somerset Thinking Skills Course Handbook* Oxford: Basil Blackwell.
Bloom, B. S. (Ed.) (1956) *Taxonomy of Educational Objectives, the classification of educational goals - Handbook I: Cognitive Domain*. New York: McKay
Bourdieu, P. (1999) *The Weight of the World.* Cambridge: Polity

Csapó, B. (1992) Improving Operational Abilities in Children In A. Demetriou, M. Shayer and A. Efklides (Eds) *Neo-Piagetian Theories of Cognitive Development: Implications and Applications For Education* London: Routledge.
Dweck. C. (1999) *Self Theories: Their Role in Motivation, Personality and Development.* Hove (Sussex): Psychology Press (Taylor & Francis).
Edwards, A. and Westgate, D. (1987) *Investigating Classroom Talk.* London: Falmer Press.
EPPI (2001) *EPPI-Centre Review Group Manual Version 1.1*, London: Social Science Research Unit, Institute of Education, University of London.
EPPI (2002) *EPPI-Centre Core Keywording Strategy: Data Collection for a Register of Educational Research Version 0.9.5* London: EPPI Centre, Social Science Research Unit, Institute of Education, University of London.
Fan, X. (2001) *Statistical significance and effect size in educational research: two sides of a coin.* Journal of Educational Research, 95 (5) 275-282
Feuerstein, R. Rand, Y., Hoffman, M.B. and Miller, R. (1980) *Instrumental Enrichment: an intervention programme for cognitive modifiability* Baltimore: University Park Press.
Freud, A. (1938) *The ego and the mechanism of defence.* New York: International Universities Press
Galton, G., Hargreaves, L., Comber, C. and Pell, A. (1999) *Inside the Primary Classroom 20 Years On*, London: Routledge

Hall, E., Wall, K., Higgins, S., Stephens, L., Pooley, I. and Welham, J. Learning to Learn with Parents: lessons from two research projects. *Improving Schools* 2005, In press.

Hall, E. and Higgins, S. *Picking the strawberries out of the jam: thinking critically about narrative reviews, systematic reviews and meta-analysis* Paper presented at the British Education Research Association conference, at Manchester Metropolitan University, September 2004 available at:
http://www.leeds.ac.uk/educol/documents/00003835.htm
Hamers, J.H.M. & Overtoom, M.Th. (Eds.). (1997). *Teaching thinking in Europe. Inventory of European programmes*. Utrecht: SARDES

Hattie, J., Biggs, J., and Purdie, N. (1996). *Effects of learning skills interventions on student learning: A meta-analysis.* Review of Educational Research, 66, 99-136.
Hattie, J. (1999) *"Influences On Student Learning".* Inaugural Lecture: Professor of Education. University of Auckland. August 2, 1999.
http://www.arts.auckland.ac.nz/edu/staff/jhattie/Inaugural.html
Hattie, J. (2004) *Factors that influence children's learning: results from a study involving 500 meta-analyses* Paper presented at the ESRC Seminar 'Effective Educational Interventions', University of Newcastle upon Tyne, July 8th, 2004

Higgins, S. (2001) *Thinking Through Primary Teaching.* Cambridge, Chris Kington Publishing.

Higgins, S. and Leat, D. (1997) "Horses for courses or courses for horses: what is effective teacher development?" *British Journal of In-Service Education* 23.3 pp 303-314

Higgins, S. and Moseley D. (2002) 'Raising Achievement in Literacy through ICT' in M. Monteith (ed), *ICT and Literacy* Buckingham: Open University Press.
Higgins S, Baumfield V, Lin M, Moseley D, Butterworth M, Downey G, Gregson M, Oberski I, Rockett M and Thacker D (forthcoming) *Thinking skills approaches to*

effective teaching and learning. In: Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education.
Hillage, J., Pearson, R., Anderson, A. and Tamkin, P. (1998). *Excellence in research on schools.* London: DfEE.
Hunter, J. E. and Schmidt, F. L. (1990) *Methods of Meta Analysis: correcting error and bias in research findings* New York: Sage
Leat, D. and Higgins S. (2002) The role of powerful pedagogical strategies in curriculum development *The Curriculum Journal* 13.1 pp 71-85 ISSN 0958-5176.
Leat, D. and Lin, M. (2003) Developing a Pedagogy of Metacognition and Transfer: some signposts for the generation and use of knowledge and the creation of research partnerships *British Educational Research Journal* 29.3 pp 383 – 414.
Lipman, M., Sharp, A. and Oscanyan, F. (1980) *Philosophy in the Classroom* Princeton: Temple University Press.
Lipsey, M. W., and Wilson, D. B. (2001). *Practical Meta-Analysis. Applied Social Research Methods Series* (Vol. 49). Thousand Oaks, CA: SAGE Publications
McGuinness, C. (1999) *From Thinking Skills to Thinking Classrooms: A review and evaluation of approaches for developing pupils' thinking*. Nottingham: DfEE Publications
McGuinness, C., Wylie, J., Greer, B. & Sheehy, N.A.F. (1995) Developing children's thinking: A tale of three projects, *Irish Journal of Psychology*, 16(4), pp. 378-388.
Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.
Mercer, N. (1995) *The guided construction of knowledge: Talk amongst teachers and learners* Clevedon: Multilingual Matters.
Nisbet, J. and Davies, P. (1990) The curriculum redefined: learning to think-thinking to learn, *Research Papers in Education,* 5.1, pp. 49-92.
Olkin, I. (1992) *Meta-analysis: methods for combining independent studies.* Statistical Science, 7, 226
Romney, D.M. and Samuels, M.T. (2001) A meta-analytic evaluation of Feuerstein's Instrumental Enrichment program *Educational and Child Psychology* 18(4) pp.19-34
Slavin, R. (2004) *Evidence-Based Education Policies: Promises and Pitfalls* Keynote Address, British Educational Research Association Conference, UMIST, Manchester, September 15-18th September.
Tooley, J. and Darby, D. (1998) *Educational research - a critique.* London: Office for Standards in Education
Torrance, H. and Pryor, J. (1998) *Investigating formative assessment: teaching, learning and assessment in the classroom* Open University Press.
White, R. and Gunstone, R. (1992). *Probing Understanding* London: Falmer Press.
Webb, N. and Farivar, S. (1994) Promoting helping behaviour in cooperative small groups in middle school mathematics. *American Educational Research Journal,* 31.2 pp. 369-396.
Wood, D. and Wood, H. (1996) Vygotsky, Tutoring and Learning, *Oxford Review of Education.* Vol. 22.1, pp. 5-16.

Studies in the meta-analysis

1. Adey, P. and M. Shayer (1990). "Accelerating the Development of Formal Thinking in Middle and High-School-Students." *Journal of Research in Science Teaching* 27(3): 267-285.

2. Adey, P., A. Robertson, et al. (2002). "Effects of a cognitive acceleration programme on Year I pupils." *British Journal of Educational Psychology* 72: 1-25.

3. Cardelle-Elawar, M. (1992). "Effects of Teaching Metacognitive Skills to Students with Low Mathematics Ability." *Teaching & Teacher Education* 8(2): 109-21.

4. Chang, C.-Y. and J. P. Barufaldi (1999). "The Use of a Problem-Solving-Based Instructional Model in Initiating Change in Students' Achievement and Alternative Frameworks." *International Journal of Science Education* 21(4): 373-88.

5. Collings, J. N. (1994). "Some fundamental questions about scientific thinking." *Research in Science & Technological Education* 12(2): 161-173.

6. Csapó, B. (1992). Improving operational abilities in children In A. Demetriou, M. Shayer and A. Efklides. *Neo-Piagetian theories of cognitive development* London, Routledge.

7. Cunningham, E.G., Brandon, C.M. and Frydenberg, E. (2002) Enhancing coping resources in early adolescence through a school-based program teaching optimistic thinking skills. *Anxiety Coping and Stress* 15(4): 369-381

8. De Koning, E. and J. H. M. Hamers (1999). Teaching inductive reasoning: theoretical background and educational implications. In J. H. M. Hamers, J. E. H. Van Luit and B. Csapó. *Teaching and Learning Thinking Skills* Abingdon, Swets and Zeitlinger.

9. Donegan, A. L. and J. O. Rust (1998). "Rational Emotive Education for Improving Self-Concept in Second-Grade Students." *Journal of Humanistic Education & Development* 36(4): 248-56

10. Greenberg, K. H. (2000). "Attending to hidden needs: The Cognitive Enrichment Advantage perspective." *Educational & Child Psychology*. Special Issue: Psychological influences upon educational intervention 17(3): 51-69.

11. Haywood, H. C. and et al. (1988). "Cognitive Education with Deaf Adolescents: Effects of Instrumental Enrichment." *Topics in Language Disorders* 8(4): 23-40.

12. Hoek, D., P. van den Eeden, et al. (1999). "The Effects of Social and Cognitive Strategies Instruction on the Mathematics Achievement in Secondary Education." *Learning & Instruction* 9: 427-448.

13. Iqbal, H. M. and M. Shayer (2000). "Accelerating the development of formal thinking in Pakistan secondary school students: Achievement effects and professional development issues." *Journal of Research in Science Teaching* 37(3): 259-274.

14. Kaniel, S. and R. Reichenberg (1992). "Instrumental Enrichment--Effects of Generalization and Durability with Talented Adolescents." *Gifted Education International* 8(3): 128-35

15. Kramarski, B. and Z. R. Mevarech (1997). "Cognitive-metacognitive training within a problem-solving based Logo environment." *British Journal of Educational Psychology* 67(4): 425-445

16. Maqsud, M. (1998). "Effects of metacognitive instruction on mathematics achievement and attitude towards mathematics of low mathematics achievers." *Educational Research* 40(2): 237-243.

17. Martin, D. S. (1984). "Cognitive Modification for the Hearing Impaired Adolescent: The Promise." *Exceptional Children* 51(3): 235-42.

18. Mercer, N., R. Wegerif, et al. (1999). "Children's talk and the development of reasoning in the classroom." *British Educational Research Journal* 25(1): 95-111.

19. Muttart, K. (1984). "Assessment of Effects of Instrumental Enrichment Cognitive Training." *Special Education in Canada* 58(3): 106-08.

20. Naval-Severino, T. (1993). "Developing creative thinking among intellectually able Filipino children from disadvantaged urban communities." *Gifted Education International* 9(2): 119-123.

21. Oladunni, M. O. (1998). "An experimental study on the effectiveness of metacognitive and heuristic problem solving techniques on computational performance of students in mathematics." *International Journal of Mathematical Education in Science & Technology* 29(6): 867-874.

22. Riding, R. J. and S. D. Powell (1987). "The Effect on Reasoning, Reading and Number Performance of Computer presented Critical Thinking Activities in Five-year-old Children." *Educational Psychology* 7(1): 55-65.

23. Ritchie, S. M. and J. Edwards (1996). "Creative Thinking Instruction for Aboriginal Children." *Learning & Instruction* 6(1): 59-75.

24. Scheinin, P. M. and J. Mehtäläinen (1999). Applying the theory of knowledge to teaching thinking. In J. H. M. Hamers, J. E. H. Van Luit and B. Csapó. *Teaching and Learning Thinking Skills* Abingdon, Swets and Zeitlinger.

25. Schmid, R. F. and G. Telaro (1990). "Concept Mapping as an Instructional Strategy for High School Biology." *Journal of Educational Research* 84(2): 78-85.

26. Shayer, M. and F. Beasley (1987). "Does instrumental enrichment work?" *British Educational Research Journal* 13(2): 101-119.

27. Strang, J. and M. Shayer (1993). "Enhancing high school students' achievement in chemistry through a thinking skills approach." *International Journal of Science Education* 15(3): 319-337.

28. Tenenbaum, G. (1986). "The Effect of Quality of Instruction on Higher and Lower Mental Processes and on the Prediction of Summative Achievement." *Journal of Educational Research* 80(2): 105-114.

29. Tzuriel, D. and M. Alfassi (1994). "Cognitive and Motivational Modifiability as a Function of the Instrumental Enrichment (IE) Program." *Special Services in the Schools* 8(2): 91-128.

30. Ward, L. and D. Traweek (1993). "Application of a Metacognitive Strategy to Assessment, Intervention, and Consultation: A Think-Aloud Technique." *Journal of School Psychology* 31(4): 469-85.