REVIEW

October 2004

Thinking skills approaches to effective teaching and learning: what is the evidence for impact on learners?

Review conducted by the Thinking Skills Review Group

The EPPI-Centre is part of the Social Science Research Unit, Institute of Education, University of London

AUTHORS

This report was written by a team of colleagues from the former Thinking Skills Research Centre (now the Centre for Learning and Teaching), based at the University of Newcastle upon Tyne: namely, Steve Higgins, Vivienne Baumfield, Mei Lin, David Moseley with Marie Butterworth, Graham Downey, Maggie Gregson, Mel Rockett (members of the Thinking Skills Core Review Group) and Iddo Oberski and David Thacker from the Consultation and Critical Advisory Panel. Other members of the Thinking Skills Review Group Consultation and Critical Advisory Panel provided valuable support and critique, influencing both the direction and scope of the review and editing drafts of the report.

This document reports on a systematic review of the impact of thinking skills approaches on learning and teaching in schools. It sets out the background, processes and findings of the review. The Review Group is co-ordinated by the Centre for Learning and Teaching at Newcastle University and the review was undertaken by a core group, including teachers, LEA advisers and academics. Practitioners played a key part in this review, particularly in defining the scope of the review, in clarifying the terms used in searching and keywording so as to identify studies with practical implications for teachers and schools.

CORE REVIEW GROUP

Dr Vivienne Baumfield	Senior Lecturer in Education, University of Newcastle
Marie Butterworth	Deputy Headteacher, Heaton Manor School
Graham Downey	Headteacher, Kramel First School, Northumberland
Maggie Gregson	Lecturer in Post-compulsory Education, University of Sunderland
Steve Higgins	Senior Lecturer in Education, University of Newcastle
Dr Mei Lin	Lecturer, University of Newcastle
David Moseley	Reader in Education, University of Newcastle
Mel Rockett	LEA Adviser, Northumberland

CONSULTATION AND CRITICAL ADVISORY PANEL

Sue Eagle	Primary School Headteacher, Norfolk
Jane Brooke	LEA Adviser, Cheshire
Dr Carol McGuinness	Professor of Educational Psychology, Queen's University, Belfast
David Thacker	Secondary Headteacher (retired)
Dr Iddo Oberski	Lecturer in Education, Stirling University
Dr Carolyn Tan	Senior Lecturer in Early Years Education, National Institute of Education (NIE), Nanyang Technological University (NTU), Singapore
Dr William Wu	Co-Director of the Thinking Qualities Initiative at the Centre for Educational Development, Hong Kong Baptist University

ACKNOWLEDGEMENTS

This Group is registered as an 'EPPI-Centre Review Group' and is part of the initiative for evidence informed practice and policy in education funded by the Department for Education and Skills (DfES) at the EPPI-Centre (Evidence for Policy and Practice Information and Co-ordinating Centre), Social Science Research Unit, University of London Institute of Education. The training and support from the EPPI-Centre team was invaluable in conducting the review, and the advice from Dina Kiwan, Katy Sutcliffe, Angela Harden, Nicholas Houghton, Zoe Garrett and James Thomas was especially helpful.

The Group also acknowledges the feedback of referees in developing the thinking of the group. Comments on the initial proposal and on the protocol were useful in sharpening aspects of the project.

The encouragement and support from Judy Sebba at the Department for Education and Skills are also acknowledged; her perspective on policy and thinking skills has been valuable. Thanks go to Jean Halligan for her support with the early stages of literature screening. Finally, a debt of gratitude is owed to Susan Angus without whose librarianship and information management skills the project would have foundered in its early stages.

Funding

The funding provided by DfES through the EPPI-Centre was essential to establish the group and cover the costs of attending training. We estimate this funding covered a little less than a third of the actual costs of the review. Additional costs have been met by the Thinking Skills Research Centre at Newcastle University and academic staff have benefited from HEFCE funding (albeit declining) for research. The support of the schools and local education authorities as well as the individuals themselves in the core Review Group and Consultation and Critical Advisory Panel was essential. Without this generous donation of time and effort, the review could not have been undertaken.

CONFLICTS OF INTEREST

Throughout the review, we have tried to be consistent and transparent in conducting the review, working within the EPPI-Centre guidelines, methodology and quality-assurance procedures for systematic reviewing, and involving members of the core Review Group and Advisory Panel. We wanted to ensure that our own pre-existing interest in the implementation and evaluation of thinking skills programmes and approaches did not influence our working processes or findings. We can, however, record our keen interest in the outcomes of the review and acknowledge that this may have influenced the review in ways which are not apparent to us.

LIST OF ABBREVIATIONS

BDI	Biblioscape Database
BEI	British Education Index
CAME	Cognitive acceleration through mathematics education
CASE	Cognitive acceleration through science education
CATE	Cognitive acceleration through technology education
CERUK	Current Educational Research in the UK
DfES	Department for Education and Skills
ECO	Electronic Collections Online
EPPI-Centre	Evidence for Policy and Practice Information and Co-ordinating Centre
ERA	Education research abstracts
ERIC	Educational Resources Information Centre
HEI	Higher education institution
IBSS	International Bibliography of the Social Sciences
LEA	Local education authority
OCLC	Online Computer Library Centre
OFSTED	Office for Standards in Education
REEL	Research Evidence in Education Library
TTA	Teacher Training Agency

This report should be cited as: Higgins S, Baumfield V, Lin M, Moseley D, Butterworth M, Downey G, Gregson M, Oberski I, Rockett M and Thacker D (2004) Thinking skills approaches to effective teaching and learning: what is the evidence for impact on learners. In: Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

First submitted: July 2003
Revised: October 2003 and January 2004
Final version: June 2004

TABLE OF CONTENTS

SUMMARY

Background

The teaching of thinking skills is an explicit part of the National Curriculum in England and contributes directly to an initiative of the Department for Education and Skills (DfES) 'Teaching and learning in the Foundation subjects' at Key Stage 3 (DfES 2003). This emphasises the importance of thinking skills approaches for the promotion of effective questioning and extending pupils' oral responses in classrooms as well as the potential contribution to assessment for learning.

Our working definition for the purposes of this review is that thinking skills interventions are 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 characterised as approaches or programmes which:

- require learners to articulate and evaluate specific learning approaches; and/or
- identify specific cognitive and related affective or conative processes that are amenable to instruction

A thinking skills approach therefore not only specifies *what* is to be taught but also *how* it is taught: the content of lessons and the teaching approach form an integral part of thinking skills approaches to teaching and learning. 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) and Somerset thinking skills (Blagg, Ballinger and Gardner, 1988). Considerable interest has also been shown by teachers and policymakers in how these more formal programmes can be integrated effectively or 'infused' into teaching approaches and adopted more widely by teachers (McGuinness, Wylie, Greer and Sheehy, 1995; McGuinness, 1999; Leat and Higgins, 2002).

A systematic review was needed:

- to provide potential users with an overview of current research and evidence in the field by updating and extending the scope of earlier reviews, which have attempted to evaluate evidence from a range of thinking skills approaches (e.g. Sternberg and Bhana, 1986) or which have focused on a particular programme (such as Romney and Samuels' (2001) meta-analysis of evidence of the impact on learners of Feuerstein's instrumental enrichment (FIE)

- to identify and analyse the empirical evidence available to support the teaching of thinking in schools in order to test the conclusions of the positive but largely descriptive reviews recently undertaken (McGuinness, 1999; Wilson 2000)

Aims

The aim of the Review Group is to investigate the impact of thinking skills interventions on teaching and learning in classrooms over a series of focused reviews. Our main review question is:

What is the impact of the implementation of thinking skills interventions on teaching and learning?

Underpinning this main question are the following considerations:

What are the parameters for defining a particular pedagogy or curriculum development as a thinking skills approach?

How closely do the findings of the studies of the impact of thinking skills approaches correlate with current understanding of effective teaching and learning?

For the in-depth review, a narrower focus was identified for the central question about the impact of thinking skills interventions:

What is the evidence for impact on learners' attainment in schools?

Methods

The review focused on studies which both explicitly and implicitly evaluated the implementation of thinking skills programmes and approaches in classrooms during the ages of compulsory schooling (5 to 16) in all areas of the curriculum. These were mainly accessed through articles published in peer-reviewed academic journals and unpublished materials (such as conference papers or LEA evaluation studies or dissertations) where the focus of the study was on the implementation and evaluation of thinking skills approaches or programmes in classrooms. Nearly 6,500 chapters, articles and papers were identified as potentially relevant from searching electronic databases of references. These were screened on the basis of title and also an abstract, if available, and about 800 of them were judged to meet the inclusion/exclusion criteria of the review. Full texts of these were ordered and those that were obtained were screened in detail against the same criteria. The 191 reports that met these criteria were entered into a database using keywords in accordance with EPPI-Centre Core Keywording Strategy (version 0.9.5). This enabled the group to build a 'map' of the literature about thinking skills and to identify more precisely those studies that might answer the review question in terms of the study type and focus.

From this database, 23 studies were selected for in-depth analysis and evaluation because they have a combination of quantitative and qualitative empirical data about the impact of thinking skills approaches on pupils' attainment. The Review Group felt that the combination of quantitative and qualitative data from empirical work in classrooms would be most likely to provide the evidence of impact that would be most relevant to users, particularly practitioners, and contain sufficient detail to evaluate how applicable the results would be to other educational settings.

Results

Identification of studies

Before we presented a systematic map linked to our review question, we made a broad sweep to see in which general areas of thinking skills figured as an aspect of research into teaching and learning. Analysis of just over 1,500 of the sources identified in the British Education Index and ERIC which used the keywords 'thinking skills' showed that the term is used broadly in the reporting of research across a number of disciplines. The majority of references are in school settings looking at pupils' thinking (61%) or teachers' thinking (3%). The term is also used in other disciplines, particularly medicine and related fields (17%), especially nursing and nurse practitioner education; there are also articles about thinking skills and critical thinking in veterinary medicine and physiotherapy. Fourteen percent of references are about undergraduate or postgraduate education. The term is also used in business education (3%), particularly accountancy and marketing, as well as social work (1%), with a handful of references (1%) in other fields appearing in our search, such as criminology and military education. We noted that there has been a shift in the use of terminology from an initial focus on thinking skills and higher order thinking to an interest in metacognition and at the time of this review on aspects of self-regulation. Many of these reports, however, deal with related aspects of teaching and learning, although practitioner interest in thinking skills has remained strong. The findings from this preliminary sweep about terminology confirmed what was found in a narrative review of thinking skills in post-compulsory education commissioned by the Learning and Skills Research Centre (Moseley *et al.,* 2004).

Systematic map

When the search was refined to focus on teaching and learning in schools in accordance with our research question, of the 191 reports identified as being relevant to the review and likely to contain empirical data, about two-thirds are from the USA (34%) and the UK (27%). Nearly half of these reports were set in secondary schools (45%) and about a third in primary schools (34%). There is a greater proportion of research reported for 11 to 14 year-olds with comparatively little about 5 to 7 year-olds (Key Stage 1 pupils). Most subjects of the curriculum are represented in these reports, although a majority are in the core areas of science (34%), literacy (20%) and mathematics (19%). The majority of reports contain data on pupil attainment, with just less than a quarter having data on pupil attitudes or beliefs, and about an eighth with data on teachers' attitudes or beliefs. For most of the reports of pupil attainment, the data are quantitative (102 out of 136). On the other hand, about a third of the reports with data on teacher attitudes or beliefs contain qualitative data only (13 out of 40).

In-depth review and synthesis

The focus was further refined for the review by selecting studies that include quantitative and qualitative data and which meet the criterion for there being a researcher-manipulated evaluation. This produced a further subset of 23 studies. The synthesis of evidence from these studies indicates the following:

- The majority of studies report positive impact on pupils' attainment across a range of non-curriculum measures (such as reasoning or problem-solving). No studies report negative impact on such measures.
- Approximately half of the studies show immediate, positive impact on learning on curricular measures of attainment (where such measures were used).
- There is some evidence that pupils can apply or translate this learning to other contexts.
- Where there is either no, or small immediate, impact on curriculum measures such improvement may appear later or increase over time.
- The impact of thinking skills approaches may not be even across all groups of pupils.
- There is some evidence that there may be greater impact on low attaining pupils, particularly when using metacognitive strategies.
- There is some evidence that pupils benefit from explicit training in the use of thinking skills strategies and approaches.
- Some of the benefits of thinking skills programmes and approaches derive from making thinking and reasoning explicit through a pedagogical emphasis on classroom talk and interaction.
- The role of the teacher is important in thinking skills programmes and approaches in establishing collaborative group work, effective patterns of talk and in eliciting pupils' responses.

Conclusions

Strengths

The review has brought some structure and order to a previously disparate field of enquiry, providing practitioners and other users of research with a map of where and how the impact of thinking skills on teaching and learning has been investigated. The map indicates which phases of schooling and which subject disciplines can draw upon research evidence on thinking skills; the in-depth review and synthesis present the evidence for the impact of thinking skills interventions on pupils in authentic classroom settings, as advocated by McGuinness (1999) in her descriptive overview of research in the field. Users of research and researchers now have a clear picture of the weight of evidence to support the use of thinking skills to improve pupil learning, the gaps in knowledge and deficiencies of research in the field.

Limitations

In the main, studies were accessed through journal articles and these did not always provide all the details needed to make an appraisal of how applicable the findings would be to other educational settings or their 'ecological validity' (Gall, Borg and Gall, 1996). Details - such as the selection of the schools involved and the specific training given to teachers implementing the intervention - were either not included or were sketchy. We were conscious of a disparity between the requirements to write up research for refereed journals and ensuring that research processes are sufficiently transparent to encourage practitioners to evaluate the significance of a study for their own practice. We found the reporting of qualitative evidence, in particular, to be limited in terms of providing detailed explanations of how the data informing the conclusions were analysed.

To some degree the strength of the review in mapping the parameters of the use of the term 'thinking skills' can also be seen as a weakness. The inclusion of studies linked to keywords with wider provenance under our broader definition means that outcomes highlighted in our synthesis could be claimed to result from collaborative learning, for example, rather than thinking skills per se. However, it has enabled us to indicate where the evidence relates to the use of specific cognitive strategies and where it is concerned with outcomes linked to classroom climate and a particular style of interaction.

Despite our best efforts, there will be studies that we did not find and so did not include in the review. The decision to limit our search to studies published in English is, of course, an obvious limitation but one that we could not overcome, given the scope, scale and funding of the review.

Implications for policy, practice and research

Whilst thinking skills programmes and approaches have a positive impact on pupils' attainment, such impact is not always consistent. The evidence from this review suggests that there is a need to select interventions carefully and to be prepared to persist with an intervention, as it may not always provide improvement on curricular measures in the short-term. Research also indicates that the causes of improvement in pupil learning are complex and a more general emphasis on making aspects of teaching and learning explicit in classrooms (particularly in terms of making reasoning explicit) may have similar benefits to those obtained through a particular programme. Further research across a wider range of subjects and age groups would be particularly useful, as would (i) comparative research to evaluate the relative benefits of different thinking skills programmes and approaches, and (ii) a comparison of such approaches with other educational interventions.

- **Policy-makers** - Provision of guidelines for the implementation and evaluation of thinking skills in classrooms based on research evidence would enable schools to make informed choices. Access to information, in order to make links between thinking skills programmes and what is known about effective teaching and learning and national policy initiatives, could be facilitated. Research could be commissioned to establish which thinking skills interventions are effective, efficient and cost-effective.

- **Practitioners** - When introducing interventions that focus on improving specific cognitive strategies, it could be more efficient to target particular groups of pupils and identify the most appropriate times for development. Interventions aimed at developing a classroom ethos conducive to making learning more explicit and fostering dialogue about teaching and learning, on the other hand, can be promoted at any time. Positive outcomes on pupil motivation and self-esteem may be registered before there is any tangible impact on attainment measured by standard assessments. There may be a delay of as much as two years in the appearance of improved attainment in tests and exams, and consequently it may be difficult to distinguish between the impact of the intervention and the effect of any subsequent teaching.

- **Researchers** – Further work is needed on identifying efficient, as well as effective, ways of intervening to promote thinking skills and raise attainment. There is a clear need for more comparative studies between

different types of intervention, and between thinking skills approaches and other strategies designed to change patterns of classroom interaction. The descriptive map shows where there are gaps in the research evidence. The in-depth review indicates where aspects of methodology and the reporting of findings could be more robust and accessible to other researchers as well as to other users of the findings.

1. BACKGROUND

1.1 Aims and rationale for current review

Our interpretation of the research literature in the field of the impact of thinking skills interventions and approaches 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, such as formative assessment and feedback in classrooms (Black and Wiliam, 1998; Torrance and Pryor, 1998) or classroom talk and interaction (Galton *et al.,* 1999; Mercer, 1995). Our intention in this systematic review was to maintain the focus on the relationship between the general characteristics of thinking skills approaches and interventions, and any impact on teaching and learning.

The review set out to provide information for a range of audiences. The descriptive map and keywording of studies provide a valuable overview of the field for practitioners (particularly those interested in, or undertaking, classroom-based research), educational researchers themselves and those who fund research. The findings from this section of the review should enable these groups to find relevant studies about the implementation of thinking skills programmes and approaches which contain evidence of impact. The in-depth review and synthesis can inform these audiences and summarise findings for a wider group of policy-makers and practitioners, with implications drawn out for policy and practice. The findings from the synthesis could also be of benefit to the wider educational community, especially parents and learners, in providing information about effective educational interventions and approaches.

1.2 Definitional and conceptual issues

The teaching of thinking skills is an explicit part of the National Curriculum in England 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) provides an overview of current research into the teaching of thinking skills and builds on the work of earlier reviews in this area. Nisbet and Davies (1990) list 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 (United Kingdom) 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 broader aspects of pedagogy. Our working definition for the purposes of this review is that thinking skills interventions are 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 characterised as approaches or programmes which:

- require learners to articulate and evaluate specific learning approaches

- 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 consists of six major categories arranged in the following order: *knowledge, comprehension*, *application*, *analysis*, *synthesis* and *evaluation*. The relationship among the categories along the continuum was presumed to constitute a cumulative hierarchy.

Appendix A of the Review Group's proposal[1] contains a discussion of some of the issues surrounding a definition of the term. With the focus on thinking skills in the curriculum in England at the time of this review, commercial interest in promoting specific programmes has created the need for teachers to have access to reliable information about the scope and impact of particular approaches for all pupils.

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) and Somerset thinking skills (Blagg, Ballinger and Gardner, 1988). Nickerson, Perkins and Smith (1985) attempted to impose a structure on these programmes by classifying them into five categories, a classification accepted by Garnham and Oakhill (1994), and Hamers *et al.* (1999), although the former authors accepted that these were only broad categories.

Cognitive operations

Programmes in this category stress the need for certain basic skills, such as classification or seriation. The obvious exemplar here is Feuerstein's instrumental enrichment (FIE). IE sets out to foster the development of what are considered to be crucial underlying skills, such as comparing, classifying and clear perception. Such skills are often thought to be missing or poorly developed in children, on account of inadequate early experiences. Feuerstein's ideas are generally acknowledged to be seminal in this area. They have directly inspired several other programmes, notably, in this country, the Somerset Thinking Skills Course (Blagg, Ballinger and Gardner, 1988), a series of generic thinking programmes, aimed at the secondary age-level, and Top Ten Thinking Tactics (Lake and Needham, 1993) aimed at primary children.

Heuristics (strategies)

The essential feature of this approach is task analysis, where a complex task is split up into more manageable chunks. Although his Cognitive Research Trust (CoRT) materials are not currently published in Britain, and are not as frequently used in British schools as several other programmes, the name of Edward de Bono is probably the one which more British people would associate with thinking

[1] http://eppi.ioe.ac.uk/EPPIWeb/home.aspx?page=/reel/review_groups/thinking_skills/home.htm

skills than any other. Throughout his writings (e.g. 1970, 1992), de Bono stresses the importance of consciously practising certain strategies in order to become a more effective thinker. His CoRT materials refer to 'thinking tools', which are made easy for children to remember, with mnemonic titles such as PMI, standing for Plus, Minus, Interesting - urging the student not to rush into a critical decision, but first to list all the things which are in favour of the idea, those which militate against it and those which are interesting, irrespective of critical orientation.

Formal thinking

In the formal thinking approach, Piaget's stage theory of development underpins the emphasis on helping pupils to make the transition from concrete to formal operational thinking. Examples of this approach would be operational enrichment (OE) (Csapó 1992) or cognitive acceleration through science education (CASE) (Adey, Shayer and Yates, 1995), in the teaching of science for secondary-age pupils, although it also uses principles from Feuerstein. CASE has developed into other curriculum areas (Shayer and Adey, 2002), such as mathematics and technology education (cognitive acceleration through mathematics education (CAME) and cognitive acceleration through technology education (CATE)), as well as for use with younger pupils in science (Let's Think).

Thinking as manipulation of language and symbols

Socio-cultural or socio-historical approaches have also influenced thinking skills programmes and approaches. Drawing on the work of the Russian psychologist, Lev Vygotsky, the emphasis is on talk, discussion and 'scaffolded' experiences where children develop understanding through communicating their ideas. The Thinking Together programme developed by a team at the Open University (Dawes, Mercer and Wegerif, 2000) draws explicitly on these ideas.

Thinking about thinking: metacognition

In thinking about thinking or metacognition it is assumed that improving understanding of one's own thinking will improve subsequent thinking. Nickerson *et al.* (1985) include in this category those programmes which focus on thinking as their subject matter. Foremost in this category is the work of the American philosopher, Matthew Lipman. His Philosophy for Children programme (e.g. Lipman, Sharp and Oscanyan, 1980) rests on certain assumptions, such as that discussion skills usually precede and form the basis of thinking skills (rather than the other way round). Through engaging in group dialogue in an open spirit of enquiry, in what is known as a 'community of enquiry', children can become more effective thinkers as they practise thinking about their thinking processes (Lipman, 1991, 2003). There are several other programmes based on the 'community of enquiry' approach, such as Karin Murris' 'Teaching Philosophy with Picture Books' (which has recently been republished as Storywise (Murris and Haynes, 2001)), or Robert Fisher's work (Fisher 1996, 1998; see also the website for the Society for the Advancement of Philosophical Enquiry and Reflection in Education (SAPERE)). An interest in a philosophical approach, as opposed to a psychological one, tends to predominate in this area.

Integration and infusion

There has been recent interest in 'infused' approaches which seek to develop teachers' pedagogy at the same time as make learners' thinking explicit. Infusion and the use of pedagogical strategies (Leat and Higgins, 2002; McGuinness *et al.*, 1995; McGuinness, 1999) tend to blend aspects of thinking skills programmes which makes classification into precise sub-categories challenging.

1.3 Policy and practice background

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, 1991/2003), '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 (Galton *et al*., 1999; Webb and Farrivar, 1994). This work has been used in research and development work with trainee and practising teachers as a means by which teachers could put into practice or 'enact' findings from educational research (Higgins, 2001; Higgins and Moseley, 2002; Leat and Higgins, 2002). In England, thinking skills approaches have been influential in the development of the National Curriculum (McGuinness, 1999) and the development of the Key Stage 3 national strategy.

1.4 Research background

There is a range of research evidence about whether thinking skills approaches and the underpinning metacognitive techniques are effective in raising pupils' attainment (e.g. Adey and Shayer, 1994; Romney and Samuels, 2001; Rosenshine and Meister, 1994; Wong *et al*., 1985). The descriptive reviews by McGuinness (1999) for the Department for Education and Employment and by Wilson (2000) for the Scottish Executive provide an overview of recent research into the teaching of thinking skills relevant to the UK. Hamers and Overtoom (1997) and Hamers and Van Luit (1999) summarise recent research in continental Europe. Most of these reviews are not explicit about their scope and scale.

Several meta-analyses have been conducted in the field of thinking skills. These techniques make it possible to compare the impact of different types of educational interventions using statistical techniques involving effect sizes. A few of these focus on specific approaches rather than general thinking skills; this review hopes to provide some answers about thinking skills as a broad approach. This will also help to place thinking skills approaches in the broader context of educational research. Two notable studies have looked at a range of interventions and included aspects of thinking skills as part of their reviews. Hattie *et al.* (1996) evaluated the effects of a wide range of learning skills interventions on students' learning. A meta-analysis by Marzano (1998) was even broader in scope and larger in scale. It categorised all studies in terms of type of intervention and area of thinking affected, although it should be noted that it has not been subject to peer review but is available as a report posted on the Internet.

Both studies are consistent in finding that techniques designed to be used by students led to significantly better results than for those designed to be used by teachers. Although there was enormous diversity in the intervention studies selected by Marzano, ranging from a focus on specific skills (such as memorisation) to the use of disposition-monitoring strategies, he made the following claim about the importance of making aspects of thinking explicit, or metacognition:

> 'Instructional techniques that employed the metacognitive system had strong effects whether they were intended to enhance the knowledge domains, the mental process within the cognitive system, the beliefs and processes within the self-system, or the processes within the metacognitive system itself' (p 127).

Overall, Marzano found that interventions which engage either the 'self system' or the 'metacognitive system' lead to better knowledge outcomes (by six and five percentile points respectively) than those which are directed only at the use of cognitive skills. Nevertheless, there are some types of very effective intervention at the cognitive skill level. These are interventions which address experimental enquiry, using analogies, comparing and contrasting, idea representation and the storage and retrieval of knowledge.

We can summarise these two meta-analyses by saying that there is powerful empirical evidence that thinking skills interventions can be very effective at all levels, but especially if they are directed at metacognitive and self-regulatory approaches. In particular, their effectiveness is likely to be greater if they are used for learner-self-regulation rather than being too closely directed by the teachers. However, in addition to this, certain types of thinking skills interventions can also be very effective if well focused at the cognitive level. These include experimental enquiry, idea representation, and using cues and questions to aid retrieval. These reviews, however, were quite broad and do not focus on teaching and learning in schools; neither do they report on the details of the implementation of the interventions, which is of particular value to both researcher and teachers wishing to replicate the effects.

There are only a few meta-analyses of the impact of specific thinking skills programmes and approaches. Our search strategy did not include identifying reviews (other than as a source of references) and such reviews were excluded before the keywording stage unless they included original empirical data. This summary of the research background may therefore have omitted studies which could be relevant. We intend to address this omission in subsequent reviews.

An example of a meta-analysis of an approach which fits our broader definition of thinking skills is Rosenshine and Meister's (1994) review of reciprocal teaching. This is a teaching approach which features 'guided practice in applying simple concrete strategies to the task of text comprehension' (Brown and Palincsar, 1989). It includes cognitive techniques, such as summarisation, question generation, clarification and prediction, supported through dialogue between teacher and students (or students and students) as they attempt to gain meaning from a text. Rosenshine and Meister's review includes 16 studies of reciprocal teaching and found an average effect size of 0.32 when the impact of the intervention was measured using standardised tests, and an average effect size of 0.88 when more specific tests developed by the researcher were used.

An example of a meta-analysis of a specific thinking skills programme is Romney and Samuels' (2001) review of Feuerstein's instrumental enrichment (FIE). Proponents of FIE claim that the programme results in an improvement in school achievement, cognitive ability and classroom behaviour. However, because some outcome studies have produced negative results, Romney and Samuels undertook a meta-analysis in order to provide a more reliable and comprehensive assessment of the efficacy of FIE. A total of 40 controlled studies, comprising 47 different samples, were examined. Significant, although modest, average effect sizes were found in all three areas - achievement, ability, and behaviour - with the most extensive improvement being made in ability. Gains in spatial/perceptual

ability were related to the length of the intervention (number of hours) and self-esteem was related to age, with older children showing increases and younger children showing decreases. This provides powerful evidence for those considering using such a programme in schools. Not only does it suggest that gains, although modest, are likely to be achieved when using FIE, it also provides some pointers about how to implement the programme and some implications about attending to the impact on self-esteem of younger learners in particular. The average effect sizes are lower that those found in other programmes, but that might also be explained by the predominant use of FIE with pupils with special needs.

In this review, we augment previous reviews by including a wider field of studies and identifying empirical investigations with ecological validity for closer analysis. These studies have then been subjected to a systematic and rigorous review procedure that enables users to form a judgement about the relevance or applicability of the evidence of the impact of thinking skills approaches to their own context.

1.5 Authors, funders and other users of the review

The funding for this review was provided by DfES through the EPPI-Centre, with additional costs being met by the Thinking Skills Research Centre at Newcastle University. The review set out to provide information for a range of audiences. The descriptive map and keywording of studies provide a valuable overview of the field for practitioners (particularly those interested in, or undertaking, classroom-based research), educational researchers themselves and those who fund research. The findings from this phase of the review should enable these groups to find relevant studies about the implementation of thinking skills programmes and approaches which contain evidence of impact. The in-depth review and synthesis are intended to inform these audiences and summarise findings for policy-makers and practitioners more widely, with implications drawn out for policy and practice.

A systematic review was therefore needed:

- to evaluate the empirical evidence highlighted in the descriptive reviews recently undertaken (McGuinness, 1999; Wilson 2000)
- to compare the findings of earlier reviews which have attempted to evaluate evidence from a range of thinking skills approaches (e.g. Sternberg and Bhana, 1986)
- to extend the scope of existing systematic reviews which have focused only on a particular programme or approach (such as the recent meta-analysis of evidence of the impact on learners of Feuerstein's instrumental enrichment by Romney and Samuels' (2001))
- to set the evidence of the impact of thinking skills programmes and approaches within the wider context of the impact of educational interventions designed to make the processes of learning more explicit to teachers and learners

1.6 Review questions

The aim of the review is to investigate the impact of thinking skills interventions on teaching and learning in classrooms.

The main question for the review is:

What is the impact of the implementation of thinking skills interventions on teaching and learning?

In order to address the question and identify keywords for the search, two further questions were considered:

What are the parameters for defining a particular pedagogy or curriculum development as a thinking skills approach?

How closely do the findings of the studies of the impact of thinking skills approaches correlate with current understanding of effective teaching and learning?

The Review Group therefore adopted a three-stage process, beginning with a survey to identify how the term is used in educational research literature. This was followed by systematic identification of relevant studies. We undertook a mapping exercise of included studies, followed by an in-depth review of a focused sample of studies. The reason for the focused sample is the potential breadth of the field of inquiry, the requirement to identify and include studies systematically and to apply criteria assessing the quality of studies in the final stage.

In Stage 1, the Review Group undertook to explore how, and in which fields, the term 'thinking skills' is used in the educational research literature. We wanted to ascertain which interventions and classroom approaches can be categorised as 'thinking skills' approaches and begin to define the parameters of the term. An overview of the literature identified from key search terms indicated the main areas where thinking skills approaches are used and evaluated (e.g. higher education, in compulsory schooling, nurse education, etc.). This also helped to develop a clearer understanding of the way that terminology like 'thinking skills' is used and which related terms are used to describe similar approaches.

In Stage 2, the Review Group addressed the main review question:

What is the impact of the implementation of thinking skills interventions on teaching and learning?

This was further refined in the data extraction phase to focus on the impact of learners in schools and to consider those studies for which both qualitative and quantitative data were available.

The Review Group intends to pursue related questions in subsequent reviews such as:

- What is the evidence for impact on teachers?
- Is there evidence that some approaches are more effective than others (such as infusion compared with enrichment)?

- Is there evidence of differential impact on different groups of learners (such as by level of prior attainment, by gender, etc.)?

Finally, we wanted to contextualise the findings of the review in the light of the broader field of educational research and the following issues were discussed by the Review Group:

- How do the findings of studies of the implementation of thinking skills relate to current understanding of effective teaching and learning?
- Do the findings from the systematic review confirm or extend our theoretical understanding of effective teaching? How do the findings of studies of the implementation of thinking skills relate to current understanding of effective teaching and learning? How do the findings of studies of the implementation of thinking skills relate to current understanding of effective teaching and learning?
- Do the findings from the systematic review confirm or extend our empirical understanding of effective learning? How do the findings of studies of the implementation of thinking skills relate to current understanding of effective teaching and learning?
- How is this evidence consistent with the wider research literature on teaching and learning?

2. METHODS USED IN THE REVIEW

2.1 User involvement

2.1.1 Approach and rationale

The Review Group aimed to include representatives from key constituencies of users, such as practitioners from primary and secondary schools, LEA advisers and the research community. It was also important to establish links across the range of thinking skills approaches and we were keen to involve people who had experience of a range of interventions, either as practitioners or as researchers. Two members of the core group were teacher researchers who had considerable experience of implementing and evaluating thinking skills approaches and were also familiar with research methods. We were also able to utilise existing links with research centres in Singapore and Hong Kong and thereby access an international perspective.

2.1.2 Methods used

The scope of the review was initially identified by members of the Thinking Skills Research Centre at Newcastle University, consolidated through discussion with members of the core Review Group and Consultation and Critical Advisory Panel, and refined in the light of comments from referees and EPPI-Centre staff.

Meetings were planned out for the year of the review and timed to follow on from EPPI-Centre training sessions held in London. The meetings were open to anyone who was able to attend and representation of key users was ensured through the core group. The training was shared across the group and meetings focused on each stage of the EPPI-Centre process as applied to our review question. In addition to the meetings, wider consultation was carried out via email and minutes of meetings, including drafts of each stage of the implementation of the protocol, were circulated widely for comment.

2.2 Identifying and describing studies

2.2.1 Defining relevant studies: inclusion and exclusion criteria

The review focused on studies which explicitly and implicitly evaluated the implementation of thinking skills programmes and approaches in classrooms during the ages of compulsory schooling (5 to 16) in all areas of the curriculum. The Review Group found it essential to define both the inclusion and corresponding exclusion criteria explicitly; the full set of inclusion and exclusion criteria can be found in Appendix 2.1.

2.2.2 Identification of potential studies: search strategy

The search terms were agreed by the core Review Group through a series of meetings that looked at definitions of thinking skills and we developed exercises loosely based on personal construct theory to establish key terms and linked terms that were seen to be relevant by the members of the group. The terms selected were then circulated to the Advisory Panel for comment and amendment. The terms were consistently applied to all the databases (see Appendix 2.2 for further details). Terms were applied either individually, or in combination, depending on the specific search interface available. The date range was determined by the database (see Appendix 2.2).

Studies were identified from the following sources:

- Bibliographic databases
- Citation searches of key authors/papers
- Reference lists of key authors/papers
- References on key websites
- Personal contacts
- Direct requests to key informants

The review primarily identified studies through articles published in peer-reviewed academic journals and unpublished materials (such as conference papers or LEA evaluation studies or dissertations) where the focus of the study was on the implementation and evaluation of thinking skills approaches, or programmes in classrooms.

2.2.3 Screening studies: applying inclusion and exclusion criteria

All the citations identified in the initial searches were subjected to the inclusion criteria, which were applied to the titles and abstracts, or full studies if the abstract was not sufficiently clear. Studies were excluded if they failed to meet any one of the inclusion criteria as they were applied in sequence from 1 to 5. Where there was any doubt, studies were included.

2.2.4 Characterising included studies

Reports which met the inclusion criteria were then keyworded using two coding tools: the EPPI-Centre Core Keywording Strategy, version 0.9.5 (EPPI-Centre, 2002a) and the review specific keywords. The EPPI-Centre keywords contain terms relevant to educational research more generally (such as phase of education, curriculum focus and educational setting). The review-specific keywords were developed by the core Review Group with support from members of the Advisory Panel and contain more detailed terms relevant to aspects of teaching and learning in schools (such as age of pupils and terms relevant to thinking skills approaches and interventions).

2.2.5 Identifying and describing studies: quality assurance process

The core Review Group moderated the use of the inclusion and exclusion criteria through meetings where members worked in pairs to apply the criteria on a sample of abstracts and full studies.

Two core Review Group meetings focused on keywording reports. At the first meeting, two reports were keyworded individually, using the EPPI-Centre and draft review specific keywords, and the results were discussed. Then the Review Group worked in pairs to apply the keywords to eight other reports to test the robustness of the review-specific terms. A selection of reports was then keyworded individually before the next meeting. At this second meeting, guidelines were drawn up for application of review-specific terms.

At the next stage, 24 reports were each keyworded by two members of the core Review Group. Subsequent batches were keyworded individually when analysis of paired keywording revealed a high level of reliability. A sample of 10 reports was also keyworded by EPPI-Centre staff (5%). This gave an overall sample of 34 reports out of 191 keyworded by two people (17%); this figure does not include the initial 10 reports which were collaboratively keyworded in the Review Group meetings. Our initial intention had been that each study would be read by at least two members of the core Review Group, but this proved impractical because of the number of reports included and identified for keywording.

2.3 In-depth review

2.3.1 Moving from broad characterisation (mapping) to in-depth review

During the course of the systematic mapping, it became clear the there is a large number of studies in the field of thinking skills. The team therefore chose to focus on studies where there were both quantitative and qualitative data on evidence of impact on pupil attainment. An additional further refinement of the studies for in-depth review was agreed on the basis of the study type, where the study type met the criteria for a 'researcher manipulated' evaluation: that is, there was an attempt on the part of the researcher(s) to change people's experience and as a consequence have control over which groups of people are 'introduced' or 'exposed' to the thinking skills intervention or approach.

Studies selected for in-depth review therefore:

- met the criteria for inclusion in the review (Appendix 2.1)
- contain qualitative as well as quantitative data on evidence of impact on pupil attainment
- reported that there was some control over which groups of pupils received thinking skills intervention

2.3.2 Detailed description of studies in the in-depth review: EPPI-Centre and review-specific data extraction

Detailed description was completed according to the EPPI-Centre Guidelines for Extracting Data (EPPI-Centre, 2002b) and the software supported the reviewers in making a careful analysis of the content of the studies. We were particularly concerned with the ecological validity of the studies. Therefore questions regarding how the intervention was implemented were important and cross-referenced with our review-specific keywords. Studies were included at this stage if they had both quantitative and qualitative data, so that both impact on attainment and a richer description of the impact of the programme could be evaluated. Methodologically studies were included if they reported a research design that involved the researcher in planning or controlling the implementation

to provide opportunities for before and after evaluation with comparison or control groups.

2.3.3 Assessing quality of studies and weight of evidence for the review question

The EPPI-Centre guidelines for assessing the quality of studies require the weight of evidence to be judged both according to the internal validity and reliability of each study, and external or ecological validity in terms of the value for our particular review.

- Weight of evidence A refers to the internal consistency of the study in the sense of the extent to which a study is carried out according to accepted practice for the method adopted, or can the reported findings be trusted in answering the study question?
- Weight of evidence B is concerned with the appropriateness or applicability of the research design for our review question.
- Weight of evidence C is concerned with the focus of the study for our review question.
- Weight of evidence D is concerned with the overall weight of evidence when A, B and C are combined.

A, B, C and D are all classified as high, medium or low. The classification of weight of evidence D is determined by the lowest grade given in either A, B or C.

Issues in establishing the weight of evidence often revolved around the transparency of reporting and whether sufficient information was provided in the study to make judgements about aspects of the research (such as fidelity of implementation of the thinking skills programme or approach).

2.3.4 Synthesis of evidence

Information from those studies which addressed the questions was brought together within the conceptual/contextual framework introduced in Chapter 1. For this evaluation, the Review Group wanted to identify how easily the intervention could be undertaken in a normal school setting (or their 'ecological validity'). The following factors were used as criteria to assess this aspect of the review: reporting details about the implementation of the thinking skills programme or approach (such as the number of lessons and the content of thinking skills activities, or how this was fitted into the school curriculum); and who undertook the teaching of the programme and the group size.

2.3.5 In-depth review: quality-assurance process

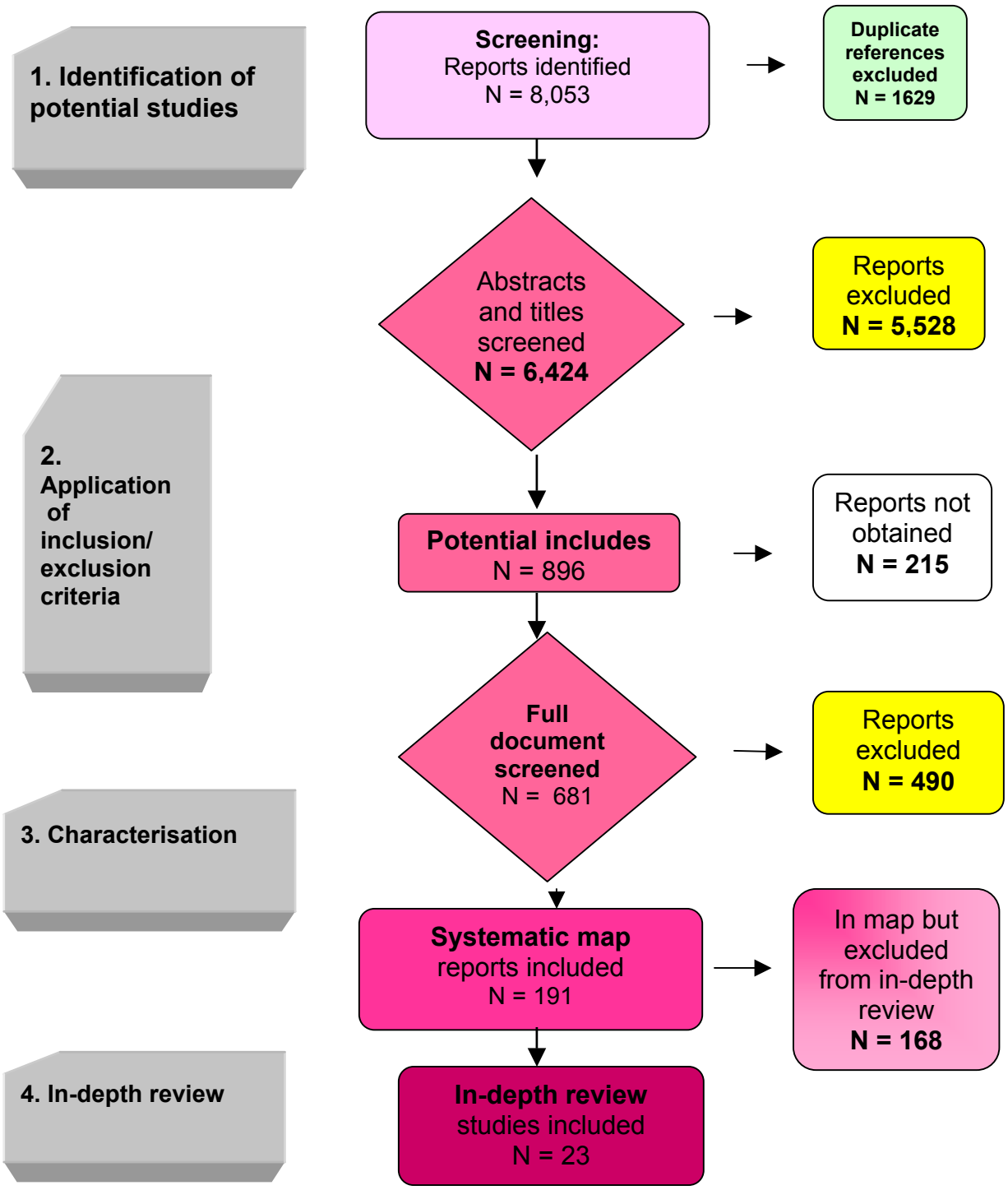
Data were double-entered onto EPPI-Reviewer (a web-based database) by two reviewers working independently. In cases where there was initial disagreement about data-extraction or quality appraisal, this was discussed and resolved. Three members of the EPPI-Centre were involved in this so as to ensure consistency across systematic reviews. Seven of the 23 studies (30%) were evaluated independently by EPPI-Centre staff, then the results compared and any differences discussed and resolved.

The synthesis of evidence was reviewed at a core Review Group meeting where the weighting of evidence was discussed and the relevance of studies for our key question debated.

3. IDENTIFYING AND DESCRIBING STUDIES: RESULTS

3.1 Studies included from searching and screening

Figure 3.1: Filtering of papers from searching to map to synthesis



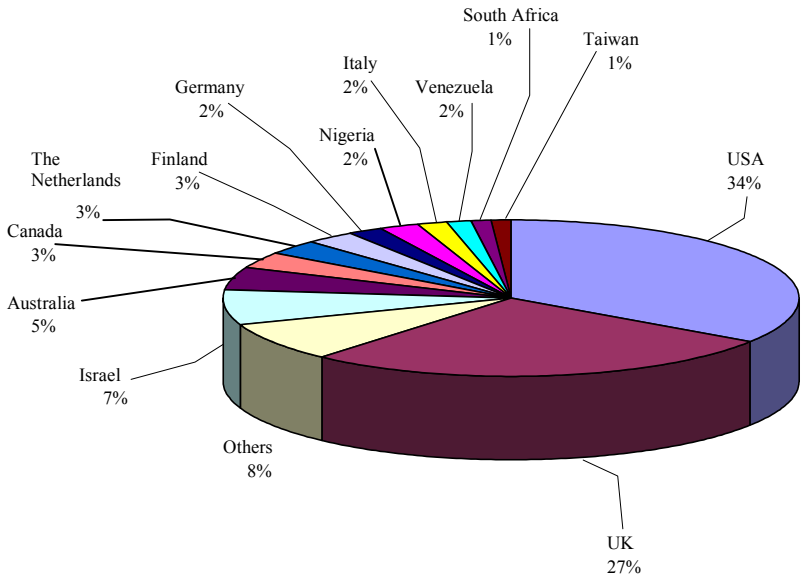
3.2 Characteristics of the included reports (systematic map)

In this section we show how the review progressively narrowed the focus from which studies were taken for the in-depth analysis. The systematic map is based on reports rather than studies. Some studies were described in more than one report and these have therefore been counted twice in the systematic map. This should be taken into account when interpreting the numbers in the figures and tables[1].

3.2.1 Countries represented

The majority of the 191 keyworded reports are from the USA (34%) and the UK (27%). Research undertaken in other countries accounts for the remaining 37%. Fifteen of the reports are from different countries as follows: Belgium, Costa Rica, Cyprus, Estonia, Guam, Hungary, Iceland, Jordan, New Zealand, Pakistan, The Philippines, Saudi Arabia, Singapore, Spain and Turkey. However, it should be noted that this review was limited to reports written in English.

Figure 3.2: International spread of keyworded reports (191 reports keyworded)



3.2.2 Educational setting

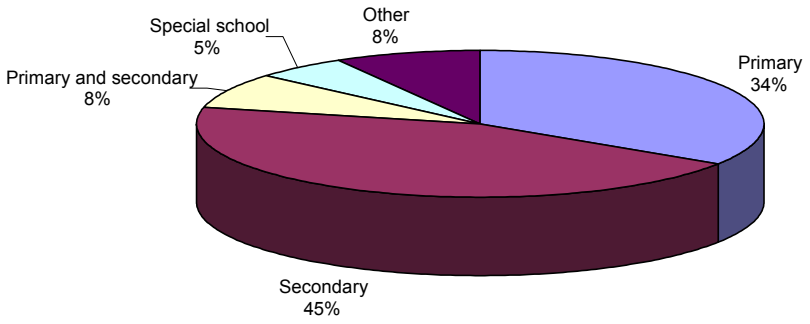
Reports were also keyworded according to the educational setting in which the research took place. Slightly more secondary schools (45%) are represented than primary (34%), with several reports set in both types of schools (8%). A number

[1] This systematic map will be updated in following reviews to be based on studies rather than reports

of reports in special schools are represented (5%) with the remainder (8%) made up of other settings, such as residential schools, community settings or higher education institutions (such as where school-age pupils were taught at a higher education institution in special classes for the purposes of the research), or where the setting is not specified.

Where a specific school type was mentioned, this was recorded, or the age of the pupils was used to categorise the school. Specific issues arose because the types of schools and the age at which pupils change schools varies in different countries and even within countries. The intention of this keywording was to see the range of educational settings represented by the reports. Most countries draw a distinction between primary education (often where one teacher has responsibility for a class) and secondary schooling where subject specialisms dominate the curriculum. We tried to retain this distinction in our keywording, so if pupils in a report were aged 12 but the school was identified by authors as an 'elementary' school, the characteristics of the school were used to determine the educational setting.

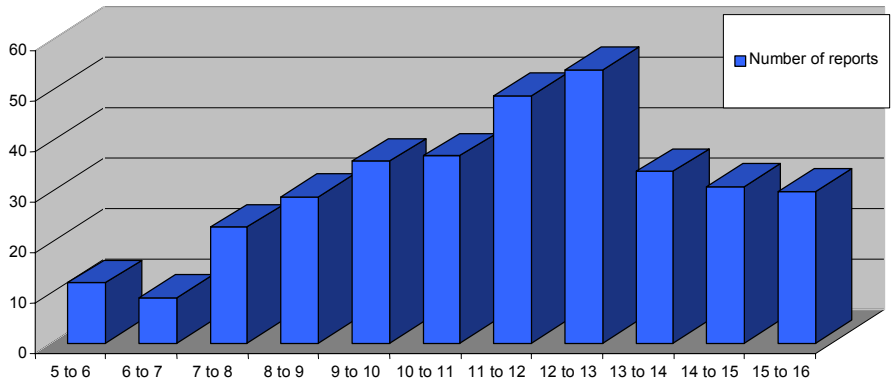
Figure 3.3: Educational setting (191 reports keyworded)



3.2.3 Ages of pupils studied

Reports of research into thinking skills interventions were keyworded according to the ages of pupils in the sample. This broadly reflects the information captured by the educational setting, showing slightly more reports involving secondary pupils. However, it also became apparent that more research is reported for the 11 to 13 age range. Where reports collected data on more than one year group, each year group has been counted separately (total year groups = 344). About half the reports describe working with more than one school year group.

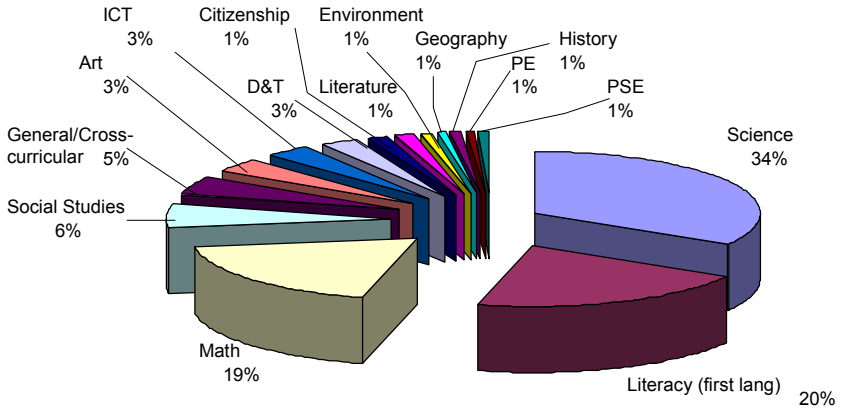
Figure 3.4: Ages of pupils (191 reports keyworded)



3.2.4 Curriculum focus

In terms of areas of the curriculum covered, 122 reports had one or more subjects or areas of the curriculum identified as a focus. Three areas account for nearly three-quarters of the reports: science (34%), literacy (20%) or mathematics (19%). Social studies (when combined with environment, geography and history) is the next largest area[2] (9%) with a few reports covering art, information and communication technology (ICT), design and technology, and other subjects. Most of the reports looking at literacy investigated the impact of thinking skills programmes and approaches on specific aspects of literacy, such as reading comprehension or writing.

Figure 3.5: Curriculum subjects (122 out of 191 reports specified a curriculum focus)

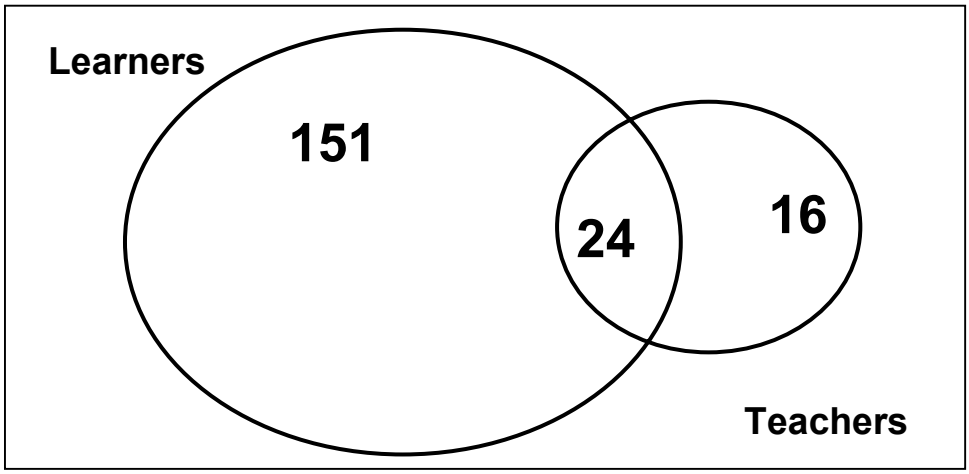


[1] The curriculum focus was coded according to the author's description; however, in some countries (such as the USA) social studies includes subjects such as history and geography.

3.2.5 Impact on learners or teachers?

Some reports investigated the impact of thinking skills programmes and approaches on either learners only (151) or teachers only (16), or both learners and teachers (24).

Figure 3.6: Impact focus (191 reports keyworded – some reported impact on both teachers and learners)



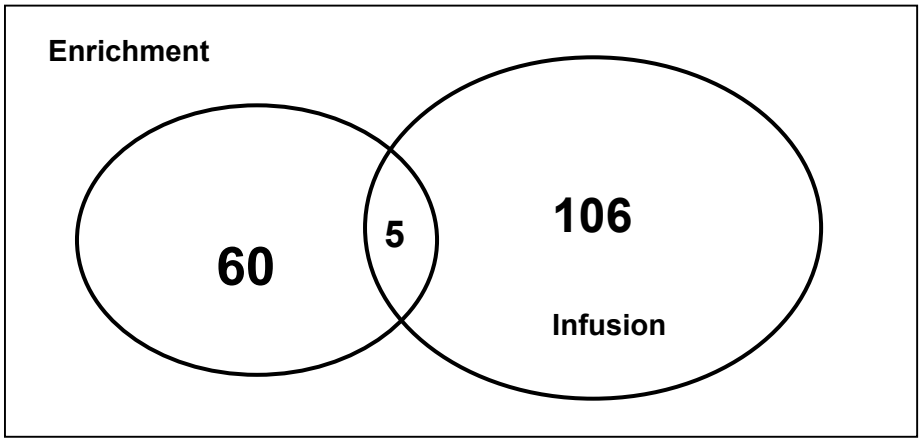
Totals: Learners 151 + 24 = 175; Teachers 16 + 24 = 40

3.2.6 Infusion or enrichment?

There have been debates about which approach to thinking skills is more effective, with current opinion (e.g. McGuinness 1999; Wilson 2000) tending to prefer an 'infusion' approach where specific skills are taught embedded in a meaningful context, rather than through 'enrichment' where thinking skills are taught separately. It is often difficult, however, to categorise programmes or approaches: whether a particular intervention counts through infusion or enrichment depends upon the pedagogical approach adopted.

Where discrete 'thinking skills' lessons were identified and taught as supplementary lessons to the usual curriculum, we classified these as 'enrichment'. If a programme or intervention was taught as part of a particular subject, it was classified as 'infusion'.

Figure 3.7: Infusion and enrichment (171 out of 191 reports were keyworded as infusion or enrichment)

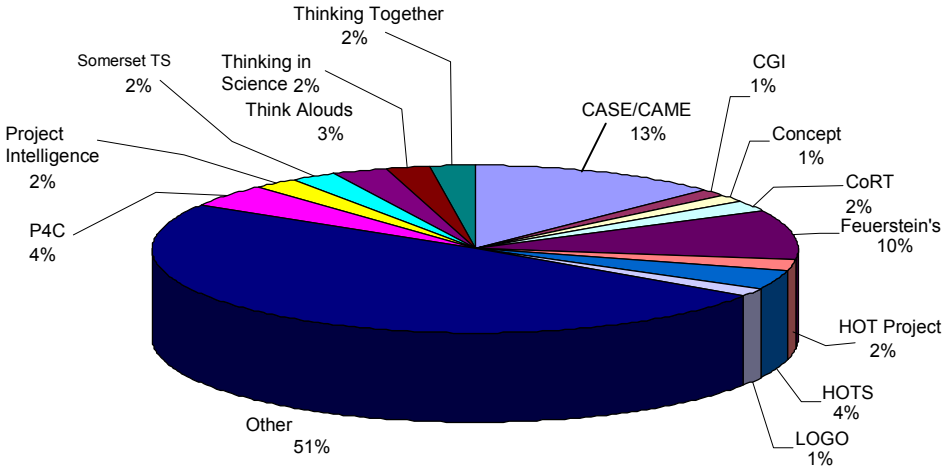


Totals: Enrichment 60 + 5 = 65; Infusion 106 + 5 = 111

3.2.7 Names of programmes and approaches

Just over 70% of the reports keyworded named either the programme or the approach that was being evaluated. Some of these were specific programmes with published materials, such as cognitive acceleration through science education (CASE) or instrumental enrichment (IE) or De Bono's cognitive research trust (CoRT) materials. Others referred to a more general approach, such as concept mapping, cognitively guided instruction (CGI) or the use of 'think-aloud' protocols which had been used in the intervention. Although there were clusters of evaluative reports around well-known programmes, over half the reports keyworded referred to a programme or approach in which that name only occurred once (Figure 3.8), such as Top Ten Thinking Tactics or Thinking Actively in a Social Context (TASC). A couple of reports evaluated more than one programme or approach.

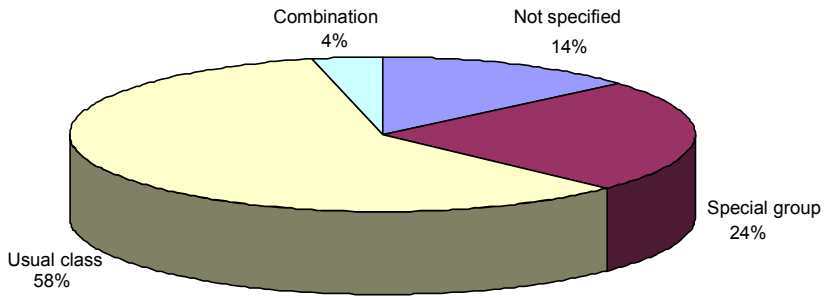
Figure 3.8: Thinking skills programmes by name
(137 out of 191 reports specified a programme name or title)



3.2.8 Class grouping

The method of class organisation was also identified by keywords assigned to reports. These identify whether the intervention was carried out with a special group, or whether it was a normal class environment. For over half the reports, the intervention was carried out with a normal class grouping (whatever was usual for that class). This reflects an aspect of the inclusion criterion for the review, which sought to identify studies which had ecological validity in terms of their applicability to classroom practice, hence studies in schools where the teaching was undertaken by a member of staff were of particular interest.

Figure 3.9: Usual class or special grouping (191 reports)



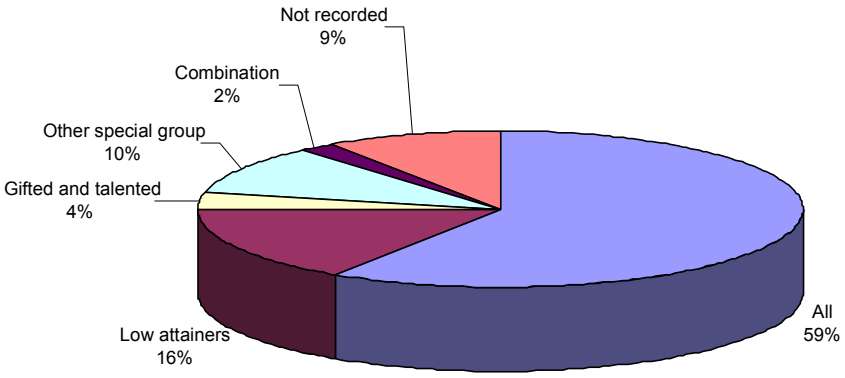
3.2.9 Teaching group size

The size of the teaching group (as opposed to the number in the research sample) was usually not recorded (58% of reports). Where it was recorded, small group sizes (15 or less) formed the largest proportion (40%), followed by medium-sized classes of 16-25 pupils, with a further 27% of reports recording group sizes of over 26 pupils.

3.2.10 Learner focus

The type of pupils who were the focus of the research was also keyworded. This was in order to distinguish between the grouping of the pupils for teaching purposes and the focus for data collection; for example, a group of low attaining pupils could be studied either when they were taught as part of a larger class or when they were withdrawn for separate lessons.

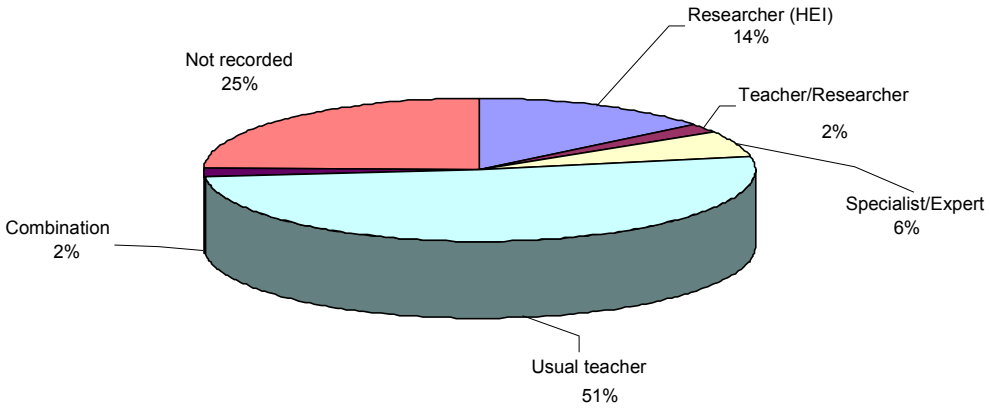
Figure 3.10: Learner focus (191 reports)



3.2.11 Teacher

The professional role of those who undertook the teaching in the different reports was keyworded according to whether they were researchers working in a higher education institution, whether they were teachers as researchers, specialists or experts trained to deliver special lessons, or the usual teacher. In just over half the reports (51%), it was the usual teacher who taught the class.

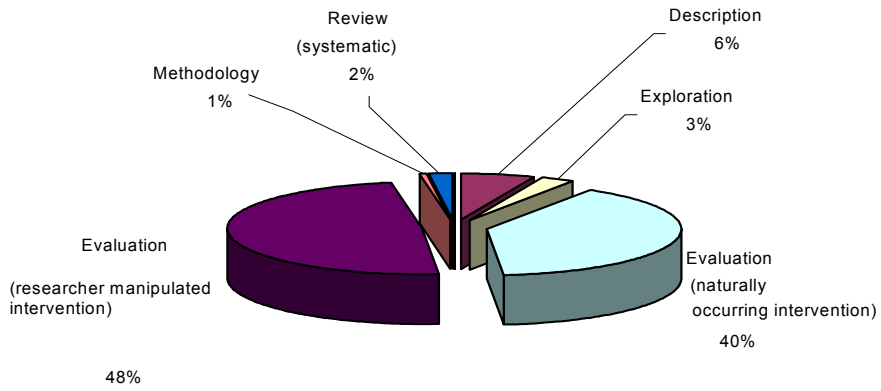
Figure 3.11: Teacher focus (191 reports)



3.2.12 Study type

Eighty-eight percent of reports are evaluations (48% are evaluations of researchers manipulated interventions and 40% evaluations of naturally occurring interventions). Six percent are descriptions and three percent exploration of relationships.

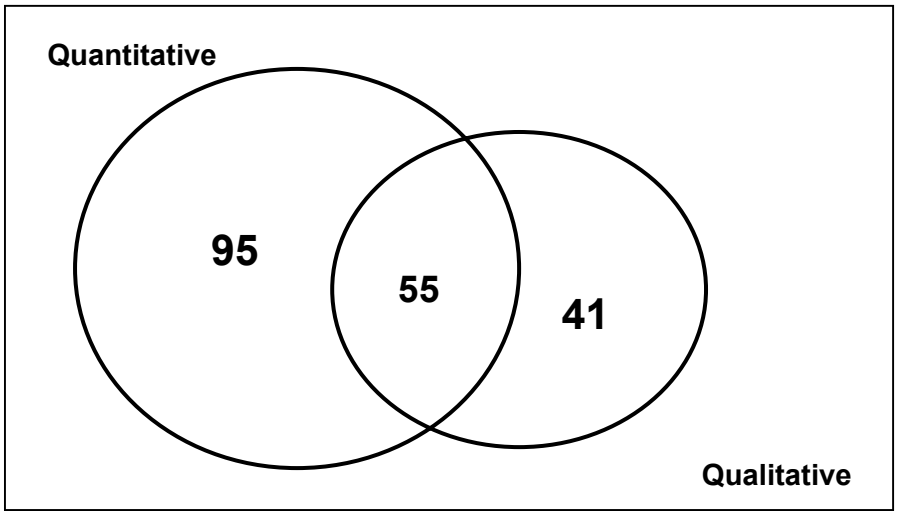
Figure 3.12: Study type (191 reports)



3.2.13 Data

The majority of reports keyworded contain quantitative outcome data, although 41 contain qualitative outcome data only. About a quarter of the reports (55) contain both.

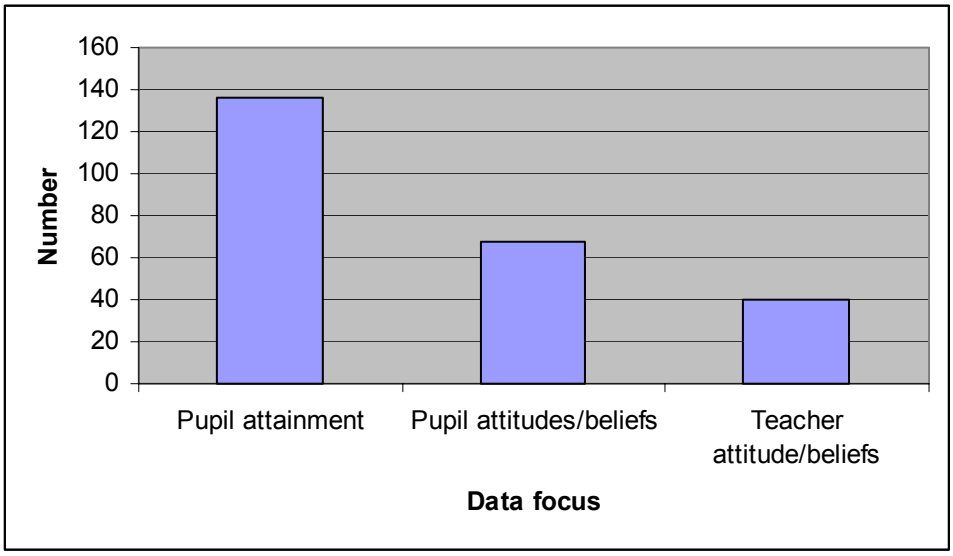
Figure 3.13: Data type (191 reports)



Totals: Quantitative 95 + 55 = 150 Qualitative 41 + 55 = 96

Keywording also identified the focus of the data (Figure 3.14).

Figure 3.14: Data focus (191 reports)



Note: codes not mutually exclusive some reports had more than one focus

The majority of reports contain data on pupil attainment (136/191) with about a third (68/191) having data on pupil attitudes or beliefs and about a fifth (40/191) with data on teachers' attitudes or beliefs. For most of the reports of pupil attainment, the data were quantitative (102/136). On the other hand, about half of the reports with data on teachers' attitudes or beliefs contained qualitative data alone.

3.3 Identifying and describing studies: quality assurance results

The core Review Group worked in pairs to apply inclusion and exclusion criteria to abstracts and full reports. Group meetings were used to review this process so that definitions were applied consistently across the group. The keywording stage was subject to similar quality assurance (QA) procedures and we found a high overall proportion of categories were coded identically (92%). The differences were usually extra coding (e.g. 'teaching staff' by one coder, and 'teaching staff' and 'senior management' by another). The most problematic category was 'Thinking skills terms' (Q17) where there was considerable overlap for one or two terms, but a different third term or use of 'other'.

The next item where most differences were found was 'Research sample' (Q15), where the number of classes, or number of pupils in the intervention and control groups was reported. Differences here were experienced either in combining number of classes for intervention and control, or in different figures reported in different parts of a report (e.g. intended, rather than actual sample, or different numbers of pupils for different outcome measures used).

Occasional differences were found for other categories such as 'Educational setting' (Q9), where identifying a school type across differing school systems is challenging; for example, middle schools in the UK) or where primary schools extend to age 12 or 13.

The iterative nature of the keywording process ensured that there was a high level of consistency in coding the reports across the Review Group and with the EPPI-Centre staff. Overall there was a high level of agreement between coders in the keywording phase of the review.

4. IN-DEPTH REVIEW: RESULTS

4.1 Selecting studies for the in-depth review

During the course of the mapping described above, it became clear that there was a large number of studies in the field of thinking skills. It was clearly not possible to review in-depth all the studies, within the timescale required for the first review. The Review Group therefore needed to identify a sub-set of the included studies for in-depth analysis. The group chose to focus on studies where there was both quantitative and qualitative evidence of impact on pupil attainment. This produced a further subset of 35 studies. However, even at this stage the amount of work required proved to be considerable due to the time required for in-depth review, with each study coded by two members of the group. A further refinement of the studies for in-depth review was therefore agreed on the basis of the study type. Studies needed to meet the criteria for a 'researcher manipulated' evaluation, where there was an attempt on the part of the researcher(s) to change people's experience and as a consequence have control over which groups of people are 'introduced' or 'exposed' to the thinking skills intervention or approach. This produced a further subset of 23 studies which met these criteria.

Studies selected for in-depth review therefore:

- meet the criteria for inclusion in the review
- contain qualitative as well as quantitative evidence of impact on pupil attainment
- report that there was some control over the groups of pupils who received thinking skills intervention so as to create a valid comparison

4.2 Comparing the studies selected for the in-depth review with the reports in the systematic map

The characteristics of studies selected for in-depth review were then compared with the reports in the descriptive map to determine how the in-depth review represented the descriptive mapping of the thinking skills literature reported above.

4.2.1 Country of included studies

The studies in the in-depth review come from a range of countries, with just over half from the US and UK.

Table 4.1: Countries of studies in in-depth review (N=23 studies)

Country	Studies
Australia	2[1]
Belgium	2
Cyprus	1
Germany[2]	1

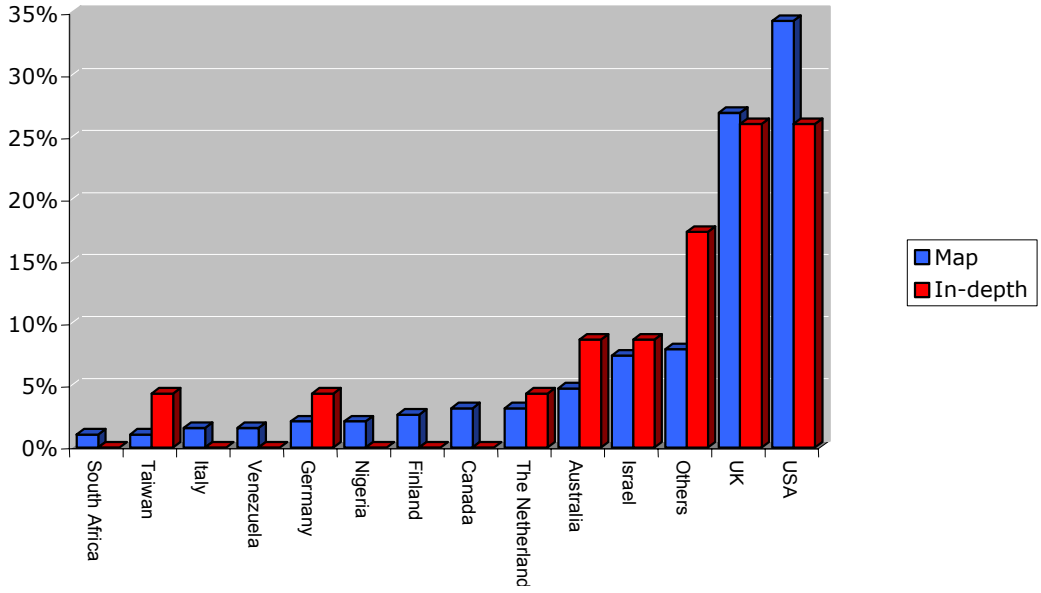
[1] Inference for one study from affiliation of author of the study
[2] The school setting was a British forces school in Germany

Hungary	1
Israel	2
The Netherlands	1
Taiwan	1
UK: England	5
UK: Scotland	1
USA	6[3]

Note: Categories are mutually exclusive

This is similar to the spread identified in the descriptive map.

Figure 4.1: Countries in the descriptive map and the in-depth review (191/23 studies)



4.2.2 Educational setting

The studies in the in-depth review represent a range of educational settings.

Table 4.2 Educational setting of studies in the in-depth review (N = 23 studies)

Educational Setting	Studies
Independent (primary) school	1
Primary school	14
Secondary school	9
Special needs school	2
Total[4]	***26***

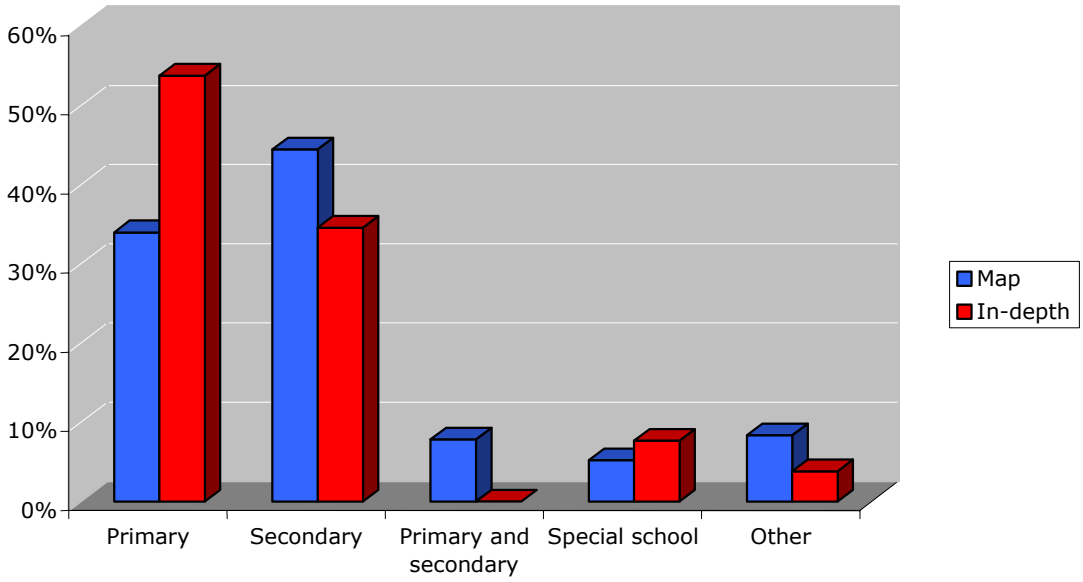
Note: Some studies were coded with more than one educational setting.

[3] Inference for one of these studies from the affiliation of authors of the study

[4] Some studies report data from more than one kind of school.

This can be compared with the settings in the descriptive map. The balance between primary and secondary schools has altered slightly, although overall the representation is similar.

Figure 4.2: Comparison of school setting between descriptive map and in-depth studies (191/23)



4.2.3 Curriculum focus of the studies in the in-depth review

The majority of the studies in the in-depth review focus on science, mathematics and aspects of literacy (mainly reading comprehension). Some studies report on impact on more than one curriculum subject. Subjects in the humanities and arts are not represented.

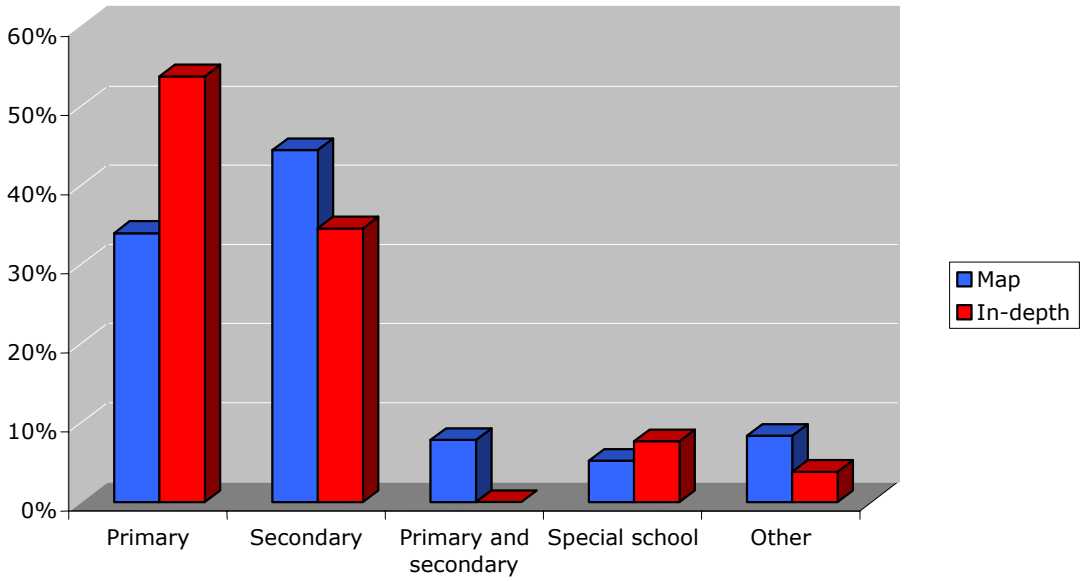
Table 4.3 Curriculum focus of the in-depth studies (N=23 studies)

Curriculum focus	Studies
Citizenship	2
General/Cross-curricular	2
ICT	2
Literacy - first languages	7
Mathematics	7
Science	10
Total	***30***

Note: Some studies were coded with more than one curriculum focus.

This is similar to the distribution of subjects described in the descriptive map.

Figure 4.3: Comparison of curriculum subjects between descriptive map (N = 191) and in-depth studies (N = 23)



4.3 Further details of studies included in the in-depth review

The majority of studies explicitly report that they include pupils of both sexes, although the reporting of some studies did not always refer to the sex of the pupils studied.

Table 4.4: Sex of pupils in the in-depth studies (N=23 studies)

Sex	Studies
Mixed sex	16
Male only	1
Not stated/unclear	6

Note: Study categories are mutually exclusive

4.3.1 Sample size

The distribution of sample size is shown in Table 4.5. The majority of studies analysed for the in-depth review have samples between 11 and 100 pupils.

Table 4.5: Distribution of sample size in the in-depth studies (N=23 studies)

Sample size (experimental and controls)	Studies
0-10	2
11-50	8

51-100	7
101-500	4
500+	2

Note: Study categories are mutually exclusive

4.3.2 Thinking skills programmes

The majority of studies included in the in-depth review are evaluations of named programmes.

Table 4.6: Thinking skills named programmes (N=23 studies)

Named programme	Studies
Yes	17
No	6

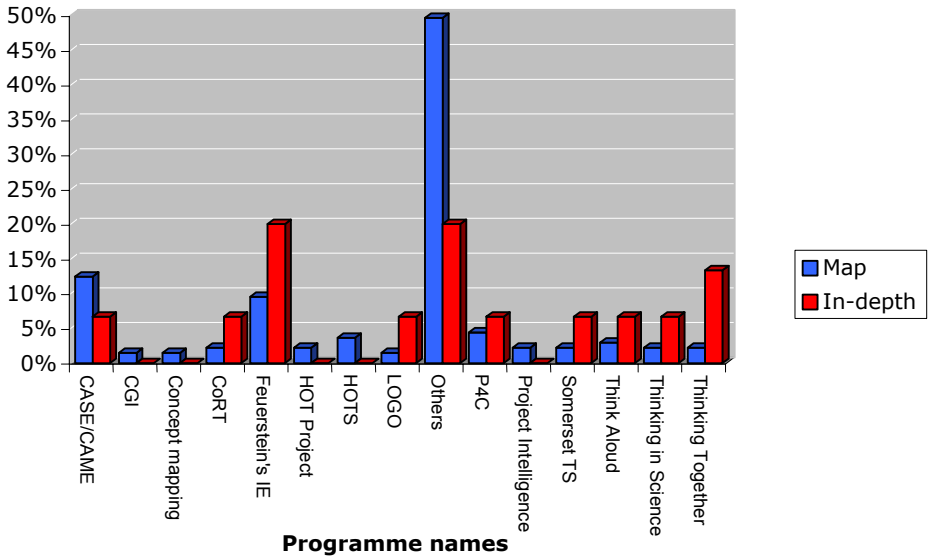
Note: Study categories are mutually exclusive

These programmes are:

- Cognitive acceleration through science education (CASE)
- CoRT (Cognitive Research Trust)
- Feuerstein's instrumental enrichment (three studies)
- LOGO/Metacognition
- Metacognitive learning cycle
- Philosophy for children
- Programme reading comprehension
- Search, solve, create, share (SSCS) (Pizzini *et al.*, 1988)
- 'SOLVE' (systematic analysis; overall planning; linking together; verification)
- Somerset thinking skills
- Think-aloud procedure (Meyers and Lytle, 1986 and Wade 1990)
- Thinking actively in a social context (TASC)
- Thinking in science
- Thinking mathematics project
- Thinking together/TRAC (talk, reasoning and computers)

This representation of named thinking skills programmes can be compared with the range of programmes in the descriptive map. Most of the programmes with three or more studies are represented and the in-depth review can certainly claim to cover the range of programmes identified in the mapping stage.

Figure 4.4: Comparison of thinking skills programmes represented in the descriptive map (137 studies) and the in-depth (17) studies



4.3.3 Research focus

Other aspects of the studies in the descriptive map were also considered. The review identified studies which investigated the impact of thinking skills on learning, so all of the included studies have a teaching and learning focus, and the majority have a specific curriculum focus. A few studies have an additional focus, such as assessment or specific aspects of research methodology.

Table 4.7: Research focus of the in-depth studies (N = 23)

Research focus	Studies
Assessment	2
Curriculum	20
Methodology	1
Teaching and learning	23
Other	1

Note: Some studies were coded with more than one type of focus.

4.3.4 Data-collection methods

A range of types of data-collection was used in the studies with all reporting either on curriculum outcomes or some form of attainment data (such as standardised tests).

Table 4.8: Methods of data-collection in the in-depth studies (N = 23)

Data collection methods	Number
Curriculum-based assessment	16
Group interview	6
One to one interview	8
Observation	14

Self-completion questionnaire	4
Self-completion report or diary	2
Exams	3
Practical test	4
Psychological test	14
School/college records	2
Other documentation	3

Note: All the studies contain more than one type of data.

4.3.5 Timing of test measures

Most of the studies in the in-depth review used a pre- and post-test design. One study used post-test measures only.

4.3.6 Quality of the studies

Most of the studies report clear aims, underpinning rationale and context. However, there are rarely details about the overall sampling strategy (the sampling frame) by which schools were involved in the research. The studies also vary on the extent to which they report on other aspects of the research, such as the classroom processes involved in ensuring that the thinking skills intervention enabled pupils to talk about their thinking ('fidelity of implementation'). The majority of studies contain clear explanations about the quantitative analysis and methods used to ensure the transparency of reporting (such as detailing reliability and validity of measures used, or the method of analysis chosen). Studies were selected because they contain both qualitative and quantitative data as the Review Group considered that studies containing both types of data would be likely to offer a more complete description of the impact of thinking skills approaches in the classroom (such as by including pupils' and teachers' perceptions, for example). However, the quality of reporting of qualitative methodology and analysis was particularly opaque. Most studies did not describe how these data were analysed (or even obtained) in sufficient detail to understand the context and therefore the interpretation. Where qualitative data are included, they therefore tend to be anecdotal.

4.4 Synthesis of evidence

Data from those studies which address the question was brought together and the weight of evidence judgements for each of the studies in the in-depth review are shown in Table 4.4.1. (See section 2.3.3 for detail on how the weight of evidence judgements were made.)

Table 4.4.1: Weight of evidence judgements for studies in the in-depth review

Study	Weight of evidence A	Weight of evidence B	Weight of evidence C	Weight of evidence D
Adey and Shayer (1990)	High	High	High	High
Blank (2000)	Medium	Medium	Medium	Medium
Bowdler *et al.* (1992)	Low	Medium	Medium	Low
Cardelle-Elawar (1992)	Medium	High	High	Medium

Chang and Barufaldi (1999)	High	High	High	High
Csapo (1992)	High	High	High	High
De Corte *et al.* (2001)	High	High	High	High
De Konig and Hamers (1999)	High	High	Medium	Medium
Fields (1995)	Low	Low	Medium	Low
Georghiades (2000)	Medium	High	Medium	Medium
Haywood *et al*. (1988)	Medium	Medium	Medium	Medium
Head and O'Neill (1999)	Low	Medium	Low	Low
Hojnacki and Grover (1992)	Medium	Low	Medium	Low
Kramarski and Mevarech (1997)	Medium	Medium	High	Medium
Maltby (1995)	Low	Low	Low	Low
Mastropieri *et al.* (1997)	Low	Low	Low	Low
Mercer *et al.* (1999)	Low	Low	Medium	Low
Ritchie and Edwards (1996)	High	High	Medium	Medium
Strang and Shayer (1993)	Medium	Medium	High	Medium
Verschaffel *et al.* (1997)	Medium	High	Medium	Medium
Ward and Traweek (1993)	Medium	High	Medium	Medium
Wegerif (1996)	Medium	Medium	High	Medium
Zohar (1996)	Medium	Medium	High	Medium

Four of the studies were judged to have ‘high’ weight of evidence overall, whilst just over half of the studies (12) were judged by the reviewers to provide 'medium' weight of evidence. Two out of these 12 studies (De Konig and Hamers, 1999; Ritchie and Edwards, 1996) were judged to be of high quality in terms of the research paradigm they used (weight of evidence A); but the evaluation indicated some limitations on possible generalisation or inference due to the context in which these studies were carried out (weight of evidence C), therefore reducing their applicability to the review. These studies were therefore considered ‘medium’ weight of evidence overall. Seven of the 23 studies were judged to be of low weight of evidence, usually due to a lack of detail in what was reported, also though some studies contain inconsistencies or actual errors in analysis. Further details for each study are given in Appendix 4.

For the purposes of this review, studies which were undertaken in school settings in usual classes and where the teaching was conducted by a class teacher (rather than a researcher) were of particular interest. These studies were considered to be valuable because the Review Group believed that the findings were more likely to be applicable than if a researcher undertook the teaching of a small group of pupils withdrawn from the class. The applicability of such research (or its ‘ecological validity’) was an important perspective for the Review Group.

The synthesis of findings from the studies reviewed indepth reveals a complex picture with no simple answer to the review question. The overall findings are clearly positive in that the majority of studies (15 out of 23) report positive impact on pupils’ attainment across a range of quantitative non-curriculum measures (such as reasoning or problem-solving). No studies report negative impact on such measures. These measures themselves were often chosen because of their general nature or the match between the intervention and the test. All four studies rated as highly applicable to the review found positive impact in this area, though the impact was not even across all the groups of pupils, either in terms of age (Adey and Shayer, 1990; Csapó, 1992), sex (Adey and Shayer, 1990) or prior attainment (Adey and Shayer, 1990; De Corte *et al*., 2001). So, whilst thinking skills programmes and approaches can perhaps be characterised as having

positive impact on cognitive abilities, they are not a panacea and can be expected to have differential effects depending upon the content of the programme and its implementation. (See Appendix 4.4 for the synthesis tables on which these findings are based.)

The evidence of impact upon pupils' attainment on curriculum measures of attainment is also complex. In terms of quantitative data, nine of the studies report immediate positive impact on learning on curricular measures of attainment (where such measures were used). This might be considered a surprising result as, compared with controls, they were probably spending less time on acquisition of curriculum content and more on aspects of the process of learning or making aspects of learning explicit. The implication is that using thinking skills approaches is at least as good as the regular curriculum and often produces better results. However, the immediate positive impact on curriculum was only found for one of the studies rated as highly applicable to the review (Chang and Barufaldi, 1999). Only one of the 23 studies found negative impact on some of curricular outcomes used to evaluate the programme (Haywood *et al.*, 1988). This was a study of the impact of the use of Feuerstein's Instrumental Enrichment on deaf adolescents where the experimental group improved significantly compared with controls on three measures of reasoning, but not on seven of ten of the measures of curricular attainment. More than half the studies report evidence from qualitative data for positive impact on pupils' attainment. However the reporting of qualitative data was problematic. Although the studies were selected because they contained qualitative as well as quantitative data, the reporting of these data does not include sufficient details (particularly about the actual sample and the method of analysis) to draw firm conclusions. Most report teachers' or pupils' views (or both) which were invariably positive overall, but anecdotal.

Where studies tracked learning gains over a longer period of time, the findings are again positive. These gains, achieved through thinking skills programmes and approaches, are usually retained or extended both in terms of curricular and non-curricular achievement (Adey and Shayer, 1990; Blank, 2000; Georghiades, 2000; Zohar, 1996). Where there is no, or small immediate, impact on curriculum measures, such improvement may appear later or increase over time (Adey and Shayer, 1990; Blank, 2000; Georghiades, 2000).

Although the issue of transfer of learning to other contexts is problematic, there is evidence from this review that pupils can apply or translate what they learn through thinking skills programmes and approaches to other contexts, when this is explicitly investigated (Adey and Shayer, 1990; De Corte *et al*., 2001; De Konig and Hamers, 1999; Georghiades, 2000). The most impressive example of this is in terms of the impact of cognitive acceleration through science education (CASE) where most pupils who had been taught using CASE in science lessons achieved significantly better GCSE results in mathematics and English, although gains were found with other age groups and subjects.

The impact of thinking skills approaches may not be the same for all pupils. Where studies report analysis of results in terms of the age of the pupils, their sex, or in terms of their prior attainment, there was variation in the impact of the programmes (Adey and Shayer, 1990; Csapó, 1992; Strang and Shayer, 1993; Verschaffel *et al*., 1997). Although it is hard to draw clear conclusions, it seems likely that this depends upon the specific content of particular programmes and approaches. There is some evidence that there may be greater impact on low attaining pupils, particularly when using metacognitive strategies (De Corte *et al*., 2001; Cardelle-Ellawar, 1992), although one study found that the impact on

higher attaining pupils was greatest (Verschaffel *et al*., 1997). This effect is the more usual one in educational interventions as prior attainment is usually the best predictor of subsequent achievement (Tymms and Fitz-Gibbon, 2001).

A final group of effects can be identified in terms of aspects of implementation of thinking skills programmes and approaches, although, again, the evidence is not straightforward. Most thinking skills programmes and approaches have a characteristic pedagogy, which involves explicit teaching strategies that include articulation and discussion of thinking. Sometimes this is through collaborative discussion and group-work (Wegerif, 1996), sometimes it is the explicit role of the teacher to 'mediate' learners' thinking as in Feuerstein's instrumental enrichment and sometimes both aspects are included, such as in CASE (Adey and Shayer, 1990). Some of the benefits of thinking skills programmes and approaches are therefore likely to derive from making thinking and reasoning explicit through this pedagogical emphasis on classroom talk and interaction. There is also some evidence that pupils benefit from explicit training in the use of specific strategies (Csapó, 1992; De Corte *et al*., 2001; Kramarski and Mevarech, 1997; Wegerif, 1996). Such changes in patterns of classroom interaction may not be easy to achieve and, in order to benefit from such approaches which use collaborative group work, pupils may need to be taught how to work in groups (Ritchie and Edwards, 1996; Wegerif, 1996). The role of the teacher is also important in thinking skills programmes and approaches in establishing effective patterns of talk and in eliciting pupils' responses to questions (Chang and Barufaldi, 1999; De Konig and Hamers, 1999; Cardelle-Ellawar, 1992; Verschaffel *et al*., 1997).

Table 4.4.2 Summary of findings and studies from which they derive

Finding	Warrant[5]
The majority of studies report positive impact on pupils' attainment across a range of non-curriculum measures (such as reasoning or problem-solving). No studies report negative impact on such measures.	15/23 studies, including all studies rated as highly applicable to the review: **Adey and Shayer, 1990; Chang and Barufaldi, 1999; Csapó, 1992; De Corte *et al*. 2001**[6]
Half the studies show immediate positive impact on learning on curricular measures of attainment (where such measures were used).	9/18, but only one of the studies rated as highly applicable to the review: **Chang and Barufaldi, 1999**
These gains are usually retained over time.	**Adey and Shayer, 1990**; Blank, 2000; Georghiades, 2000; Zohar, 1996
There is evidence that pupils can apply or translate this learning to other contexts.	**Adey and Shayer, 1990**; **De Corte *et al.,* 2001;** De Konig and Hamers, 1999; Georghiades, 2000
Where there is no or small immediate impact on curriculum measures, such improvement may appear later or increase over time.	**Adey and Shayer, 1990**; Blank, 2000; Georghiades, 2000
The impact of thinking skills approaches may not be even across all groups of pupils.	**Adey and Shayer, 1990**; **Csapo, 1992**; Strang and

[5] Studies given an overall 'High' weight of evidence are represented in bold.
[6] Adey and Shayer (1990) and Caspo (1992) also found uneven evidence of impact (i.e. some groups did better than others such as particular age groups).

	Shayer, 1993
There is some evidence that there may be greater impact on low attaining pupils, particularly when using metacognitive strategies.	**De Corte *et al.*, 2001**; Cardelle-Ellawar, 1992 [7]
There is some evidence that pupils benefit from explicit training in the use of thinking skills strategies and approaches.	**De Corte *et al.*, 2001**; **Csapó, 1992**; Kramarski and Mevarech, 1997; Wegerif, 1996
Some of the benefits of thinking skills programmes and approaches derive from making thinking and reasoning explicit through a pedagogical emphasis on classroom talk and interaction.	**Adey and Shayer, 1990**; **De Corte *et al.*, 2001**; Blank, 2000; Cardelle-Ellawar, 1992
The role of the teacher is important in thinking skills programmes and approaches in establishing effective patterns of talk and in eliciting pupils' responses.	**Chang and Barufaldi, 1999**; De Konig and Hamers, 1999; Cardelle-Ellawar, 1992; Verschaffel *et al.*, 1997
In order to benefit from thinking skills approaches which use collaborative group work, pupils may need to be taught how to work in groups.	Ritchie and Edwards, 1996; Wegerif, 1996

4.5 In-depth review: quality-assurance results

All studies were data-extracted by two reviewers independently. Eight of the studies were data-extracted by members of the EPPI-Centre. All differences in coding were resolved and an agreed version of the data-extraction uploaded prior to the synthesis stage of the review.

4.6 Nature of actual involvement of users in the review and its impact

Users were fully integrated into the Review Group and took a full and active part in each stage of the review. However, the online data-extraction was largely completed by the higher education institution (HEI) members of the group and then discussed with the wider group. This was a direct consequence of the time pressures and the lack of funding to release school-based staff for the length of time required to complete the exercise. We regard this as an issue of the best deployment of resources rather than any invocation of a hierarchy of expertise. Each member of the Review Group was involved in a paper exercise of data-extraction so that we were equally well informed as to the nature of the process; the subsequent synthesis of evidence was discussed at Review Group meetings and its significance for our review question was validated via consultation with colleagues from across the Review Group.

The impact of the review on all the participants has been to highlight the need to report studies in a more accessible and transparent manner, so that the methods and context of the study can be fully understood. Greater clarity in the use and provenance of the term 'thinking skills' has also been achieved within the group

[7] There is some counter-evidence from Verschaffel *et al.* (1997)

and now needs to be shared more widely, along with the key findings of the studies included in the in-depth review.

5. FINDINGS AND IMPLICATIONS

5.1 Summary of principal findings

5.1.1 Identification of studies

The identification of studies demonstrates how the use of terminology in the broad field of thinking skills has varied across the studies covered by the map. The implications of this for the identification of relevant studies is that cognate terms may be used to describe studies but their underpinning theoretical orientation may be different in ways that can become significant when seeking to review issues such as impact on pupils. The review suggests that there may be a need for greater conceptual clarity between interventions focused on cognitive education and approaches that have a philosophical basis and which seek to develop dispositions and real-life problem-solving.

We have been able to map out the areas in which research has been done and this can inform further enquiry into the impact of thinking skills. However, we also found that attempts to identify studies can be frustrated by the fact that the academic conventions of reporting, principally through publication in journals, militate against the transparency of description that we required. We were seeking to establish evidence for the impact of interventions on pupil learning in authentic classroom contexts. Although studies frequently reported detailed literature reviews, there were often only general details about the actual practicalities of the study, such as: was the group a normal class working with their usual teacher? We would also like to make a plea for greater standardisation of what should be included in an abstract.

5.1.2 Mapping of all included studies

Of the 191 reports identified as being relevant to the review and likely to contain data, about two-thirds are from the USA (34%) and the UK (27%). Nearly half of these reports were set in secondary schools (45%) and about a third in primary schools (34%). There is a greater proportion of research reported for the 11 to 14 year-old year groups with comparatively little about 5 to 7 year-olds (Key Stage 1 pupils). Most subjects of the curriculum are represented in these reports, although most were in the core areas of science (34%), literacy (20%) and mathematics (19%). The majority of reports contain data on pupil attainment, with just less than a quarter having data on pupils' attitudes or beliefs, and about an eighth with data on teachers' attitudes or beliefs. For most of the reports of pupil attainment, the data were quantitative (102 out of 136). On the other hand, about a third of the reports with data on teacher attitude or beliefs contained qualitative data only (13 out of 40).

5.1.3. Nature of studies selected for in-depth review

The 23 studies selected for in-depth review are a subset of the 191 reports keyworded in the earlier stage of the review. Of these 23 studies, 17 were judged to be of high or medium overall weight of evidence, and used in the synthesis of evidence. Just over half of the studies come from the USA or UK. They contain qualitative and quantitative data evaluating the impact of thinking skills programmes or approaches on pupil attainment. Most of the attainment data

relate to science, mathematics or aspects of literacy. Most of the studies evaluate the impact of named thinking skills programmes, which broadly cover the range of programmes identified in the mapping stage. A summary synthesis of the studies can be found in Appendix 4 along with synthesis tables describing the aims, research questions and findings of the studies selected for in-depth review.

5.1.4 Synthesis of findings from studies in in-depth review

The following synthesis of findings are from studies reviewed in depth (that is, studies with high and medium weights of evidence):

- The majority of studies report positive impact on pupils' attainment across a range of non-curriculum measures (such as reasoning or problem-solving). No studies reported negative impact on such measures.
- Half the studies found immediate positive impact on learning with curricular measures of attainment (where such measures were used).
- These gains are usually retained over time.
- There is evidence that pupils can apply, or translate, this learning to other contexts.
- Where there is no, or small immediate, impact on curriculum measures, such improvement may appear later or increase over time.
- The impact of thinking skills approaches may not be even across all groups of pupils.
- There is some evidence that there may be greater impact on low attaining pupils, particularly when using metacognitive strategies.
- There is some evidence that pupils benefit from explicit training in the use of thinking skills strategies and approaches.
- Some of the benefits of thinking skills programmes and approaches derive from making thinking and reasoning explicit through a pedagogical emphasis on classroom talk and interaction.
- The role of the teacher is important in thinking skills programmes and approaches in establishing effective patterns of talk and in eliciting pupils' responses.
- In order to benefit from thinking skills approaches which use collaborative group work, pupils may need to be taught how to work in groups.

This review has focused on studies that reflect the provenance of approaches to thinking skills in compulsory schooling that have ecological validity as they are classroom-based investigations and so the findings have direct relevance for users.

5.2 Strengths and limitations of this systematic review

5.2.1 Strengths

- *A strength of the review is the extent of the literature included in the review, the mapping of this literature to describe this research and the number of studies included in the in-depth review.*
  - This systematic review is based on an extensive search across a wide range of literature and descriptive map of reports included which helps to contextualise the research which has emerged through the process for in-depth analysis.

- *A strength of this review is the close involvement of users groups in setting and refining the questions and interpreting and disseminating the findings.*
  - The Review Group was keen to ensure that the perspectives of practitioners were included in the review and at every stage involved those working in, and with experience of schools, in order to maintain the link between research, the interpretation of that research and the development of practice in schools.

- *The Review Group is in a position to build on both the findings and experiences of this first review.*
  - The review provides the basis from which to investigate further the impact of thinking skills programmes and approaches. In particular a key question is to identify the size of the effect of thinking skills programmes and interventions through meta-analysis and to try to identify any further variables associated with effect size (such as intervention type, length of intervention, etc.) as well as the size of effect compared with other researched interventions.

  - The review details the impact of a range of thinking skills programmes and approaches on pupils' attainment. These findings have the potential to inform approaches to pedagogy in policy and practice. In particular the development of the Key Stage 3 strategy and the emerging of the national primary strategy could benefit from the findings.

5.2.2 Limitations

The Review Group was conscious throughout of the complexity of the questions we were trying to answer and limitations of the data in the separate studies that we examined in the review. Identifying the causes of impact on pupil attainment in classrooms is complex and the focus of the studies did not always acknowledge this complexity. In particular, we found the following:

- A tendency for the study reports to concentrate on either quantitative outcomes or details of implementation and classroom processes but rarely on both these types of data. It is therefore difficult to draw conclusions about common features of programmes and approaches which may account for the positive impact reported.
- A surprising lack of detail about qualitative data and its analysis, even where this was an explicit part of the research design. Analysis of these data (particularly teachers' and pupils' views) may have helped to explain some of the variability in the impact found, but was rarely presented in a systematic way.
- We were only able to retrieve 76% of the reports identified through searching and this may have had an impact on the focus of the in-depth review.
- In particular, there may well have been valuable data in a number of doctoral studies. However, we were unable to retrieve these within our timescale and note that these data remain unexplored.

5.3 Implications

Whilst thinking skills programmes and approaches have a positive impact on pupils' attainment, such impact is not always consistent. The evidence from this review suggests that there is a need to select interventions carefully and to be prepared to persist with an intervention, as it may not always provide

improvement on curricular measures in the short-term. Research also indicates that the causes of improvement in pupil learning are complex and a more general emphasis on making aspects of teaching and learning explicit in classrooms (particularly in terms of making reasoning explicit) may have similar benefits to those obtained through a particular programme of intervention. Further research across a wider range of subjects and age groups would be particularly useful, as would comparative research to evaluate the relative benefits of different thinking skills programmes and approaches, as well as a comparison of such approaches with other educational interventions.

5.3.1 Policy-makers

Provision of guidelines for the implementation and evaluation of thinking skills in classrooms based on research evidence would enable schools to make informed choices. Access to information to make links between thinking skills programmes and what is known about effective teaching and learning and national policy initiatives could be facilitated. Finally, research could be commissioned to establish what is both effective and efficient (particularly what is cost-effective) in terms of thinking skills interventions.

5.3.2 Practitioners

When introducing interventions that focus on improving specific cognitive strategies, it could be more efficient to target particular groups of pupils and identify the most appropriate times for development. On the other hand interventions aimed at developing a classroom ethos conducive to making learning more explicit and fostering dialogue about teaching and learning, can be promoted at any time. Positive outcomes on pupil motivation and self-esteem may be registered before any tangible impact on attainment as measured by standard assessments. There may be a delay of as much as two years in the appearance of improved attainment in tests and exams and, as a consequence, it may be difficult to distinguish between the impact of the intervention and the effect of any subsequent teaching.

5.3.3 Researchers

Further work is needed on identifying efficient, as well as effective, ways of intervening to promote thinking skills and raise attainment. There is a clear need for more comparative studies between different types of intervention and between thinking skills approaches and other strategies designed to change patterns of classroom interactions. The descriptive map shows where there are gaps in the research evidence. The in-depth review indicates where aspects of methodology and the reporting of findings could be more robust and accessible both to other researchers and other users of the findings.

6. REFERENCES

6.1 Reports included in the systematic map

Studies marked with an asterisk are also included in the in-depth review

Adams MJ (1989) Thinking skills curricula: their promise and progress. *Educational Psychologist* **24:** 25-77.

Adetula LO (1996) Effects of counting and thinking strategies in teaching addition and subtraction problems. *Educational Research* **38:** 183-198.

Adey P (1987) Science develops logical thinking - doesn't it? Part II. The CASE for science. *School Science Review* **69:** 17-27.

Adey P (1988) Cognitive acceleration: review and prospects. *International Journal of Science Education* **10:** 121-134.

Adey P (1992) The CASE results: implications for science teaching. *International Journal of Science Education* **14:** 137-146.

Adey P (1999) Thinking science: science as a gateway to general thinking ability, in: Hamers JHM, Van Luit JEH, Csapó B (eds.) *Teaching and Learning Thinking Skills*. Abingdon: Swets and Zeitlinger. (pp63-80).

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

Adey P, Shayer M (1988) Strategies for meta-learning in physics. *Physics Education* **23:** 97-104.

*Adey P, Shayer M (1990) Accelerating the development of formal thinking in middle and high-school students. *Journal of Research in Science Teaching* **27:** 267-285.

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 and Instruction* **11:** 1-29.

Adeyemi MB (1992) The relative effectiveness of the reflective and the lecture approach methods on the achievement of high school social studies students. *Educational Studies* **18:** 49-56.

Al-Qahtani SA (1995) Teaching thinking skills in the social studies curriculum of Saudi Arabian secondary schools. *International Journal of Educational Development* **15:** 155 -163.

Alaiyemola FF, Okebukola PAO, Jegede OJ (1990) The effect of a metacognitive strategy of instruction on the anxiety level of students in science classes. *International Journal of Science Education* **12:** 95-99.

Alexander PA (1987) Effects of teacher training on children's analogical reasoning performance. *Teaching and Teacher Education* **3:** 275-285.

Armstrong CL (1993) Effect of training in an art production questioning method on teacher questioning and student responses. *Studies in Art Education* **34:** 209-212.

Ayaduray J, Jacobs GM (1997) Can learner strategy instruction succeed? The case of higher order questions and elaborated responses. *System* **25:** 561-570.

Baird JR (1986) Improving learning through enhanced metacognition: a classroom study. *European Journal of Science Education* **8:** 263-282.

Barak M, Doppelt Y (1999) Integrating the Cognitive Research Trust (CoRT) Programme for creative thinking into a project-based technology curriculum. *Research in Science and Technological Education* **17:**139-151.

Barba RH, Merchant LJ (1990) The effects of embedding generative cognitive strategies in science software. *Journal of Computers in Mathematics and Science Teaching* **10:** 59-65.

Bass GM Jr, Perkins HW (1984) Teaching critical thinking skills with CAI. *Electronic Learning* **4:** 32-34.

Baumfield V, Higgins S (1997) But no one has maths at a party: pupils' critical thinking strategies in a thinking skills programme. *Curriculum* **18:** 140-148.

Baumfield V, Oberski I (1998) What do teachers think about thinking skills? *Quality Assurance in Education* **6:** 44-51.

Baumfield V, Mroz M (2002) Investigating pupils' questions in the primary classroom. *Educational Research* **44:** 129-140.

Beishuizen J, Le Grand J, Van der Schalk J (1999) No correlation between inferencing causal relations and text comprehension? *Learning and Instruction* **9:** 37-56.

Ben-Chaim D (1998) Proportional reasoning among 7th grade students with different curricular experiences. *Educational Studies in Mathematics* **36:** 247-273.

Benito YM (1993) The effect of instruction in question-answer relationships and metacognition on social studies comprehension. *Journal of Research in Reading* **16:** 20-29.

*Blank LM (2000) A metacognitive learning cycle: a better warranty for student understanding. *Science Education* **84:** 86-506.

*Bowdler D, Webb T, Dyke SAF (1992) Using 'Somerset Thinking Skills' to promote oracy in the classroom. In: Jones N and Jones N (eds.) *Learning to Behave: Curriculum and Whole School Management Approaches to Discipline*. London: Kogan Page (pp159-169).

Bryson M, Scardamalia M (1991*) Teaching Writing to Students at Risk for Academic Failure.* San Francisco, Jossey-Bass.

Burden R, Nichols L (2000) Evaluating the process of introducing a thinking skills programme into the secondary school curriculum. *Research Papers in Education* **15:** 293-306.

*Cardelle-Elawar M (1992) Effects of teaching metacognitive skills to students with low mathematics ability. *Teaching and Teacher Education* **8:** 109-121.

Cardelle-Elawar M (1995) Effects of metacognitive instruction on low achievers in mathematics problems. *Teaching and Teacher Education* **11:** 81-95.

*Chang C.-Y, Barufaldi JP (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:*** 373-388.

Chang L.-W, Kyriacou C (1996) Teaching thinking skills to primary school pupils in Taiwan. *Education Section Review* **20:** 28-30.

Chinnappan M, Lawson MJ (1996) The effects of training in the use of executive strategies in geometry problem solving. *Learning and Instruction* **6:** 1-17.

Clement JJ (1998) Expert novice similarities and instruction using analogies. *International Journal of Science Education* **20:** 1271-1286.

Cobb P (1999) Individual and collective mathematical development: the case of statistical data analysis. *Mathematical Thinking and Learning* **1:** 5-43.

Cohen L, Williams S, Keane KJ (1993) Mediating a single student: a case study. *International Journal of Cognitive Education and Mediated Learning* **3:** 39-46.

Coles MJ (1995) Critical thinking, talk and a community of enquiry in the primary school. *Language and Education* **9:** 161-177.

Collings JN (1994) Some fundamental questions about scientific thinking. *Research in Science and Technological Education* **12:** 161-173.

Crump WD, Schlichter CL, Palk BE (1988) Teaching HOTS in the middle and high school: a district-level initiative in developing higher order thinking skills. *Roeper Review* **10:** 205-211.

*Csapó B (1992) Improving operational abilities in children. In: Demetriou A, Shayer M, Efklides A (eds.) *Neo-Piagetian Theories of Cognitive Development.* London: Routledge (pp144-159).

D'Agostino JV (1996) Authentic instruction and academic achievement in compensatory education classrooms. *Studies in Educational Evaluation* **22:** 139-155.

*De Corte E, Verschaffel L, De Ven AV (2001) Improving text comprehension strategies in upper primary school children: a design experiment. *British Journal of Educational Psychology* **71:** 531-559.

*De Koning E, Hamers JHM (1999) Teaching inductive reasoning: theoretical background and educational implications. In: Hamers JHM, Van Luit JEH, Csapó B (eds.) *Teaching and Learning Thinking Skills.* Abingdon: Swets and Zeitlinger. (pp157-188).

Deadman G (1997) An analysis of pupils' reflective writing within a hypermedia framework. *Journal of Computer Assisted Learning* **13:** 16-25.

Donnelly FC, Helion J, Fry F (1999) Modifying teacher behaviors to promote critical thinking in K-12 physical education. *Journal of Teaching in Physical Education* **18:** 199-215.

Endler LC, Bond T (2001) Cognitive development in a secondary science setting. *Research in Science Education* **30:** 403-416.

Ertepinar H, Geban O (1996) Effect of instruction supplied with the investigative-oriented laboratory approach on achievement in a science course. *Educational Research* **38:** 333-341.

Ewing JM, Dowling JD, Coutts N (1998) Teaching thinking IS possible through information and communications technology. *Virtual University Journal* **1:** 127-141.

Farha JL, Milbrandt MS (1989) Beyond lollipop trees: teaching thinking skills through art. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA.

*Fields JI (1995) Empirical data research into the claims for using philosophy techniques with young children. *Early Child Development and Care* **107:** 115-128.

Fisher R (1997) Thinking about thinking. *Curriculum* **18:** 117-128.

Fisher R (2001) Philosophy in primary schools: fostering thinking skills and literacy. *Reading* **35:** 67-73.

Foster V (1994) Brill to the rescue. *Special Children*. **75:** 11-15

Franke ML, Fennema E, Carpenter T, Ansell E, Behrend J (1998) Understanding teachers' self-sustaining, generative change in the context of professional development. *Teaching and Teacher Education* **14:** 67-80.

Franke ML, Kazemi E (2001) Learning to teach mathematics: focus on student thinking. *Theory Into Practice* **40:** 102-109.

Fisher N, Gerdes K, Logue T, Smith L, Zimmerman I (1998) Improving students' knowledge and attitudes of science through the use of hands-on activities. Unpublished Masters action research project. Illinois: Saint Xavier University.

*Georghiades P (2000) Beyond conceptual change learning in science education: focusing on transfer. *Educational Research* **42:** 119-139.

Gerber BL, Cavallo AML, Marek EA (2001) Relationships among informal learning environments, teaching procedures and scientific reasoning ability. *International Journal of Science Education* **23:** 535-549.

Greenberg KH (2000) Attending to hidden needs: the cognitive enrichment advantage perspective. *Educational and Child Psychology. Special Issue: Psychological influences upon educational intervention* **17:** 51-69.

Guthrie JT, Van Meter P, McCann A, Wigfield A, Bennett L, Poundstone C, Rice ME, Fabisch F, Hunt B, Mitchell A (1996) Growth of literacy engagement: changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly* **31:** 306-332.

Hager W, Hasselhorn M (1998) The effectiveness of the cognitive training for children from a differential perspective: a meta-evaluation. *Learning and Instruction* **8:** 411-438.

Hamaker A, Jordan P, Backwell J (1998) An evaluation of a two year cognitive intervention programme in technology education for Key Stage 4. *Journal of Design and Technology Education* **3:** 26-33.

Hansler DD (1985) Studies on the effectiveness of the cognition enhancement technique for teaching thinking skills. Unpublished research report. Washington, US. (ERIC Doc no. 266432).

Haywood HC, Brooks P, Burns S (1986) Stimulating cognitive development at developmental level: a tested, non-remedial preschool curriculum for preschoolers and older retarded children. *Special Services in the Schools* **3:** 127-147.

*Haywood HC, Towery-Woolsey J, Arbitman-Smith R, Aldridge AH (1988) Cognitive education with deaf adolescents: effects of instrumental enrichment. *Topics in Language Disorders* **8:** 23-40.

*Head G, O'Neill W (1999) Introducing Feuerstein's instrumental enrichment in a school for children with social, emotional and behavioural difficulties. *Support for Learning* **14:** 122-128.

Heather A, Vinson M (1988) The Kent experiment. *Special Children* **24:** 12-14.

Hernstein RJ, Nickerson RS, Sanchez M, Swets JA (1986) Psychology in action: teaching thinking skills. *American Psychologist* **41:** 1279-1289.

Hoek D, van den Eeden P, Terwel, J. (1997) The effects of social and cognitive strategies instruction on the mathematics achievement in secondary education. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL, March 1997 (ED 406 213).

Hogan K, Nastasi BK, Pressley M (1999) Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction* **17:** 379-432.

*Hojnacki SK, Grover BW (1992) Thinking mathematics: what's in it for the students? Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA: April 21.

Hooper M-L (1994) The effects of high and low level cognitive and literacy language arts tasks on motivation and learning in multiability, multicultural classrooms: developmental studies. *Learning and Instruction* **4:** 233-251.

Howie DR, Thickpenny JP, Leaf CA, Absolum MA (1985) The piloting of 'instrumental enrichment' in New Zealand with eight mildly retarded children. *Australia and New Zealand Journal of Developmental Disabilities* **11:** 3-16.

Huberty CJ, Davis EJ (1998) Evaluation of a state critical thinking skills training program. *Studies in Educational Evaluation* **24:** 45-69.

Hudgins B, Edelman S (1986) Teaching critical thinking skills to fourth and fifth graders through teacher-led small-group discussions. *Journal of Educational Research* **79:** 333-342.

Hymans MH (1994) Impulsive behaviour: a case for helping children 'think' about change. *Educational Psychology in Practice* **10:** 141-148.

Iqbal HM Shayer M (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:** 259-274.

Jarvinen E-M, Twyford J (2000) The influences of socio-cultural interaction upon children's thinking and actions in prescribed and open-ended problem solving situations: an investigation involving design and technology lessons in English and Finnish primary schools. *International Journal of Technology and Design Education* **10:** 21-41.

Kagan DM (1988) Evaluating a language arts program designed to teach higher level thinking skills. *Reading Improvement* **25:** 29-33.

Kaniel S, Licht P, Peled B (2000) The influence of metacognitive instruction of reading and writing strategies on positive transfer. *Gifted Education International* **15:** 45-63.

Kaniel S, Reichenberg R (1992) Instrumental enrichment: effects of generalization and durability with talented adolescents. *Gifted Education International* **8:** 128-135.

Keane KJ, Kretschmer RE (1987) Effect of mediated learning intervention on cognitive task performance with a deaf population. *Journal of Educational Psychology* **79:** 49-53.

Keys CW (1995) An interpretive study of students' use of scientific reasoning during a collaborative report writing intervention in ninth grade general science. *Science Education* **79:** 415-435.

Kirkwood M (2000) Infusing higher-order thinking and learning to learn into content instruction: a case study of secondary computing studies in Scotland. *Journal of Curriculum Studies* **32:** 509-535.

Klauer KJ (1996) Teaching inductive reasoning: some theory and three experimental studies. *Learning and Instruction* **6:** 37-57.

Klein PS, Nir Gal O (1992) Effects of computerized mediation of analogical thinking in kindergartens. *Journal of Computer Assisted Learning* **8:** 244-254.

Kosonen P, Winne PH (1995) Effects of teaching statistical laws on reasoning about everyday problems. *Journal of Educational Psychology* **87:** 33-46.

Koufetta-Menicou C, Scaife JA (2000) Teachers' questions types and significance in science education. *School Science Review* **81:** 79-84.

*Kramarski B, Feldman Y (2000) Internet in the classroom: effects on reading comprehension, motivation and metacognitive awareness. *Educational Media International* **37:** 149-155.

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

Ladwig JG, King MB (1992) Restructuring secondary social studies: the association of organizational features and classroom thoughtfulness. *American Educational Research Journal* **29:** 695-714.

Lauchlan F, Elliott J (1997) Using dynamic assessment materials as a tool for providing cognitive intervention to children with complex learning difficulties. *Educational and Child Psychology* **14:** 137-148.

Leat D (1997) Getting ambiguous. *Educating Able Children* **1:** 17-25.

Leat D (1999) Rolling the stone uphill: teacher development and the implementation of thinking skills programmes. *Oxford Review of Education* **25:** 387-403.

Leshowitz B, Jenkins K, Heaton S, Bough TL (1993) Fostering critical thinking skills in students with learning disabilities: an instructional program. *Journal of Learning Disabilities* **26:** 483-490.

Luther M, Wylie B (1990) Feuerstein's thinking skills program with children, adolescents, and adults in 'special needs' schools. *International Journal of Dynamic Assessment and Instruction* **1:** 49-70.

MacArthur RP, Okolo CD, Ferretti CM (2001) Teaching for historical understanding in inclusive classrooms. *Learning Disability Quarterly* **24:** 59-71.

*Maltby F (1995) The use of TASC to develop a selection of tools for effective thinking. *Gifted Education International* **11:** 18-23.

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

Markovits Z, Hershkowitz R (1997) Relative and absolute thinking in visual estimation processes. *Educational Studies in Mathematics* **32:** 29-47.

Martin DS (1984) Cognitive modification for the hearing impaired adolescent: the promise. *Exceptional Children* **51:** 235-242.

Mason L (1994a) Analogy, metaconceptual awareness and conceptual change: a classroom study. *Educational Studies* **20:** 267-291.

Mason L (1994b) Cognitive and metacognitive aspects in conceptual change by analogy. *Instructional Science* **22:** 157-187.

Mason L, Santi M (1998) Discussing the greenhouse effect: children's collaborative discourse reasoning and conceptual change. *Environmental Education Research* **4:** 67-85.

*Mastropieri MA, Scruggs TE, Shian R-L (1997) Can computers teach problem-solving strategies to students with mild mental retardation? *Rase: Remedial and Special Education* **18:** 157-165.

McCartney C, Schrag F (1990) Departmental and school leadership in promoting higher-order thinking. *Journal of Curriculum Studies* **22:** 529-543.

McGregor D, Gunter B (2001) Changing pedagogy of secondary science teachers: the impact of a two-year professional development programme. *Teacher Development* **5:** 59-74.

Medrano C, De la Caba MA (1994) A model of intervention for improving moral reasoning: an experiment in the Basque country. *Journal of Moral Education* **23:** 427-437.

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

Mevarech ZR, Kramarski B (1993) Vygotsky and Papert: social-cognitive interactions within Logo environments. *British Journal of Educational Psychology* **63:** 96-109.

Mioduser D, Santa Maria M (1995) Students' construction of structured knowledge representations. *Journal of Research in Computing in Education* **28:** 63-84.

Moore PJ (1993) Metacognitive processing of diagrams, maps and graphs. *Learning and Instruction* **3:** 215-226.

Muttart K (1984) Assessment of effects of instrumental enrichment cognitive training. *Special Education in Canada* **58:** 106-108.

Naisbett A (1997) Policy and provision for more able pupils: the Nunthorpe experience. *Support for Learning* **12:** 83-89.

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

Newmann FM (1990) Qualities of thoughtful social studies classes: an empirical profile. *Journal of Curriculum Studies* **22:** 253-275.

Nickerson RS (1986) Project intelligence: an account and some reflections. *Special Services in the Schools* **3:** 83-102.

Niedelman M (1991) Problem solving and transfer. *Journal of Learning Disabilities* **24:** 322-329.

Northing C (1989) Teaching thinking skills: the development of design concepts in school pupils. *Journal of Art and Design Education* **8:** 23-36.

Oladunni MO (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 and Technology* **29:** 867-874.

Onosko JJ (1992) Exploring the thinking of thoughtful teachers. *Educational Leadership* **49:** 40-43.

Pasnak R, Hansbarger A, Dodson S, Hart J Blaha, J (1996) Differential results of instruction at the preoperational/concrete operational transition. *Psychology in the Schools* **33:** 70-83.

Peat D, Wilgosh L, Mulcahy R (1996) Efficacy of cognitive strategy-based instruction for elementary students with learning disabilities: a retrospective study. *Canadian Journal of School Psychology* **12:** 135-142.

Peterson D, Kromrey J, Borg J, Lewis A (1990) Defining and establishing relationships between essential and higher order teaching skills. *Journal of Educational Research* **84:** 5-12.

Pogrow S (undated) Systematically using powerful learning environments to accelerate the learning of disadvantaged students in grades 4-8. Unpublished research report. Tucson, AZ: University of Arizona.

Pogrow S (1988a) HOTS: a thinking skills program for at risk students. *Principal* **67:** 19-24.

Pogrow S (1988b) Teaching thinking to at-risk elementary students *Educational Leadership* **45:** 79-85.

Pogrow S (1995) A revalidation of the effectiveness of the HOTS program. Unpublished research report prepared for review by the PEP committee of the National Diffusion Network. Tucson, AZ: University of Arizona.

Popham WJ (2000) Putting instruction on the line. *School Administrator* **57:** 46-48.

Powell SD, Makin M (1994) Enabling pupils with learning difficulties to reflect on their own thinking. *British Educational Research Journal* **20:** 579-593.

Quicke J, Winter C (1994) Teaching the language of learning: towards a metacognitive approach to pupil empowerment. *British Educational Research Journal* **20:** 429-445.

Rafoth MA (1998) Sharing the secret: peer tutoring and independent learning. *Teaching and Change* **5:** 245-260.

Ratcliffe M (1997) Pupil decision-making about socio-scientific issues within the science curriculum. *International Journal of Science Education* **19:** 167-182.

Resnick LB, Bill V, Lesgold S (1992) Developing thinking abilities in arithmetic class. In: Demetriou A, Shayer M, Efklides A (eds.) *Neo-PiagetianTheories of Cognitive Development.* London: Routledge (pp 210-230).

Riding RJ, Powell SD (1987) The effect on reasoning, reading and number performance of computer presented critical thinking activities in five-year-old children. *Educational Psychology* **7:** 55-65.

Ristow RS (1988) The teaching of thinking skills: does it improve creativity? *Gifted Child Today* **11:** 44-46.

*Ritchie SM, Edwards J (1996) Creative thinking instruction for aboriginal children. *Learning and Instruction* **6:** 59-75.

Roberts C, Ingram C, Harris C (1992) The effect of special versus regular classroom programming on higher cognitive processes of intermediate elementary aged gifted and average ability students. *Journal for the Education of the Gifted* **15:** 332-343.

Ross JA, Bradley CJ (1993) Patterns of student growth in reasoning about correlational problems. *Journal of Educational Psychology* **85:** 49-65.

Ross JA, Cousins JB (1993) Enhancing secondary school students' acquisition of correlational reasoning skills. *Research in Science and Technological Education* **11:** 191-205.

Rottman TR, Cross DR (1990) Using informed strategies for learning to enhance the reading and thinking skills of children with learning disabilities. *Journal of Learning Disabilities* **23:** 270-278.

Rust AL, Donegan JO (1998) Rational emotive education for improving self-concept in second-grade students. *Journal of Humanistic Education and Development* **36:** 248-256.

Sarapuu T, Adojaan K (1999) Usage of educational web pages to develop students' higher order thinking skills. In: Proceedings of SITE 99 - 10th International Conference, AACE, San Antonio, 257-262.

Scheinin PM, Mehtäläinen J (1999) Applying the theory of knowledge to teaching thinking. In: Hamers JHM, Van Luit JEH ,Csapó B (eds) *Teaching and Learning Thinking Skills.* Abingdon: Swets and Zeitlinger. (pp 81-104)

Schmid RF, Gelaro T (1990) Concept mapping as an instructional strategy for high school biology. *Journal of Educational Research* **84:** 78-85.

Scruggs TE, Thomas E, Mastropieri, Sullivan MA, Sharon G (1994) Promoting relational thinking: elaborative interrogation for students with mild disabilities. *Exceptional Children* **60:** 450-457.

Shayer M (1999) Cognitive acceleration through science education II: its effects and scope. *International Journal of Science Education* **21:** 883-902.

Shayer M, Adey PS (1992a) Accelerating the development of formal thinking in middle and high school students III: testing the permanency of effects. *Journal of Research in Science Teaching* **29:** 1101-1115.

Shayer M, Adey PS (1992b) Accelerating the development of formal thinking in middle and high-school students II: postproject effects on science achievement. *Journal of Research in Science Teaching* **29:** 81-92.

Shayer M, Adey PS (1993) Accelerating the development of formal thinking in middle and high-school students IV: three years after a two-year intervention. *Journal of Research in Science Teaching* **30:** 351-366.

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

Sigurborsdottir I (1998) Philosophy with children in Foldaborg: development project in Foldaborg, a preschool in Reykjavik for children from 1-6 years. *International Journal of Early Childhood* **30:** 14-16.

Silven M (1992a) The role of metacognition in reading instruction. *Scandinavian Journal of Educational Research* **36:** 211-221.

Silven M, Vauren M (1992b) Improving reading through thinking aloud. *Learning and Instruction* **2:** 69-88.

Silverman H, Waksman M (1988) Feuerstein's instrumental enrichment elicitation of cognitive interaction in the classroom. *Canadian Journal of Special Education* **4:** 133-150.

Sinatra R, Beaudry J, Pizzo J, Geisert G (1994) Using a computer-based semantic mapping, reading and writing approach with at-risk fourth graders. *Journal of Computing in Childhood Education* **5:** 93-112.

Skuy M, Mentis M, Durbach F, Cockcroft K, Fridjhon P (1995) Crosscultural comparison of effects of instrumental enrichment on children in a South African mining town. *School Psychology International* **16:** 265-282.

Smith MW (1992a) Effects of direct instruction on ninth graders' understanding of unreliable narrators. *Journal of Educational Research* **85:** 339-347.

Smith SW (1992b) *Effects of a Metacognitive Strategy on Aggressive Acts and Anger Behavior of Elementary and Secondary-Aged Students.* Florida: Florida Educational Research Council.

Smith T (2000) Bridging the research-practice gap: developing a pedagogical framework that promotes mathematical thinking and understanding. *Mathematics Teacher Education and Development* **2:** 4-16.

Stein MK, Lane S (1996) Instructional tasks and the development of student capacity to think and reason: an analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation* **2:** 50-80.

Stevenson RB (1990) Engagement and cognitive challenge in thoughtful social studies classes: a study of student perspectives. *Journal of Curriculum Studies* **22:** 329-341.

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

Subhi T (1999) The impact of LOGO on gifted children's achievement and creativity. *Journal of Computer Assisted Learning* **15:** 98-108.

Tanner H, Jones S (1994) Using peer and self-assessment to develop modeling skills with students aged 11 to 16: a socio-constructive view. *Educational Studies in Mathematics* **27:** 413-431.

Taverner S (2001) A case study of the professional development of a main scale teacher of mathematics. *Mathematics Education Review* **13:** 1-9.

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:** 105-114.

Tzuriel D (2000) The Seria-Think Instrument: development of a dynamic test for young children. *School Psychology International* **21:** 177-194.

Tzuriel D, Alfassi M (1994) Cognitive and motivational modifiability as a function of the Instrumental Enrichment (IE) Program. *Special Services in the Schools* **8:** 91-128.

Van Luit JEH (1999) Teaching mathematical thinking to children with special needs. In: Hamers JHM, Van Luit JEH , Csapó B (eds) *Teaching and Learning Thinking Skills*. Abingdon: Swets and Zeitlinger. (pp241-258).

Vauras M (1999) Motivational vulnerability as a challenge for educational interventions. *International Journal of Educational Research* **31:** 515-531.

*Verschaffel L, De Corte E (1997) Teaching realistic mathematical modeling in the elementary school: a teaching experiment with fifth graders. *Journal for Research in Mathematics Education* **28:** 577-601.

*Ward L, Traweek D (1993) Application of a metacognitive strategy to assessment, intervention and consultation: a think-aloud technique. *Journal of School Psychology* **31:** 469-485.

*Wegerif R (1996) Using computers to help coach exploratory talk across the curriculum. *Computers and Education* **26:** 51-60.

Wegerif R, Mercer N (1996) Computers and reasoning through talk in the classroom. *Language and Education* **10:** 47-64.

Wegerif R, Mercer N, Dawes L (1999) From social interaction to individual reasoning: an empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction* **9:** 493-516.

Wiersema B, Van Oudenhoven JP (1992) Effects of cooperation on spelling achievement at three age levels (grades 2, 4, and 6). *European Journal of Psychology of Education* **7:** 95-108.

Wilks S (1997) Teaching aesthetics: improving critical thinking in visual arts education. *Curriculum* **18:** 171-189.

Wilks S, Emery L (1998) Aesthetics and critical thinking in visual arts education. *Australian Art Education* **21:** 61-70.

Williams M (2000) The part which metacognition can play in raising standards in English at Key Stage 2. *Reading* **34:** 3-8.

Wright J, Cashdan A (1991) Training metacognitive skills in backward readers: a pilot study. *Educational Psychology in Practice* **7:** 153-162.

Yarrow F, Topping KJ (2001) Collaborative writing: the effects of metacognitive prompting and structured peer interaction. *British Journal of Educational Psychology* **71:** 261-282.

Yates C (1987) Teaching correlational reasoning to eleven-to-thirteen year olds. *Journal of Biological Education* **21:** 197-202.

Yildirim A (1994) Teachers' theoretical orientations toward teaching thinking. *Journal of Educational Research* **88:** 28-35.

Zenke L, Alexander L (1984) Teaching thinking in Tulsa. *Educational Leadership* **42:** 81-84.

Zohar A (1996) Transfer and retention of reasoning strategies taught in biological contexts. *Research in Science and Technological Education* **14:** 205-219.

Zohar A (1998) Result or conclusion? Students' differentiation between experimental results and conclusions. *Journal of Biological Education* **32:** 53-59.

Zohar A (1999) Teachers' metacognitive knowledge and the instruction of higher order thinking. *Teaching and Teacher Education* **15:** 413-429.

Zohar A, Degani A, Vaaknin E (2001) Teachers' beliefs about low-achieving students and higher order thinking. *Teaching and Teacher Education* **17:** 469-485.

6.2 Studies selected for in-depth review

Adey P, Shayer M (1990) Accelerating the development of formal thinking in middle and high-school-students. *Journal of Research in Science Teaching* **27:** 267-285.

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 and Instruction* **11:** 1-29.[1]

Blank LM (2000) A metacognitive learning cycle: a better warranty for student understanding. *Science Education* **84:** 486-506.

Bowdler D, Webb T, Dyke SAF (1992) Using 'Somerset Thinking Skills' to promote oracy in the classroom. In: Jones N, Jones N (eds.) *Learning to Behave: Curriculum and Whole School Management Approaches to Discipline*. London: Kogan Page (pp 159-169).

Cardelle-Elawar M (1992) Effects of teaching metacognitive skills to students with low mathematics ability. *Teaching and Teacher Education* **8:** 109-121.

Chang C-Y, Barufaldi JP (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:** 373-388.

Csapó B (1992) Improving operational abilities in children. In: Demetriou A, Shayer M, Efklides A (eds.) *Neo-Piagetian Theories of Cognitive Development*. London: Routledge, (pp 144-159).

De Corte E, Verschaffel L, Van De Ven A (2001) Improving text comprehension strategies in upper primary school children: a design experiment. *British Journal of Educational Psychology* **71:** 531-559.

De Koning E, Hamers JHM (1999) Teaching inductive reasoning: theoretical background and educational implications. In: Hamers JHM, Van Luit JEH, Csapó B (eds.) *Teaching and Learning Thinking Skills*. Abingdon: Swets and Zeitlinger. (pp 157-188).

Fields JI (1995) Empirical data research into the claims for using philosophy techniques with young children. *Early Child Development and Care* **107:** 115-128.

Georghiades P (2000) Beyond conceptual change learning in science education: focusing on transfer. *Educational Research* **42:** 119-139.

Haywood HC, Towery-Woolsey J, Arbitman-Smith R, Aldridge AH (1988) Cognitive education with deaf adolescents: effects of instrumental enrichment. *Topics in Language Disorders* **8:** 23-40.

[1] Additional data was taken from this report to supplement that of Adey and Shayer (1990)

Head G, O'Neill W (1999) Introducing Feuerstein's instrumental enrichment in a school for children with social, emotional and behavioural difficulties. *Support for Learning* **14:** 122-128.

Hojnacki SK, Grover BW (1992) Thinking mathematics: what's in it for the students? Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA: April 21.

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

Maltby F (1995) The use of TASC to develop a selection of tools for effective thinking. *Gifted Education International* **11:** 18-23.

Mastropieri MA, Scruggs TE, Shiah R-L (1997) Can computers teach problem-solving strategies to students with mild mental retardation? *Rase: Remedial and Special Education* **18:** 157-165.

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

Ritchie SM, Edwards J (1996) Creative thinking instruction for aboriginal children. *Learning and Instruction* **6:** 59-75.

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

Verschaffel L, De Corte E (1997) Teaching realistic mathematical modeling in the elementary school: a teaching experiment with fifth graders. *Journal for Research in Mathematics Education* **28:** 577-601.

Ward L, Traweek D (1993) Application of a metacognitive strategy to assessment, intervention, and consultation: a think-aloud technique. *Journal of School Psychology* **31:** 469-485.

Wegerif R (1996) Using computers to help coach exploratory talk across the curriculum. *Computers and Education* **26:** 51-60.

Zohar A (1996) Transfer and retention of reasoning strategies taught in biological contexts. *Research in Science and Technological Education* **14:** 205-219.

6.3 Other references used in the text of the report

Adey P, Shayer M (1994) *Really Raising Standards.* London: Routledge.

Adey PS, Shayer M, Yates C (1995) *Thinking Science: The Curriculum Materials of the CASE Project.* London: Thomas Nelson and Sons.

Baumfield VM, Oberski IO (1998) What do teachers think about thinking skills? *Quality Assurance in Education* **6:** 44-51.

Biggs JB, Collis KF (1982) *Evaluating the Quality of Learning-the SOLO Taxonomy* (1st ed.). New York: Academic Press.

Black P, Wiliam D (1998) Assessment and classroom learning. *Assessment in Education* **5:** 7-71.

Blagg N, Ballinger M, Gardner R (1988) *Somerset Thinking Skills Course Handbook.* Oxford: Basil Blackwell.

Bloom BS (ed.) (1956) *Taxonomy of Educational Objectives, the Classification of Educational Goals - Handbook I: Cognitive Domain*. New York: McKay.

Brown AL, Palincsar AS (1989) Guided cooperative learning and individual knowledge acquisition. In: Resnick L (ed.) *Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser.* Hillsdale, NJ: Erlbaum (pp 393-451).

Csapó B (1992) Improving operational abilities in children. In: Demetriou A, Shayer M, Efklides A (eds.) *Neo-Piagetian Theories of Cognitive Development: Implications and Applications For Education.* London: Routledge (pp 144-159).

Dawes L, Mercer N, Wegerif R (2000) *Thinking Together: A Programme of Activities for Developing Thinking Skills at KS2.* Birmingham: Questions Publishing.

DeBono E (1970) *Lateral Thinking.* London: Penguin.

DeBono E (1992) *Teach Your Child to Think.* London: Penguin.

Dweck C (1999) *Self Theories: Their Role in Motivation, Personality and Development.* Hove, Sussex: Psychology Press.

Edwards A, Westgate D (1987) *Investigating Classroom Talk.* London: Falmer Press.

Elliot J (2001) Making evidence-based practice educational. *British Educational Research Journal* **27:** 555-574.

EPPI-Centre (2001) *EPPI-Centre Review Group Manual Version 1.1*. London: Social Science Research Unit, Institute of Education, University of London.

EPPI-Centre (2002a) *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.

EPPI-Centre (2002b) *Guidelines for Extracting Data and Quality Assessing Primary Studies in Educational Research. (Version 0.9.7).* London: EPPI-Centre, Social Science Research Unit, Institute of Education, Universtiy of London.

Evans J, Benefield P (2001) Systematic reviews of educational research: does the medical model fit? *British Educational Research Journal* **27:** 527-541.

Feuerstein R, Rand Y, Hoffman MB, Miller R (1980) *Instrumental Enrichment: an Intervention Programme for Cognitive Modifiability.* Baltimore: University Park Press.

Fisher R (1996) *Stories for Thinking.* Oxford: Nash Pollock.

Fisher R (1998) *Teaching Thinking: Philosophical Enquiry in the Classroom.* London: Cassell.

Gall MD, Borg WR, Gall JP (1996) *Educational Research: An Introduction* (6th ed.). London: Longman.

Galton G, Hargreaves L, Comber C, Pell A (1999) *Inside the Primary Classroom 20 Years On*. London: Routledge.

Garnham A, Oakhill J (1994) *Thinking and Reasoning.* Oxford: Blackwell.

Gunstone R, White R (1992) *Probing Understanding*: London: Falmer Press.

Hamers JHM, Van Luit JEH (1999) *Teaching and Learning Thinking Skills.* Abingdon: Swets and Zeitlinger.

Hamers JHM, Overtoom MT (eds.) (1997) *Teaching thinking in Europe: Inventory of European programmes*. Utrecht: SARDES.

Hammersley M (2001) On 'systematic' reviews of research literatures: a 'narrative' response to Evans and Benefield. *British Educational Research Journal* **27:** 543-554.

Hattie J, Biggs J, Purdie N (1996) Effects of learning skills interventions on student learning: a meta-analysis. *Review of Educational Research* **66:** 99-136.

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

Higgins S, Leat D (1997) Horses for courses or courses for horses: what is effective teacher development? *British Journal of In-Service Education* **23:** 303-314.

Higgins S, Moseley D (2002) Raising achievement in literacy through ICT. In Monteith M (ed.) *Teaching Primary Literacy with ICT.* Buckingham: Open University Press. (pp 30-45).

Lake M, Needham M (1993) *Top Ten Thinking Tactics.* Birmingham: Questions Publishing Company.

Leat D, Higgins S (2002) The role of powerful pedagogical strategies in curriculum development. *The Curriculum Journal* **13:** 71-85.

Lipman M (1991) *Thinking in Education* (1st. edition). Cambridge: Cambridge University Press.

Lipman M (1995) Caring as thinking. *Inquiry: Critical Thinking Across the Disciplines* **15:** 1-13.

Lipman M (2003) *Thinking in Education* (2nd. edition). Cambridge: Cambridge University Press.

Lipman M, Sharp A, Oscanyan F (1980) *Philosophy in the Classroom.* Princeton: Temple University Press.

Marzano RJ (1998) *A Theory-Based Meta-Analysis of Research on Instruction*. Aurora, Colorado: Mid-continent Regional Educational Laboratory. Available from: http://www.mcrel.org/topics/productDetail.asp?topicsID=6&productID=83.

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 NAF (1995) Developing children's thinking: a tale of three projects. *Irish Journal of Psychology* **16:** 378-388.

Mercer N (1995) *The Guided Construction of Knowledge: Talk Amongst Teachers and Learners.* Clevedon: Multilingual Matters.

Moseley D, Baumfield V, Higgins S, Lin M , Miller J, Newton D, Robson S, Elliott J, Gregson M (2004) *Thinking Skill Frameworks for Post-16 Learners: An Evaluation.* A research report for the learning and skills research centre. Trowbridge: Cromwell Press. Available from: http://www.lsda.org.uk/files/pdf/1541.pdf

Murris K, Haynes J (2001) *Storywise: Thinking Through Stories.* Dialogue Works: Newport.

Nickerson R, Perkins D, Smith E (1985) *The Teaching of Thinking*. London: Lawrence Earlbaum Associates.

Nisbet J, Davies P (1990) The curriculum redefined: learning to think-thinking to learn. *Research Papers in Education* **5:** 49-92.

Oakley A (2001) Making evidence-based practice educational: a rejoinder to John Elliot. *British Educational Research Journal* **27:** 575-6.

Palincsar A, Brown A (1984) Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction* **1:** 117-175.

Romney DM, Samuels MT (2001) A meta-analytic evaluation of Feuerstein's Instrumental Enrichment program. *Educational and Child Psychology* **18:** 19-34.

Rosenshine B, Meister C (1994) Reciprocal teaching: a review of the research. *Review of Educational Research* **64:** 479-530.

Shayer M, Adey P (2002) *Learning Intelligence.* Buckingham: Open University Press.

Sternberg RG, Bhana K (1986) Synthesis of research on the effectiveness of intellectual skills programs: snake-oil remedies or miracle cures? *Educational Leadership* **44:** 60-67.

Torrance H, Pryor J (1998) *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom.* Buckingham: Open University Press.

Tymms PB, Fitz-Gibbon CT (2001) Standards, achievement and educational performance, 1976-2001: a cause for celebration? In: Phillips R, Furlong J (eds.) *Education, Reform and the State: Politics, Policy and Practice, 1976-2001*. London: Routledge (pp 157-174).

Webb N, Farivar S (1994) Promoting helping behaviour in cooperative small groups in middle school mathematics. *American Educational Research Journal* **31:** 369-396.

White R, Gunstone R (1992) *Probing Understanding.* London: Falmer Press.

Wilson V (2000) *Education Forum on Teaching Thinking Skills Report.* Edinburgh: Scottish Executive. Available from: http://www.scotland.gov.uk/library3/education/ftts-00.asp.

Wong BYL, Worsham AW, Austin GR (1985) Self-questioning instructional research: a review of effects of teaching thinking skills on SAT scores. *Review of Educational Research* **55:** 227-268.

Wood D, Wood H (1996) Vygotsky, tutoring and learning. *Oxford Review of*

Education **22:** 5-16.

APPENDIX 1.1: Consultation and critical advisory panel

Sue Eagle	Primary School Headteacher, Norfolk
Jane Brooke	LEA Adviser, Cheshire
Dr Carol McGuinness	Professor of Educational Psychology, Queen's University, Belfast
David Thacker	Secondary Headteacher (retired)
Dr Iddo Oberski	Lecturer in Education, Stirling University
Dr Carolyn Tan	Senior Lecturer in Early Years Education, NIE, Nanyang Technological University (NTU), Singapore
Dr William Wu	Co-Director of the Thinking Qualities Initiative at the Centre for Educational Development, Hong Kong Baptist University

APPENDIX 2.1: Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
	We excluded studies which:
1. We included studies which are set in a school or schools and are concerned with any section of the school population (including pupils with special needs).	1. We excluded studies which are not set in a school or schools.
2. We included studies which evaluate the impact of the implementation of thinking skills interventions on teaching and/or learning. Thinking skills interventions are 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 teaching and/or learning. These interventions may be taught as separate programmes or infused into curriculum teaching. Impact includes, for example, pupil and/or teacher motivation and engagement, and/or patterns of classroom interaction, and/or self-regulation and/or metacognitive monitoring and/or pupil attainment.	2. We excluded studies which do not evaluate the impact of the implementation of thinking skills interventions on teaching and/or learning. We excluded studies which do not evaluate programmes or approaches which require the learners to articulate and evaluate the learning strategies that they are using and/or which do not identify specific thinking processes that are amenable to instruction in order to improve teaching and/or learning. We excluded studies which describe pupils' thinking or learning without any evaluation of a thinking skills intervention, strategy or approach. We excluded studies which do not evaluate the impact of thinking skills programmes and/or approaches.
3. We included studies which are concerned with the phases of compulsory schooling (5 –16).	3. We excluded studies which are about pre-school, further and higher education, sixth form (A-level or equivalent).
4. We included studies which contain empirical classroom research with data or evidence (pupil outcomes, classroom processes, teacher role).	4. We excluded studies which are editorials, book reviews, policy documents, resources, guides, manuals, bibliographies, theoretical papers, philosophical papers, unevaluated interventions.
5. We included studies which are written in English[1].	5. We excluded studies which are not written in English.

[1] It was beyond the funding of the review to translate papers from other languages. Though this may have restricted literature identified, every effort was made to identify studies from non-English speaking countries, but published in English.

APPENDIX 2.2: Search strategy

Via BIDS

- British Education Index (from 1986)
- ERIC (Educational Resources Information Center) (from 1985)
- IBSS (International Bibliography of the Social Sciences) (from 1980)
- Ingenta Journals (full text of a large number of journals)
- PsycINFO (extensive catalogue of psychology related publications)

Via Web of Science

- Social Sciences Citation Index (SSCI) (from 1981)

Via FirstSearch

Article1st	Articles and tables of contents of journals in all subjects
Dissertations	Dissertation abstracts, theses in all subjects
ECO	(Electronic Collections Online)
EducationAbs	Education abstracts
PapersFirst	Conference papers in all subjects
Proceedings	Conference proceedings in all subjects
SIRS Researcher	Social sciences
SocialSciAbs	Social sciences
WorldCat	Books and other materials on all subjects

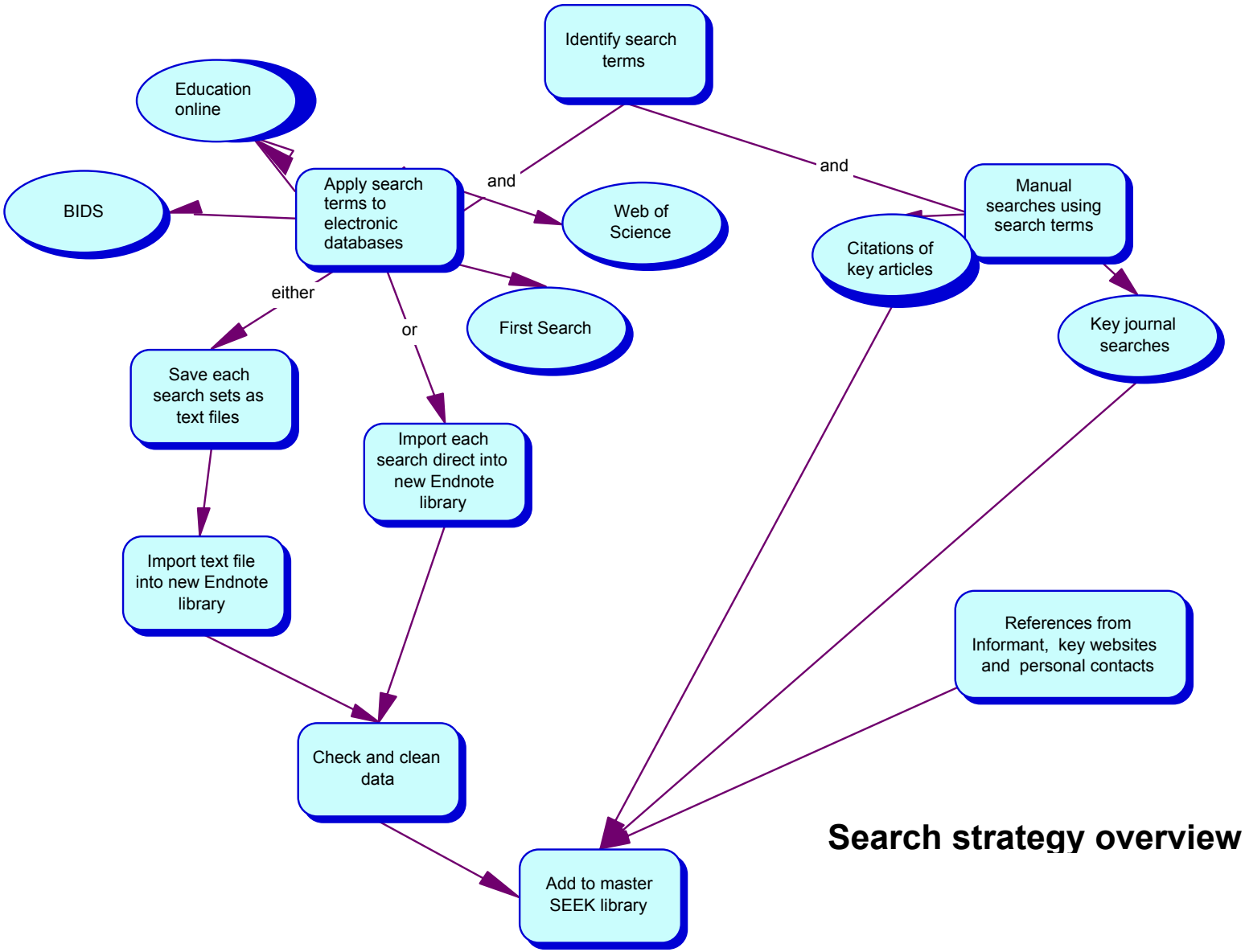
Educati*on-line*	Conference papers and studies

Key search terms applied to each database were:

- thinking, thinking skills, thinking skills program(me), thinking strategies;
- critical thinking, critical thinking skills
- creative thinking skills
- higher order thinking skills (HOTS)
- metacognition, metacognitive, meta-cognitive/ition
- community of inquiry/enquiry/learners
- transfer, near-transfer, far-transfer, bridging, teaching for transfer
- reasoning, argument
- Socratic questioning
- mediated learning

The names of specific thinking skills programmes and approaches and their authors were also applied:

- Instrumental Enrichment / Feuerstein
- Somerset Thinking Skills / Blagg
- Top Ten Thinking Tactics / Lake
- Cognitive Acceleration in Science/Maths/Technology Education (CASE/CAME/CATE) / Adey, Shayer, Adhami
- Philosophy for/with Children (P4C) / Lipman
- Thinking Actively in a Social Context (TASC) / Wallace
- Activating Children's Thinking Skills (ACTS) / McGuinness
- CoRT (Cognitive Research Trust), Six Thinking Hats / deBono
- Storywise, Philosophy with Picture Books / Murris
- Reason!Able / van Gelder


Identify search terms
Education online
BIDS
Apply search terms to electronic databases
and
Web of Science
and
Manual searches using search terms
Citations of key articles
Key journal searches
either
or
First Search
Save each search sets as text files
Import each search direct into new Endnote library
Import text file into new Endnote library
References from Informant, key websites and personal contacts
Check and clean data
Add to master SEEK library
Search strategy overview

APPENDIX 2.3: EPPI-Centre keyword sheet including review specific keywords

1. Identification of report
Citation
Contact
Handsearch
Unknown
Electronic database
(Please specify.) ……

2. Status
Published
In press
Unpublished

3. Linked reports
Is this report linked to one or more other reports in such a way that they also report the same study?

Not linked

Linked (Please provide bibliographical details and/or unique identifier.)
……
……
……
……..

4. Language (Please specify.)

……

5. In which country/countries was the study carried out? (Please specify.)

……
……

6. What is/are the topic focus/foci of the study?
Assessment
Classroom management
Curriculum*
Equal opportunities
Methodology
Organisation and management
Policy
Teacher careers
Teaching and learning
Other (Please specify.)……

***6a. Curriculum**
Art
Business studies
Citizenship
Cross-curricular
Design & Technology
Environment
General
Geography
Hidden
History
ICT
Literacy – first language
Literacy further languages
Literature
Maths
Music
PSE
Phys. Ed.
Religious Ed.
Science
Vocational
Other curriculum (Please specify.)……..

7. Programme name (Please specify.)
……
……

8. What is/are the population focus/foci of the study?
Learners*
Senior management
Teaching staff
Non-teaching staff
Other education practitioners
Government
Local education authority officers
Parents
Governors
Other (Please specify.)……

***8a Age of learners** (years)
0-4
5-10
11-16
17-20
21 and over

***8b. Sex of learners**
Female only
Male only
Mixed sex

9. What is/are the educational setting(s) of the study?
Community centre
Correctional institution
Government department
Higher education institution
Home
Independent school
Local education authority
Nursery school
Post-compulsory education institution
Primary school
Pupil referral unit
Residential school
Secondary school
Special needs school
Workplace
Other educational setting (Please specify.)……..

10. Which type(s) of study does this report describe?

A. Description
B. Exploration of relationships
C. Evaluation
   a. Naturally occurring
   b. Researcher-manipulated
D. Methodology
E. Review
   a. Systematic review
   b. Other review

Please state here if keywords have not been applied from any particular category (1-10) and the reason why (e.g. no information provided in the text).

……
……
……
……
……
……
……..

Keyworded by…… Date…….

REVIEW SPECIFIC KEYWORDS

11. Pupil ages
How old were the pupils?

5-6	11-12
6-7	12-13
7-8	13-14
8-9	14-15
9-10	15-16
10-11	

12. Teaching grouping
How were the pupils grouped for teaching?
Usual class
Set / Banded
Mixed attainment/ability
Not specified
Special group

13. Teaching group size
(NB: This might not be the same as Q15.)
Less than 15
16-25
26+
Not recorded

14. Teacher
Who was the teacher?
Usual teacher
Specialist/Expert
Researcher as teacher (HEI staff)
Teacher as researcher (school staff)
(Please specify.) _____

15. The research sample
How many schools were involved? _____
How many classes? _____
How many teachers involved? _____
How many pupils? ___.*(Intervention/Control)*

16. Does the study sample focus on a particular group of learners?
All
Special group
Gifted and Talented
EAL
Low attainers
Other (Please specify.) _____

17. Thinking skills terms
Mark up to 3 categories for the main focus.
Argumentation
Community of enquiry/learners
Co-operative learning
Creative Thinking
Critical thinking
Decision making
Discussion
Enquiry based learning
Higher order thinking
Logical thinking
Mediation/mediated learning
Metacognition
Problem solving
Reflection
Scaffolding
Self-regulation
Socratic questioning
Systems thinking
Transfer
Others (Please specify.) _____

18. Thinking skills approach
Infused
Enrichment

19. Type of data *(Mark all that apply)*
Quantitative
Qualitative

Interactions
Non-verbal behaviours
Classroom talk/discourse
Pupil attainment
Pupil attitude/beliefs/dispositions
Teacher attitude/beliefs/dispositions
Other (Please specify.)_____

20. Length of intervention *(teaching time)*
_____ lessons/hours (Delete as applicable.)

Not recorded

21. Duration of intervention *(from first lesson to last)*
_____ weeks/months (Delete as applicable.)

Not recorded

22. Method of data-collection
(Mark all that apply.)
Observation
Video
Audio recording
Test *(standardised, criterion referenced, SAT, GCSE, etc*)
Questionnaire/Survey/Rating scale
Interview
Document analysis
Other *(*Please specify.)_____

Notes
v 0.1.1

APPENDIX 4.1: In-depth review studies: topic focus, educational setting and study type

Item ID	Study	What is/are the topic focus/foci of the study?	What is/are the educational setting(s) of the study?	Which type(s) of study does this report describe?
IT14198	Adey and Shayer (1990)	Curriculum Teaching and learning: Study evaluates cognitive acceleration through science education (CASE) (p265 #837). Coding is based on authors' description.	Secondary school Other educational setting middle schools	Evaluation: researcher-manipulated CASE intervention sessions were introduced into classes in eight schools.
IT13612	Blank (2000)	Curriculum Teaching and learning: the impact of structured metacognitive reflection on students' understanding of science by bridging students' science experiences, science ideas and science understanding. Coding is based on authors' description.	Secondary school Junior high school	Evaluation: researcher-manipulated
IT13531	Bowdler *et al.* (1992)	Curriculum Teaching and learning: using Somerset thinking skills as an intervention to promote oracy across the curriculum Coding is based on authors' description.	Primary school Service Children's middle school (p 154)	Evaluation: researcher-manipulated
IT13471	Cardelle-Elawar (1992)	Curriculum Teaching and learning: aspects of metacognition and feedback to students.	Primary school US 6th grade in a public elementary school (p 112)	Evaluation: researcher-manipulated Intervention study with pre and post tests and control group. Follow-up investigation with new control group to eliminate confounding variable in the first phase when the researcher was also the instructor for the experimental group.
IT11895	Chang and Barufaldi (1999)	Curriculum Teaching and learning: looks at the potential of an innovative approach on raising attainment and investigates pupil responses to methods used. Coding is based on authors' description.	Secondary school public Junior high school	Evaluation: researcher-manipulated Four classes participated, taught by the same teacher. Researcher randomly assigned students from two intact classes to the control group and from two intact classes to the experimental group. Cluster randomised controlled trial.
IT13508	Csapo (1992)	Curriculum Teaching and learning: practical forms of instruction infused into subject teaching which improve learners' thinking.	Primary school Secondary school	Evaluation: researcher-manipulated

IT13509	De Corte *et al.* (2001)	Curriculum Teaching and learning: reading. comprehension strategies; learning environment; transfer; self-regulation.	Primary school 'elementary schools'	Evaluation: researcher-manipulated
IT13840	De Konig and Hamers (1999)	Curriculum: reading comprehension Teaching and learning: inductive reasoning.	Primary school 'second grade students (7and 8 years of age) from primary education, living in a backward social-economic home situation' (pp 177-8)	Evaluation: researcher-manipulated
IT13530	Fields (1995)	Curriculum Teaching and learning: pedagogical approach to support Socratic dialogue.	Independent school 'private and state sector' (p 116) Primary school	Evaluation: researcher-manipulated
IT11899	Georghiades (2000)	Curriculum Teaching and learning: conceptual change and the impact of metacognition on the transfer and durability of science conceptions in primary school learners.	Primary school (p 129)	Evaluation: researcher-manipulated
IT13879	Haywood *et al*. (1988)	Curriculum: impact of programme on transfer to the curriculum Teaching and learning: evaluation of Feuerstein's instrumental enrichment programme.	Special needs school 'two public residential schools ... for deaf students' (p 28)	Evaluation: researcher-manipulated The study was funded as a demonstration project, but the researchers were not in a position to control all relevant variables.
IT13562	Head and O'Neill (1999)	Teaching and learning: the impact of mediated learning experiences on learners cognitive development and academic achievement. Other: behaviour Coding is based on authors' description.	Special needs school for students with SEBD	Evaluation: researcher-manipulated
IT11910	Hojnacki and Grover (1992)	Curriculum Teaching and learning: 'the extent of benefits to students when teachers engage in reform of maths teaching program' (p2).	Primary school	Evaluation: researcher-manipulated
IT13839	Kramarski and Mevarech (1997)	Curriculum Teaching and learning: Main focus is the impact of metacognitive instruction on a range of learning outcomes, including transfer.	Secondary school: a junior high school	Evaluation: researcher-manipulated

IT13503	Maltby(1995)	Teaching and learning: The study related to developing more effective thinking strategies in primary age learners. Coding is based on authors' description.	Primary school	Evaluation: researcher-manipulated Children were divided into two groups: one receives strategy training while completing a task; the other played a game in between task completion.
IT13885	Mastropieri *et al.* (1997)	Curriculum Teaching and learning	Primary school	Evaluation: researcher-manipulated
IT13613	Mercer *et al.* (1999)	Assessment: the use of contextually checked quantitative discourse measures; the innovative use of group and individual non-verbal communication. Curriculum: science and other problem-solving activities, using ICT. Methodology: observers present and video-recording in control as well as the experimental classes. Teaching and learning: training teachers in one day to implement the programme by teaching nine lessons; ground rules for exploratory talk established in each class; some of the small group activities (TRAC: talking, reasoning and computers) were computer-based.	Primary school 'state middle schools' (p 101)	Evaluation: researcher-manipulated
IT13887	Ritchie and Edwards (1996).	Teaching and learning	Secondary school	Evaluation: researcher-manipulated The researcher trained three teachers in three different schools to deliver CoRT lessons. Two of the three control classes were in other schools, as there were not enough classes in the same schools with a high proportion of Aboriginal students.
IT13507	Strang and Shayer (1993)	Curriculum: The study focused on teaching a chemistry chemical reactions module (p 323). Teaching and learning: Students were taught in a manner which compensated for the difficulties found previously to be related to learning about chemical reactions (p 323). Coding is based on authors' description.	Secondary school Setting is described as 'a London comprehensive school' (p 323)	Evaluation: researcher-manipulated A class of students was divided into two groups: one group received lessons designed to compensate for previously identified cognitive deficits; the other group received normal lessons taught by the head of science.
IT11897	Verschaffel and De Corte (1997)	Curriculum Teaching and learning	Primary school Elementary boys school in a small Flemish town.	Evaluation: researcher-manipulated

IT13466	Ward and Traweek (1993)	Assessment: assessment by school psychologists of reading Curriculum Teaching and learning: investigates the use of a think-aloud technique.	Primary school: fifth-grade with mean age of 10 years 10 months (p 473)	Evaluation: researcher-manipulated Experimental and control groups with students assigned on the basis of reading test.
IT12365	Wegerif (1996)	Curriculum Teaching and learning: the paper reports on the evaluation of an educational programme that coaches the practice of reasoning together through talk.	Primary school: state middle school	Evaluation: researcher-manipulated: Intervention programme implemented in a classroom; additional classes served as control groups.
IT13838	Zohar (1996)	Curriculum Teaching and learning: transfer of reasoning strategies	Secondary school (p 209)	Evaluation: researcher-manipulated Experience of the students was altered (through the protocols) (p 209); no comparison or control groups, although eighth-grade, post-intervention reasoning compared with ninth-grade baseline.

APPENDIX 4.2 In-depth review studies: aims, theoretical base and research questions

EPPI Item ID	Study(s)	What are the broad aims of the study?	Was the study informed by, or linked to, an existing body of empirical and/or theoretical research?	What are the study research questions and/or hypotheses?
IT14198	Adey and Shayer (1990)	Explicitly stated: To test the hypothesis that it is the process of constructing their own meanings which leads students to the cognitive restructuring responsible for increased scores of the experimental group (p269 #837).	Explicitly stated: The intervention is linked to the work of Piaget and previous work on attempts to accelerate cognitive development (p268 #837). 'It is relevant to note here that the work reported here grew out of results obtained at Chelsea College London in the 1970s based on a broadly Piagetian paradigm' (p2 #747).	Explicitly stated: 'We hypothesise that it is the process of constructing their own meanings which leads students to the cognitive restructuring responsible for increased scores of experimental groups' (p269 #837). Implicit: Can the development of formal operations in average young adolescents be accelerated? (#837) What is the impact of using specially designed lessons (the CASE programme) on pupils' performance on Pigetian tests and science curriculum tests? (#837, #747) Is there any evidence of transfer to other curriculum areas after the programme? (#747)
IT13612	Blank (2000)	Explicitly stated: 'This study proposes a revised learning cycle model, termed the Metacognitve Learning Cycle, which emphasises formal opportunities for teachers and students to talk about their science ideas' (Abstract p486).	Explicitly stated: The study draws upon a number of theoretical models and empirical work in science education. The Science Curriculum Improvement Study (SCIS) Learning Cycle (Figure 1 p488). Good and Lavoie's flexible learning cycle with prediction power (Figure 2 on p488) Good and Lavoie altered the SCIS learning cycle to include prediction. The metacognitive learning cycle (Figure 3 p489) also incorporates Barman's four-phase approach concept assessment/status check, concept exploration, concept introduction/status check and concept application/status check.	Explicitly stated: 'The study presented here asked two questions: Does a revised learning cycle help students make connections between science activities and their science understanding? Will these connections result in higher level of science understanding?' (p490)

IT13531	Bowdler *et al.* (1992)	Explicitly stated: 'Can children be taught to think critically, not incidentally as part of opportunities within specific subjects, but through planned provision that values speaking, listening and reasoning?' (p155)	Explicitly stated: Main references are to policy initiatives in education, such as development of the National Curriculum and the National Oracy Project, as well as research that that has identified the need for problem-solving and improved oracy and studies from the Assessment Performance Unit. Also work by Wilkinson (1971) on the unrealised potential of pupils to use language and Blagg *et al.* (1988) in the materials for the programme.	Explicitly stated: 'Can children be taught to think critically, not incidentally as part of opportunities within specific subjects, but through planned provision that values speaking, listening and reasoning?' (p154)
IT13471	Cardelle-Elawar (1992)	Explicitly stated: The aim is to investigate the effects of metacognitive instruction in mathematics on low ability sixth grade students (abstract).	Explicitly stated: The rationale and procedure for the investigation were based on existing research into: (a) metacognitive theory (citing several key theorists), (b) Mayer's (1987) model for solving mathematics problems, and (c) feedback tailored to students' individual needs (p 110).	Explicitly stated: (1) What is the effect of this instruction on students' mathematics achievement? (2) What is the effect of this instruction on students' attitudes toward mathematics? (p 112)
IT11895	Chang and Barufeldi (1999)	Explicitly stated: To investigate the effects of a problem-solving-based instructional model on the achievement and alternative frameworks of ninth grade earth science students in Taiwan. In addition, students' opinions toward the problem-solving-based instructional method are investigated.	Explicitly stated: Intervention used search, solve, create and share (SSCS) programme devised by Pizzini *et al.* (1988).	Explicitly stated: To see if problem-based instruction has a positive impact on achievement and alternative frameworks - particularly application of conceptual understanding. Also to look at students' views on the problem-based instruction method.
IT13508	Csapo (1992)	Explicitly stated: To examine the effectiveness of teaching materials designed to stimulate the development of thinking abilities (logical, combinative and systematizing operations) of grade 4 and grade 7 pupils through their infusion in the grammar and science curriculum (p145).	Explicitly stated: Extensive reference made to theoretical backgound - e.g. Fischer's skill theory (Fischer, 1980; 1992) and experiential structuralism (Demetriou and Efklides, 1988; Demetriou, Gustafsson, Efklides and Platsidou, 1992) as well as empirical studies of training in binary operations of propositional logic or combinatorial reasoning or operational thinking in the context of	Implicit: 'On the basis of this earlier work the teaching material of some school subjects was analysed and methods devised for the improvement of children's operational abilities. The present one year experiment was designed to study the changes caused in the student's [sic] cognition by these method' (p145).

			school subjects (e.g. Fishbein, Pampu and Minzat, 1970; Siegler, Liebert and Liebert, 1973; Collis, 1980 in maths; and Jurd, 1978, in history). Links with Shayer and Adey's CASE intervention is also mentioned. Also builds on earlier work (Csapó, 1988) (p144-5).	
IT13509	De Corte *et al.* (2001)	Explicitly stated: The design experiment aimed at developing, implementing and evaluating a research-based, but also practically applicable learning environment for enhancing skilled strategy use when reading a text with four experimental classes of upper primary children (p531).	Explicitly stated: Reading comprehension research into text-related and student-related factors (e.g. Hiebert & Raphael, 1996). Reading mastery and use of good strategies is also a major component that distinguishes poor and skilled readers (Brand-Gruwel, Aarnoutse and Van den Bos, 1995; Cross and Paris, 1988; Palincsar and Brown, 1984; Paris and Myers, 1981; Pressley, Goodchild, Fleet, Zajchowski and Evans, 1989). 'In the USA this [the lack of strategy instruction] was already observed and reported in the 1970s by Durkin (1978, 1979; see also Paris and Oka, 1986); classroom observations showed that hardly 1% of the instruction time addressed comprehension strategies' (p 532). Also research into reciprocal teaching (e.g. Palincsar and Brown, 1989; Rosenshine & Meister, 1994) transactional instruction of comprehension strategies (Brown, Pressley, Van Meter and Schuder, 1996; Pressley *et al.*, 1989; Pressley *et al.*, 1992); and metacognition (Garner, 1987).	Explicitly stated: 'The major hypothesis of the study was that the learning environment would have a significant positive impact on pupils' adoption and use of reading comprehension strategies' (p 541). We hypothesised that the learning environment would have a favourable effect on the experimental pupils' general reading comprehension ability as well. Therefore, we predicted that at the post-test stage the children from the experimental classes would show significantly more progress on a standardised reading comprehension test than the pupils of the control classes (p 541). We also expected that, as a result of the learning environment, the experimental pupils would make more use of the strategies than the control children outside the context of reading comprehension instruction. Therefore, we predicted that the experimental classes would significantly outperform the control classes on a transfer test administered after the intervention (pp 541-2).

IT13840	De Konig and Hamers (1999)	Explicitly stated: Can a reading comprehension programme devised by the authors be implemented in the classroom? Does the programme significantly improve reading comprehension? Does the programme significantly improve inductive reasoning?	Explicitly stated: The authors drew on existing lines of research dealing with 'teaching group and row schemes' (p169), top-down (metacognitive) and bottom-up (performance) processes in reasoning and in reading, especially the work of Klauer. They sought to make 'intersubjectivity and active participation in the classroom dialogue' and linguistic and text structures explicit, drawing on the work of Piaget and Vygotsky among others (pp 170-2).	Explicitly stated: 'The following issues were investigated: (a) Can the programme be implemented in the classroom? (b) Does the reading comprehension of students who follow the programme improve to a greater extent than that of students who do not follow the programme? and (c) Does the inductive reasoning skill of students who follow the programme improve to a greater extent that that of students who do not follow the programme?' (pp177 and 179).
IT13530	Fields (1995)	Explicitly stated: 'The aim of the study was to ascertain whether, after undergoing a philosophy programme based on the materials outline below, young children could be emergent philosophers. And if so whether there were any significant changes in children's academic achievement, and reasoning skills...' Abstract (p115)	Explicitly stated: as above. Also refers to Lane and Lane (1986), an earlier evaluation.	Explicitly stated: 'The aim of the study was to ascertain whether, after undergoing a philosophy programme based on the materials outlined below, young children could be emergent philosophers. And if so whether there were any significant changes in children's academic achievement, and reasoning skills...' (Abstract, p115).
IT11899	Georghiades (2000)	Explicitly stated: Its main scope was to study the notions of transfer and durability of newly acquired scientific conceptions, and to investigate any positive impact of metacognitive instruction on the two areas (p 130).	Explicitly stated: The study draws upon four major areas of theory and research: (1) research on conceptual change learning (CCL) (pp119-22), (2) research on transfer and durability of scientific conceptions (pp122-6), including research into cognitive acceleration through science education (CASE) by Adey and Shayer, (3) research on metacognitive instruction (pp126-9), and (4) empirical data, based on the theoretical background of the above to support this thesis.	Explicitly stated: 'Its main scope was to study the notions of transfer and durability of newly acquired scientific conceptions, and to investigate any positive impact of metacognitive instruction on the two areas' (p130).

IT13879	Haywood *et al.* (1988)	Explicitly stated: 'The principle purpose of this study was to assess the relative effects of a particular programme of cognitive education (IE) on the cognitive functioning and general intellectual levels of deaf students in a residential educational environment' (p28).	Explicitly stated: The authors refer to research and development work by Feuerstein and his colleagues, together with evaluation studies of instrumental enrichment (pp25-8) and reviews of these as well as and other work in cognitive education (p23-4).	Explicitly stated: 'Since we were interested in assessing the effects of the IE on general intellectual functioning or scholastic aptitude, mastery of the program itself, transfer of some principles to new problems, and to some extent, school achievement, specific tests were selected for each of these areas' (p28). Implicit; Study also evaluated the impact on learners' attitudes towards IE, motivation to learn and views of themselves as learners.
IT13562	Head and O'Neil (1999)	Explicitly stated: To evaluate the impact of Feuerstein's instrumental enrichment programme on academic ability and behaviours of a groups of students with social, emotional and behavioural difficulties.	Explicitly stated: General literature on thinking skills approaches (such as Ashman and Conway 1993) 'Within the SEBD sector, there has been, traditionally, an emphasis on eradicating unacceptable behaviours and replacing them with newly learned ones, an approach which does not take into consideration the child's capacity to think and reason' (p122). Specific literature on IE, such as Feuerstein *et al.* (1980), due to its 'objectives of helping pupils to learn how to learn, of helping them to define and solve problems whilst restraining their impulsivity, of assisting them to address their behavioural difficulties' (p123). Literature on mediation (e.g. Burden and Florek, 1989).	Implicit: What is the impact of the use of IE on the academic performance and behaviour of students?
IT11910	Hojnacki and Grover (1992)	Explicitly stated (p2): 'Does empowering teachers empower students, and if so, how? ... The purpose of this paper is to study preliminary findings about the extent of benefits to students when teachers engage in the Thinking Mathematics (TM) program ... The goal of the	Explicitly stated: Theoretical (p3) section entitled 'Teachers change as a context for student change' cites theoretical research addressing 'the difficultly of making lasting change (Cohen, 1987; Cuban, 1990; Passow, 1986). In addition: 'Teachers are asked to commit to reform without sufficient support from systems that, in fact, are set up to punish risk-	Explicitly stated (p5): 'Empowerment of both teachers and students was hypothesized to result (from the program) because teachers were involved in all phases of development and in dissemination, and attempts were made to assess the impact of the project on the students involved.'

		project was to develop more effective ways of disseminating new knowledge about mathematics instruction and learning ... The present paper examines the ways in which this particular teacher-researcher collaboration impacted upon the teachers and students.'	taking behaviour more often than to reward them (e.g. Silberman, 1970; Smith 1991). Empirical research (pp3-4) that (1) provides a rational for involving teachers in 'decision making that influences their working conditions. Romberg (1986), (2) provides examples of successful programs described in a book by Lieberman (1986) and (3) looks at the involvement of empowering teachers through the implementation of programs that they have been involved in researching (Schwartz, 1986, p204).	
IT13839	Kramarski and Mevarech (1997)	Explicitly stated: To compare the performance of students who learned to construct graphs in the same problem-solving Logo environment that was either embedded, with or with no, metacognitive training To examine the difference in students' cognitive metacognitive behaviours under the different conditions. (p425 – abstract)	Explicitly stated: Paper reviews two main strands of research into the use of computer learning environments and metacogntive training and research into cognitive learning particularly in mathematics and the effect of metacognitive strategies. The authors and others have used the SOLVE guided questioning strategy used in this study and evaluated aspects of computer-based learning environments, but not carried out a controlled experiment to isolate the relevant variables.	Explicitly stated: Students who learn to construct graphs in a problem-solving Logo environment and use the SOLVE metacognitive strategy will do better and will be more able to reflect on their learning than the control group who work in the same Logo environment for the same duration of time and with the same teacher, but without the SOLVE strategy.
IT13503	Maltby (1995)	Explicitly stated: To see whether students improved in their thinking strategies as a result of training in the use of the TASC method of problem –solving.	Explicitly stated: The study is linked to theory explaining the TASC method which is being developed by the author and a colleague. The study does not draw on existing empirical evidence.	Implicit: No specific research questions are provided, but it was implicit that the investigation is trying to ascertain whether students exposed to TASC would be more successful at a problem-solving activity.
IT13885	Mastropieri *et al.* (1997	Explicitly stated: 'The purpose of the investigation was to determine whether students with mental retardation might benefit from computer-assisted instruction that includes a cognitive problem-	Explicitly stated: Work cited on the value of problem-solving strategies by the authors (Shiah, Mastropieri, Scruggs and Fulk, 1995) and by others working in the field of learning disabilities. Drill and practice CAI also referenced (p158).	Explicitly stated: Will students with 'mild mental retardation' (i.e. with IQs around 70) benefit from computer-assisted instruction in answering simple maths word problems? After instruction, will they perform on paper-and-pencil tests as well as on the

		solving strategy, animated depictions of the problems in a tutorial format and minimised or eliminated reading requirements' (p158).		computer?
IT13613	Mercer *et al.* (1999)	Explicitly stated: 'to improve the quality of children's reasoning and collaborative activity by developing their awareness of language use and promoting certain 'ground rules' for talking together' 'The aim of the research described here was to develop and evaluate a teaching programme for 'scaffolding' children's effective use of language as a tool for reasoning...' (p98) To explore and evaluate three hypotheses: '(1) that using exploratory talk will help children to reason together more effectively and this can be shown by an improvement in their scores when they jointly tackle the problems of a test of reasoning, (2) that children's use of exploratory talk in joint classroom activities can be increased by using specially-designed teacher-led and peer-group activities, and (3) using exploratory talk for joint reasoning will help children develop better ways of using language as a tool for reasoning individually. This will lead to improvements in the scores children achieve when working alone on a reasoning test' (p 98).	Explicitly stated: It was informed by sociocultural theories of intellectual development and by Vygotsky's idea of the teacher providing 'scaffolding'. It was also influenced by research in which attempts have been made to teach transferable skills in using language for reasoning.	Explicitly stated: To explore and evaluate these hypotheses: (1) that using exploratory talk will help children to reason together more effectively and this can be shown by an improvement in their scores when they jointly tackle the problems of a test of reasoning, (2) that children's use of exploratory talk in joint classroom activities can be increased by using specially-designed teacher-led and peer-group activities, and (3) using exploratory talk for joint reasoning will help children develop better ways of using language as a tool for reasoning individually. This will lead to improvements in the scores children achieve when working alone on a reasoning test.

IT13887	Ritchie and Edwards (1996)	Explicitly stated: To find out if the teaching of general thinking skills using DeBono's cognitive research trust (CoRT) programme promotes creative thinking and improves the academic performance and internal locus of control of urban Aboriginal children.	Explicitly stated: References cite work by Taylor and de Lacey on teaching divergent thinking to Aboriginal children and a number of research evaluations of DeBono's CoRT thinking programme	Explicitly stated in hypotheses: Hypothesis 1: The creative thinking of urban Aboriginal students will be enhanced as a result of their participation in a twenty lesson CoRT program. Hypothesis 2: The scholastic aptitude of urban Aboriginal students will be enhanced as a result of their participation in a twenty lesson CoRT program. Hypothesis 3: The school achievement of urban Aboriginal students will improve as a result of their participation in a twenty lesson CoRT program. Hypothesis 4: Urban Aboriginal children will demonstrate an improved application of CoRT thinking approaches as a result of their participation in a twenty lesson CoRT program. Hypothesis 5: The self-concept as a thinker of urban Aboriginal students will not change as a result of their participation in a twenty lesson CoRT program. Hypothesis 6: The locus of control of urban Aboriginal students will become more internal as a result of their participation in a twenty lesson CoRT program (p63).
IT13507	Strang and Shayer (1993)	Explicitly stated: The study considered the application of a thinking skills programme - Feuerstein's Instrumental Enrichment - to the teaching of chemistry to 14 year-olds (p319).	Explicitly stated: The study is linked to empirical and theoretical work regarding thinking skills strategies and their evaluation.	Explicitly stated: The study hypothesis was that teaching the chemical reactions module, using methods which compensate for the cognitive deficiencies found previously to be associated with learning about chemical reactions, will result in a better understanding of chemical change (p323).
IT11897	Verschaffel *et al.* (1997)	Explicitly stated: 'To test the hypothesis that it is feasible to develop in pupils a disposition toward (more) realistic mathematical modelling ... by immersing	Explicitly stated: Theoretical research around mathematical modelling (p577). 'As several authors have stressed, this process of mathematical modelling and problem solving has to be considered cyclic,	Explicitly stated (pp587-9): Very clear presentation of hypotheses and other research questions. Hypothesis 1: 'We hypothesized that the pupils would ... demonstrate a strong tendency to exclude real world knowledge when

		them in a classroom culture in which word problems are conceived as exercises in mathematical modelling, with a focus on the assumptions and the appropriateness of the model underlying any proposed solution' (p577).	rather than linear, progression from givens to goals (Burkhardt, 1994; Greer, 1997; Lesh and Lamon, 1993)'. Empirical research to demonstrate (a) the tendency of pupils towards routine-based unrealistic modelling and (b) the extent of the problem (pp578-9). (a) 'According to many authros (Davis, 1989; Freudenthal, 1991; Greer, 1993, in press; Gravemeijer, 1994; Kilpatrrick, 1987; Nesher, 1980; Reusser and Stebler, 1997; Schoenfield, 1991; Selter, 1994' Silver *et al.*, 1993; Teffers and De Moor, 1990; Verschaffel *et al.*, 1994, in press), children's strong tendency to suspend real-world knowledge and realistic considerations in solving real word problems develops as a result of schooling' (p579). (b) For example, a study in the US where only 24% of a national sample was able to solve a problem involving mathematical modelling for a real world problem (p578).	confronted with the problematic versions of the problems'. (The pre-test contained a number of easily solvable maths problems and some that were problematic so that they required an element of 'real world modelling'.) Hypothesis 2: 'We also hypothesized that the experimental programme would have a positive effect on pupils' disposition towards realistic modelling and interpretation of arithmetic word problems' (results of post-test). Hypothesis 3: 'We hypothesized that the positive effect of the experimental program would also transfer to the P items representing the same mathematical modelling difficulties as those encountered during the training, but embedded in considerably different text' (near-transfer). Hypothesis 4: 'A final hypothesis was that the positive effect of the experimental programme would be lasting' (tested by the retention-test). Other research questions: (1) Was the experimental programme equally effective for pupils with different levels of mathematical ability (tested through analysis of variance, using measures for strong, average and weak maths students in the experimental class)? (2) Was the programme equally effective for the five types of mathematical modelling difficulties in the programme? (3) What particular difficulties did the teacher and the pupils experience with respect to the implementation of the three pillars of the experimental programme?

IT13466	Ward and Traweek (1993)	Explicitly stated: 'To demonstrate the application of a metacognitive strategy - the think-aloud technique - in each of the school psychologists' roles of assessment, intervention, and consultation' (p469).	Explicitly stated: Refers to general literature on metacognition (e.g. Campione and Brown; Flavell - pp469-70) as well as specific literature relating to metacognition and reading (p 471) and assessment issues in metacognition and reading (pp472-3).	Explicitly stated: 'To evaluate the effect of using think-aloud procedure on their reading comprehension of school-age children' (p473).
IT12365	Wegerif (1996)	Explicitly stated: To explore and assess an educational programme based on socio-cultural theory that incorporated the use of computers to coach the practice of reasoning together through talk (p53).	Explicitly stated: The study is linked to a theoretical framework based on socio-cultural theory with reference made to the work of Vygotsky and Habermas (p52). Reference is also made to existing thinking skills programmes, such as Lipman's 'Philosophy for children' and the Oxfordshire thinking skills programme (p52). A number of references are also made to the nature and significance of exploratory talk (p51) and the role of computers in supporting exploratory talk (p52).	Implicit: Some information is given as to specific research questions but no specific questions are stated. It would appear that the overall research question was to consider the impact of the programme as a whole in integrating exploratory talk into the curriculum areas of citizenship and science.
IT13838	Zohar (1996)	Explicitly stated: 'The purpose of this article is to investigate three learning environments focusing on strategic aspects of variable control. The article describes students' initial thinking strategies, and investigates the influence of the learning environment on three dependent variables: (a) progress in students' thinking strategies following instruction; (b) transfer of newly acquired strategies to a new problem, taken from a new biological topic; and (c) retention of acquired thinking strategies across time' (p206).	'A constructivist approach was assumed in the course designing the learning activities. Such an approach suggests that fruitful teaching of scientific reasoning should include challenges to students' initial non-scientific reasoning strategies and opportunities for the development of more accurate thinking strategies' (p206). 'The idea for the learning environments described here originated in the tasks used in a set of theoretical studies which were designed to investigate the development of scientific reasoning skills (Kuhn *et al.*, 1992, 1995). Progress in subjects' thinking skills was observed in these studies. This progress suggested that the tasks used in those studies might be applied to practical educational use.	Explicitly stated: To investigate the impact of learning environments, which encouraged cognitive conflicts with students' initial non-scientific reasoning strategies prior to a structured instruction of variable control, on the progress, transfer and retention of the learners' higher-order thinking strategies.

APPENDIX 4.3 In-depth review studies: findings and conclusions

Item ID	Study(s) (Review no.)	What are the findings of the study as reported by authors?	What do the author(s) conclude about the findings of the study?
IT14198	Adey and Shayer (1990)	Overall, the experimental group made gains in levels of cognitive development which were statistically greater than those made by the control group over the two-year period of the intervention. Yet the relative gain was only 0.21 levels or 0.20 standard deviation (p272 #837). All the extra gain by the experimental group as compared with the control group appears to be concentrated in the 12+ age group....Only the boys in the 12+ age group appear to have been affected by the intervention (p273 #837). All subgroups of the 12+ experimental show substantial gains over controls, regardless of the starting level. Among the 11+ experimental subgroups, there are no gains whatsoever in relation to the controls (p274 #837). There are suggested results that gains made are not the result of direct teaching of the subject of the tests but reflect deeper changes in cognitive structure (p278 #837). No differences are shown in performance on the science assessment (p280 #837). One year after the end of the intervention, none of the experimental groups showed any overall difference from the control groups in these measures of cognitive development (Piagetian reasoning tasks). The gains that were present immediately after the intervention apparently dissipated (p16 #747). Delayed science assessment: The 12+ boys showed a very strong effect and the bimodality noted previously. The 11+ girls also showed a significant effect, confirming the suspicion raised about an effect with this group (p18 #747). GCSE science: The 12+ boys averaged one grade higher than the controls, after individual pre-test differences are taken into account. The 11+ girls also showed a significant effect with their grades 2/3 of a standard deviation better than the controls (p18 #747). GCSE maths: The results follow a similar pattern to those in science, with significant effects achieved in the 12+ boys and 11+ girls (p 9 #747). GCSE English: All groups apart from 11+ boys showed a significant effect.	The size of the experimental effect for the boys in the intervention classes starting at 12+ ...suggests to us very strongly that the CSMS British population survey data do not represent an unalterable feature of human development (p280 #837). A major decision point reached by the project has been resolved in support of using schooling to alter cognitive development, against a notion of merely adapting to existing norms. However, while we not believe that schooling can alter development, it is clear that this project is at best a first step to enabling schooling to change to do it (p280 #837). Evidence has been presented of substantial and long-lasting effects on general academic achievement of a cognitive acceleration program that concentrates on cognitive conflict, metacognition, and bridging and that uses the schemata of formal operations as a framework for the development of activities. In particular, boys starting a 12+ and girls starting at 11+ showed strong and actually increasing effects over the period following the intervention programme (p27 #747).

IT13612	Blank (2000)	The curriculum intervention did not result in any significant differences in the level of ecological understanding across treatment groups: no significant difference between either the pre- or post-test mean scores was observed between the two treatment groups. The SCIS class increased from a pre-test mean of 51.0 (SD=15.1) to a post-test mean of 79.95 (SD=12.09); the MLC class increased from a pretext mean of 58.4 (SD=10.8) to a post-test mean of 81.95 (SD=10.18). No significant difference in ecology understanding across the two treatment groups either before or after the curriculum intervention: pre-test: F= 3.3585; p-value=0.0743; post-test=0.3398; p-value=0.5632. But there is significant difference between the SCIS and the MLC students on both administrations of the delayed ecology assessments (Tables 2-5). 'This difference suggests that true restructuring or meaningful learning may have occurred in the MLC classroom to a greater extent than in the SCIS classroom. This result would be consistent with Gauld's (1986) findings that true restructuring occurs only after reflective practice or metacognition on the part of the student' (p493). Student dialogue differed across the two classrooms. The MLC discussions were particularly engaging and thoughtful, while in the SCIS classroom, no consistent struggle was observed and they were unable to explain or support their understandings correctly or incorrectly. 'These observation suggest that a strong relationship exists between high-quality student dialogue and strong retention scores….further research is needed to confirm this relationship' (p502).	'While the (MLC) model is not a warranty for student understanding it can be a useful guide for helping teachers make the most of science learning experiences. It can help cultivate classrooms in which students are asked to think explicitly about their ideas and to value the reasons behind their conceptions. ... By providing a formal interaction between students' science ideas, experience, and understanding through the use of a metacognitive learning cycle, students did not necessarily gain a greater content knowledge of ecology, but they did appear to have permanently restructured their understanding of ecology. Further research would be needed to generalise this conclusion beyond the parameters of this study.' (p503)
IT13531	Bowdler *et al.* (1992)	Attitude questionnaire: Four items 'showed a positive shift in favour of the STS class over the control class'. Explanation (+10 points), less impulsive (+6 points) attention and concentration (+10 points), confident to contribute ideas (+25 points). Bristol Social Adjustment Guide: gains in •facing new learning tasks •having better strategies •asking teacher's help •answering questions •ways with other children BAS verbal fluency 'The results indicated that all the students gained in verbal fluency on the six measures employed.' Cognitive Instrument/Oral Argument: 'The results indicated that the pupils' use of speculative thinking and argumentative	'The introduction of STS modules into the curriculum of these 9-10 year old children has led to marked improvements in their oracy skills. The children's use of higher-order language skills and cognitive strategies, their social awareness, and attitudes to new work had all developed apace. They have learnt the value of enjoying argumentative discourse and appreciated some of its importance for the development of their own thinking.' (p167)

		terms has increased'. Results from informal assessments are presented as enthusiastic quotations from pupils with two examples of transfer with teachers' positive comments.	
IT13471	Cardelle-Elawar (1992)	Phase 1 (p116 Table 1) and studying of ANCOVA analysis shows that there were significant differences between the pre- and post-tests for the experimental group and that they scored higher than the control group on all the measures post test. Phase 2 Table 1 and ANCOVA scores (p117) and Table 2 (p118) show that the experimental group had higher scores than the control group in all post-test measures and these are statistically significant. Discussion (p119) refers to analysis of improvements in aspects of problem-solving and says that students have got better at understanding how to approach a problem, identifying the appropriate schema for organizing the information, recognizing that there may be more than one way to solve a problem and in verifying their solutions. Improvements are interpreted as arising from increasing their linguistic comprehension of key words and sentences leading to heightened concentration and reduced impulsivity and an ability to reflect on their own thinking (p119).	Recommendations offered for teachers trying to help under performers in maths: Focus on the individual behaviour of pupils and not on their labels (low attainer, minority, etc.) and address their uniqueness in terms of both strengths and weaknesses. Low performing students need a supportive atmosphere in which mistakes lead to positive feedback and direction. Classrooms need to be very structured with tasks broken down into small doses of learning and mastery. Problem presentation should be well organised and structured, and a great deal of interaction between teacher and student is required so that understanding of what is required can be mediated and constructive feedback given. Low attaining students have metacognitive potential when stimulated by explicit instruction and given encouragement and tools for reflection.
IT11895	Chang and Barufaldi (1999)	Problem-based instructional method produced significantly greater achievement of ninth grade earth science students than did conventional teaching method, especially at the application level (Table 3, p381). It also statistically supported that students in experimental group experienced a significant conceptual change (Table 4, p381) even though both the intervention and traditional methods were successful in modifying AFs. Student opinion in experimental group showed no particular perceptions toward the teaching approach but did express advantages in terms of helping them develop analytical and observing skills and improving their thinking skills (pupil quotes on pp381-2). Pupils had concerns of value of the approach, given the pressure of exams.	These findings support the notion that teachers need to encourage students to develop their process skills as early as possible in the educational system in order to promote science learning in the classroom. Science teaching should encourage students' own meaningful learning and process skills. This clearly implies that application of science concepts rather than rote-learning or memorisation should be emphasised and stressed in the science classroom to help children to develop higher-level thinking skills. Teachers should also identify the prior knowledge or concepts that students possess before instruction which might facilitate students' learning of science concepts and promote conceptual change.
IT13508	Csapo (1992)	The structured task systems in the intervention had different effects across the various training contexts. Systematising ability: develops rapidly across the age range studied but is not accelerated by the intervention and tasks designed to improve it had only a weak effect on other abilities. Logical operations: develop slowly but	Cognitive acceleration in a school year is limited. After a certain number of exercises, the effect can be exhausted and may be counter-productive. Application should be planned judiciously to maximise effect. In regular school practice, ten to twenty exercises in the same group of operations might be enough per school

		can be improved at the younger age. Intervention also had a significant effect on the other abilities at an older age. The intervention did improve thinking but this change was not detected by the measurement used (possible ceiling effect) suggesting that improvement in advanced formal thinking may not be characterised in terms of formal logic. Combinative operations: develop at an intermediate pace and the intervention achieved considerable acceleration at both ages. Differences between these 3 abilities are greater than their similarities. Effects less in groups where the intervention applied in more than 1 subject. Enrichment materials to improve logical and combinative ability are worthwhile but less certain for systematizing ability which appears already well covered by normal teaching and is not significantly improved by the tasks. (p157).	year and their total number must not exceed thirty. Consequent enrichment of the teaching material with this activity in several school subjects over subsequent years may be the best way to apply these exercises in educational practice (p157). '...children with less mature operational thought benefit more from the training chosen than do the more mature children. Training might be more economic if these less mature students were targeted, but this requires more complicated classroom procedures' (pp157-8) or a co-ordinated school approach (p157).
IT13509	De Corte *et al.* (2001)	Significant gains made in strategy use, maintained in the retention test (Figure 4, p547) but not the reading comprehension test (Figure 6, p550) or the reading attitude test (Figure 7, p551). Transfer demonstrated for experimental group. Lower-scoring pupils made greater gains in strategy use than higher scoring ones on the post-test, although not on the retention and transfer tests (Figure 5, p549) More comprehension strategies reported by experimental group (Table 5, p553). Fidelity of implementation high except for the role of the learner as group leader (52%). Teacher content-oriented interventions were largely directed at control of strategy implementation (53%). In 16% of these interventions, teachers provided the answers and in 24% stimulated pupils' thinking. Teachers were generally positive about the intervention (p554).	The major results can be summarised as follows. There was significant gain by the experimental group on the Reading Strategy post-test and retained on the retention test. This effect was greatest for the low-scoring pupils. The experimental classes scored higher on the standardised Reading Comprehension Test than the control group, but this difference was not statistically significant. There was very little change either for the experimental or control groups on the Reading Attitude Scale. But the experimental children scored high on this scale at the pre-test stage and the intervention was quite short for achieving a significant attitude change. The results on the Transfer Test showed that the experimental pupils were able to apply the reading comprehension strategies acquired in the learning environment successfully in a different context from reading comprehension instruction. The analysis of the interviews with a subgroup of pupils of all the participating classes revealed that after the intervention the children of the experimental classes reported significantly more strategy use than those in the control condition (pp554-5).
IT13840	De Konig and Hamers (1999)	'The results of the observations showed that the teacher was able to deal with the main didactical requirements of the programme. However it was not easy for her to continually look past the familiar reading comprehension domain to the underlying mental processes' (p180). The	The results show that an 'across the curriculum' approach to teaching inductive reasoning can be effective when tied to the instructional domain of reading comprehension and using tasks which are 'recognisable from daily situations' (p182). It is 'not

		experimental programme produced the expected positive learning effects in reading comprehension, as well as a 'far transfer' effect in inductive reasoning.	possible to say what particular aspect of the training was helpful' (p182).
IT13530	Fields (1995)	'The results of Lane and Lane (1986) is [sic] supported by the results of the current study as there was a significant difference between the experimental and control groups (p0.05 df8) in all three reasoning tests used' (p117). 'The results of this study do not support Lipman's assertion that reading ability and fluency is increased' (p117). 'It is interesting to note that while there was no significant difference found between the groups or between the schools or the sex of the subject using the Standard Attainment Mathematics Task profiles once mathematical problems were embedded within a narrative there was a significant difference in the results for both experimental groups' (p117). 'There was no significant difference between the performances of the experimental groups and control groups in the academic subject English' (p117) however, 'the independent observer' recorded that there had been a discernible increase in 'displayed self-confidence' of some subjects (pp117-8). Based on the teachers' appraisal checklist, findings indicate more *motivation *curiosity *commitment *concentration *increased ability to stay on task *clarity in explanations *greater ability to reason and communicate reasoning coherently *'a noticeably and definably [sic] measure of social skills and general behaviour, such as appropriate/realistic goal setting, planning and self-monitoring *heightening of the child's self image and view of themselves as thinkers (p118). In part II, the transcript is claimed to show emergent philosophical enquiry.	'In conclusion, therefore, the empirical data results of this suggests that there is significant evidence that using teaching materials specifically designed for philosophical inquiry has effect on specific areas of learning and behaviour' (p118). From the evidence presented in parts I and II above, this empirical data study found that (a) children can and do philosophise and have profound thoughts..., (b) children who engaged in philosophical enquiry showed improvement in specific academic achievement, and (c) there is identifiable oral and written evidence to support the thesis that young children have sophisticated reasoning abilities which allows for them to be classed emergent philosophers' (p128).
IT11899	Georghiades (2000)	Metacognitive instruction is feasible with Year 5 students (p130). Metacognitive instruction is best given in small groups, rather than as whole-class instruction (p130). Children who received metacognitive instruction performed better (p131). Tables on pages 132-136 show positive impact of metacognitive instruction on transfer and durability of conceptions.	'To recap on the above study, metacognitive instruction has given signs of long-term improved performance in general, and of increased ability to utilize learned material in particular. However, the exact extent of this improvement and its significance in both statistical and educational terms is to be discussed in the near future.' (pp133-4)
IT13879	Haywood *et al.* (1988)	As anecdotal quotations in the text (pp38-9, in textual form and in tables 2, 3, 4 and 5 (pp31-4). Significant gains were made by the experimental group on 6 out of 7 of the mastery, transfer and IE vocabulary tests. The IE students gained significantly more than control students on all three	'The required demonstration that the students did indeed learn a significant amount of what was taught in IE can be seen in their performance on the mastery and transfer tests' (p35). Improvement in general intellectual functioning and scholastic aptitude was

		aptitude tests (Raven's Matrices and PMA reasoning and spatial relations). The control students gained significantly more than the IE students on Stanford Maths applications and (although the differences were not statistically significant) did better in absolute terms on 7 out of the 10 other measures of scholastic achievement. The multiple discriminant analysis correctly identified 75% of the high and low gain groups, using four steps: sex, group (IE versus control), number of IE instruments taught and race. It is not stated how each variable was related to gain. Informal observations by the teachers and investigators and comments gathered by one teacher suggested that IE helped improve motivation to learn, attitude towards learning, and attitude towards the self as a learner. Anecdotal studies from the teachers indicated improvements in social behaviour as well as in academic performance and generally enthusiastic responses.	accomplished and seems to have been a function of the instrumental enrichment treatment. It 'seems possible that there were positive changes in students' motivation to learn and in important attitudes toward learning and toward themselves as learners, and that these changes might have catalyzed the measured changes in general intellectual functioning' (p 39). Improvement of learning in curriculum content areas as a result of the improved cognitive processes 'has not been demonstrated' (p 35).
IT13562	Head and O'Neill (1999)	Table 1 (p 125) shows that all except one of the six pupils who completed the IE course showed an overall decrease in cognitive deficiency by a significant factor. The mean factor of change was 9% and the average among those showing a decrease in cognitive deficiency was 12%. Classroom teachers had noticed significant qualitative changes in the behaviour exhibited by the students in the experimental group. 'We felt that the group with taught IE had gained significantly both academically and socially. The control group became increasingly troublesome as they came under pressure to produce work for Standard Grades. .. None of the control group gained the maturity to go about arranging their future in the same way as the members of the IE group…There was, especially, a decrease in impulsivity among the IE students, and this was the distinguishing factor between the groups…Academically the control group were outperformed by the IE group, although it had been predicted, prior to the IE course, that the control group was likely to prove better. In IE group, every student made positive progress, although to different degrees…There was a feeling among staff that task-intrinsic motivation had increased within the IE group and that this was shown through a willingness to cope with difficult and challenging material and through conscious preparation for exams. There was also a recognition of increased confidence in	The instrumental enrichment programme 'can positively affect the lives of children with emotional difficulties'. However, it 'offers no magical cures or quick solutions,... demands a high level of commitment and input from staff, and much agonising on the part of students as they learn to think things through, often for the first time…The overall benefit of the IE programme can perhaps be summarised by saying that it lies in the realisation that the planning behaviour and self-regulation fostered by the programme lead to successful living' (p 128).

		most of the pupils in the group, especially in comparison with those in the control group.' (p126)	
IT11910	Hojnacki and Grover (1992)	Findings reported in text and tables 1 & 2 and figures 1-4 in Appendices 1a (mid-year teacher survey): 'What was quite clear was that the teachers reported their instructional practices to be changing'. Some teachers (90%) made adjustment to their teaching timelines. Some (80%) changed use of textbooks (with some teachers dropping them completely). 65% had altered their grading and/or assessment practices. 59% said that what pleased them most about TM programme were student related factors (e.g. enhanced learning, increased motivation and greater enjoyment in doing maths). Appendices 1b (end-of-year teacher evaluation): Teachers noted both direct and indirect benefits to students. Direct - affective benefits: student self-confidence and self-esteem; cognitive benefits - deeper student understanding of maths and greater number sense; indirect benefits – teachers' own professional development, improved instruction/pedagogy, teacher enthusiasm and confidence imparted to students, integration of maths throughout the curriculum Problems identified in 1b: Not enough time to plan work; concern with maintaining a balance between allowing students to develop deep understanding and covering the curriculum by the end of the year; concern that TM would disadvantage students' performance on standardised tests Computational scores between TM and non-TM classes are relatively small (grades 2 and 3); non-TM score is higher than TM score (grades 4 and 5); TM score is slightly higher than non-TM score (grade 1); TM score is 25% higher than non-TM score (Figure 1); concepts and applications subtests = mean scores for TM classes in grades 1, 3, 4 and 5 are higher than for non-TM; grade 2 TM class mean is equal to non-TM score (Figure 2). Student attitude towards maths survey: overall the mean percentage of positive responses (responses indicating positive attitude to maths) across all scales and grades was 79% = mean percentage for individual grades and	The available data on teachers' experiences with Thinking Mathematics support the view that teachers are changed in empowering ways. The project's teacher self-study data show that these teachers perceived empowering changes in their students as well. Performance on the Wood-Cobb tests indicated notable improvements in student problem-solving ability: positive student attitudes towards mathematics were revealed. Students' lower motivation scores, however, suggest they are not yet ready to match their efforts with their generally positive regard for mathematics. TM students performed as well or better than their non-TM peers on both computational and concepts and applications subsections of the standardised test. There are multiple indications that student learning and attitudes were enhanced by their participation in the programme. The collaborative model developed by the Thinking Mathematics project seems to be paying off for both teachers and students.

		individual scales all 70%+; item with highest percentage of positive responses (97%) was from the confidence scale question 'I am sure that I can learn math'. More than 90% of the students gave a 'no' response (indicating a positive attitude toward maths) to two other items from the general mathematics scale. 'I dislike everything about math' and 'Only a few people can learn math': lowest figures of positive responses were found in the motivation scale (with scores of 53% and 57% found on this scale), and one score of 56% positive response on the confidence scale. Student problem performance: post-test accuracy rates were much higher that pre-test rates for all three tests on paired one-tailed t-tests (Gr.1: t(37) = +11.37; Gr.2: t(44) = +19.65; Gr.3: t(74) = 11.12, all three at p< .0001) with gains ranging from 25% on the Grade 3 tests to 45% on the Grade 1 and Grade 2 tests. Errors: The proportions of idiosyncratic errors were similar to common errors at both pre- (27% for both common and idiosyncratic) and post- test (14% for both common and idiosyncratic). Paired two-tailed t-tests conducted on the data relating to the range of answers indicate that the range of answer narrowed significantly from the pre-test to the post-test on the Grade 1 and Grade 2 tests; there were no significant differences on the Grade 3 test. Types of common errors: categorised into computational (miscalculation by small margin), conceptual/procedural, operational, problem misunderstanding and indeterminate strategy. The single consistent change across all three tests (all three grades) is that the percentage of indeterminate strategy errors decreased the most from pre- to post-test.	
IT13839	Kramarski and Mevarech (1997)	The SOLVE trained classes outperformed the control classes on the graph construction test at the 10% level of significance, using a two-tailed test (Table 1, and accompanying text pp 435-6). Members of the metacognitively trained classes outperformed controls on the following measures: social-cognitive interaction (p<0.01), self-reflection (p<0.01), number of students recalling problem-solving strategies (p<0.01), LOGO recall (p<0.05), information processing (p<0.05), error detection in a histogram (p<0.05), the daily life usefulness of LOGO (p<0.05) and the daily life usefulness of mathematics (p<0.05). No significant differences were found in the use	After being taught to use the metacognitive SOLVE strategy, most pupils saw it as being generally relevant for solving problems. In relation to the learning of course content, they tended to construct graphs better, showed better understanding of graphs and were more likely to detect an error in a graph. They were also better able to reflect on their learning and showed more positive social-cognitive interaction than their counterparts not trained to use SOLVE. In short, metacognitive training ‘exerts positive effects on students' achievement outcomes and cognitive-metacognitive behaviours’.

		of extrapolation, in LOGO error detection or in the number of students recalling statistics and research methodology or geometrical knowledge and the daily life usefulness of literature (Table 2 and accompanying text, pp432-440).	
IT13503	Maltby (1995)	The group who received strategy training had (1) fewer repeats of ineffective folds, (2) more decreases in folds and fewer increases in folds, and (3) more students who achieved success. Students in the group receiving training were observed to be slower in completing folds, appeared more thoughtful and manipulated the paper more than those in the other group. Students in the second group folded more quickly with less thought; they appeared to have no plan of action; and did not efficiently re-assess their incorrect responses. The students in the group receiving strategy training felt in interviews that they had been more successful and agreed that that it was better to go slowly and think carefully about the options. Group 2 students agreed that they wanted help and some students said that they would have liked help.	The authors conclude that the higher frequency of successes by Group 1 (who received training) may be evidence that training in thinking strategies aids problem-solving. Encouraging students to be less impulsive, to consider alternatives and revise them if they are ineffective, can enable them to think more efficiently and perhaps achieve more success. TASC appears positively to promote the development of cognitive skills and strategies. Self-monitoring and reflection enables students to improve their solutions and to become more effective thinkers.
IT13885	Mastropieri *et al.* (1997)	'All students increased significantly in the number of word problems solved correctly from the pre-test to the on-line computer post-test. Interestingly, all students descriptively increased their performance on the paper-and-pencil post-test, but the difference was not statistically significant' (p161). The students 'increased in familiarity with the CAI program and became more independent, as assessed by the number of times they requested assistance form the trainer' (p163). Three of the four students recalled the majority of the strategy steps after training. All students enjoyed using the computer and the CAI program. Two students responded positively and two negatively to a question about whether the animations taught them anything (p165).	'Findings from this project suggest that students with mild mental retardation can successfully learn problem-solving skills from a CAI tutorial program designed to include effective instructional components and animation when teacher assistance is provided initially' (p163). 'CAI tutorial programs may be beneficial supplements to instructional programs' (p164). 'Transfer of computer-assisted problem-solving to paper-and pencil problem-solving was less consistent' (p157).
IT13613	Mercer *et al.* (1999)	'We therefore have evidence that the use of exploratory talk helps the joint solution of problems' (p105). 'We have also shown that the intervention programme increased the amount of exploratory talk used by the target focal groups' (p107). 'It can be seen that there is a relatively greater improvement in the scores for the target classes, which is in accord with our hypothesis' though 'not statistically significant' (p107). 'The gains made by the individual target class children were significantly greater than those made by children in control classes' (p107). (a) Using the	'…being taught to use exploratory talk helps develop children's individual reasoning skills' (p108). 'Our results support the view that the induction of children into cultural practices influences their use of language as a cognitive tool' (p108). 'Teacher-directed activity can have a significant influence on the development of children's reasoning' (p109). 'A sociocultural perspective provides the best available theoretical basis for the critical analysis and improvement of educational practice' (p109). They therefore make

		kind of language we call 'exploratory talk' helps children to work more effectively together on problem-solving tasks; (b) using a specially designed programme of teacher-led and group-based activities, teachers can increase the amount of exploratory talk used by children working together in the classroom; and (c) children who have been taught to use more exploratory talk make greater gains in their individual scores on the Raven's test of reasoning than do children who have had no such teaching (p108). The intervention programme increased the amount of 'exploratory talk' used by the target focal groups when solving Raven's Matrices reasoning problems (Tables IIb and IIIb, pp106-7). (Please note that this finding is as reported by the authors, but is based on an error in interpreting the supporting statistical analyses.) Adherence to the ground rules helped groups solve the reasoning test problems, as it was found that when arriving at correct group solutions, there was a high level of 'exploratory talk' in one group of three children Table I, p105). Target children's individual (but not group) performance on Raven's Matrices improved (Tables IV and V, p107).	strong and broad claims: (a) the study supports the view that a sociocultural perspective provides the best available theoretical basis for the critical analysis and improvement of educational practice; (b) it supports the part played by cultural practices and the interactive role of adults and peers in the development of children's reasoning; (c) being taught to use exploratory talk helps develop children's individual reasoning skills; (d) teachers need to explain, justify and scaffold the 'ground rules' of educational language practices; and (e) teacher-directed activity can have a significant influence on the development of children's reasoning.
IT13887	Ritchie and Edwards (1996)	The CoRT lessons did not significantly affect cognitive ability, teacher-rated school achievement, self-reported use of CoRT thinking approaches, self-concept as a thinker or internal locus of control. Significant and substantial overall gains were made on the Torrance Test measures (fluency, flexibility and originality). Trend analysis showed a consistent upwards movement in fluency and flexibility, but a levelling off after an initial gain in originality. Implementation integrity was satisfactory, although there were some problems in achieving effective group work. The teachers' familiarity and acceptance of the CoRT approach was found wanting in two respects: they felt uneasy about using the CoRT skill acronyms and they did not always demonstrate enthusiasm and confidence. However, students experienced a high level of success in the lessons.	Pupils may need to be taught how to work in groups. Creative thinking can be taught with CoRT and the decision to capitalise on a perceived area of relative strength in Aboriginal children was supported. Success in CoRT lessons is not enough to produce more generalised gains. The study suggests that the CoRT approach is not immune to teacher effects and successful implementation may require commitment to the CoRT materials and goals.
IT13507	Strang and Shayer (1993)	The authors report a significant difference between the post-test scores of the experimental and control group. The results of the regression equation suggest a bimodal distribution with some participants benefiting more than others. This means that the t-test score is misleading as	The post-test result showed that some students in the experimental group had obtained a better understanding of chemical change than the students in the control group. It does appear feasible that Feuerstein's theory might have applications within the

		most of the effect is located in four high scoring participants. The result of the experiment is not uniform for the experimental group; the reason for the effect appears to lie neither in the gender of the student nor in their initial abilities.	science curriculum both to aid concept formation and to improve general cognitive skills for problem-solving. For the teacher, the theory provides insight into the cognitive difficulties associated with particular concepts so that these can be compensated for in the selection of materials and the teaching approach adopted (p336).
IT11897	Verschaffel and De Corte (1997)	'First the results on the pre-test confirm the finding from previous studies (see Greer, 1993; Reusser & Stebler, 1997; Verschaffel *et al.,* 1994) that the current culture and practice of mathematics education elicits in pupils a strong tendency to exclude real-world knowledge and sense making when representing and solving arithmetic word problems. (Percentages of realistic reactions for E, C1 and C2 were 7%, 20% and 18% respectively.) Second, the experimental program had a positive effect on the disposition of the pupils in the E class toward realistic modelling and interpretation of arithmetic word problems. In addition, the relatively small and insignificant increase in RRs of the C1 class indicates that it is not sufficient merely to tell and to illustrate that routine solutions for word problems are not always appropriate to transform pupils into more critical and realistic problem-solvers. (Tukey a posteriori tests showed a significant increase in realistic reactions in the E group from 7% to 51%.) Third, the positive effect of the experimental program was not restricted to problematic items that resembled those used during the intervention, but transferred to P items about contexts differing considerably from those encountered during training. (There was a significant increase in near transfer items from 6% to 41% in the post-test for the E group.) Fourth, the results of the E group on the retention test showed that the positive effect of the experimental programme did not disappear after the training had stopped, and they even provided some evidence of far-transfer effects. (Percentage of realistic reactions on retention test was 40% - near equivalence to immediate post-test (41%))' (p 597, details in brackets from pp 592-3).	'These results warrant a positive conclusion about the feasibility of developing in upper elementary schools a disposition toward realistic mathematical modelling or wor(l)d problems. However, some caution is in order. We note, first, that the overall percentages of RRs of the E class on the post-test and the retention test were still relatively low (i.e., 51% and 40% respectively). A close look at the individual test scores revealed large differences in the increase in RRs from pre-test to post-test to retention test between pupils with different mathematical abilities; the high ability pupils tended to benefit more from the intervention than their average and low ability peers. Similarly, the increase in the number of RRs was considerably greater for items about the interpretation of the outcome of division with a remainder. The analysis of the videotapes of the E class not only revealed some difficulties with respect to the design and implementation of each of the three pillars of the instructional program, which allowed us to explain some of the moderate results mentioned above, but also pointed to possible improvements of the experimental program' (p 597).
IT13466	Ward and Traweek (1993)	'Students who followed the think-aloud procedure requiring them to talk about their reading strategies as they performed a cloze task improved their reading comprehension scores significantly more than those who	'The use of a think-aloud procedure for the assessment of reading may alter outcome in favour of better performance scores. This information may be relevant in dynamic assessment of whether students

		performed the cloze task without using the think-aloud procedure (word identification: mean scores of 54.9 against 43.0; mean scores passage comprehension: 79.4 against 49.3)' (p 475). 'Five of the ten questions used during reading that were believed to be most salient in generating general metacognitive strategies were analysed and no differences were found between the groups' (p 476).	only need a prompt for activating metacognitive awareness and knowledge of strategic processing while they read.' (p 476)
IT12365	Wegerif (1996)	Comparing the control and experimental groups completing the computer tasks the authors report that the experimental group: (1) Asked each other more task-focussed questions. (2) Gave reasons for statements and challenges. (3) Considered more than one possible position. (4) Drew opinions from all in the group. (5) Reached agreement before acting. Most control children did the following: (1) Unilateral action of the child using the mouse. (2) Accepted the choice of the most dominant child without reasoning together. (3) Drifted together to one or other choice without debating any alternatives. Quantitative analysis showed that target children used more questions and had larger scores of total words, and used slightly more 'because' phrases and 'if' phrases. Results from the pre- and post- tests using Ravens matrices showed: (1) All the group scores in both target and control classes increased over the period of the intervention programme. (2) The target class group score increased by 32% while the control class group score increased by 15%. (3) The differences in the pre- and post-test group scores for the intervention group were significantly different from the pre- and post-test scores of the control group. (4) The mean individual test score in the control class remained approximately the same, while that of the intervention group improved by 10%. (5) A significant difference was shown between the pre- and post-test intervention scores of the control group and intervention group. Discourse analysis showed that problems that had not been solved in the pre-test were solved in the post-test leading to the marked increase in the group scores. It is stated that these were solved as a result of group interaction strategies associated with exploratory talk and coached in the intervention programme. Frequency count data from three focal	The research described in this paper indicates the following: (1) Coaching exploratory talk leads to improved group problem-solving. (2) Coaching exploratory talk appears to improve the scores of some individuals on reasoning tests. (3) Computers can be used effectively to support exploratory talk amongst groups of children and to direct this towards curriculum ends (p 59).

		groups in the intervention group showed an increase in test scores, questions, 'because phrases', 'if phrases' and total words.	
IT13838	Zohar (1996)	'The findings from this study show that when students in eighth and ninth grade first encounter problems described here (science problems), most of them do not use accurate scientific reasoning strategies. Interaction with the learning environment increased the rate of students' valid inferences from 11% to 77%' (Table II, p205). The average percentage of valid inferences made by nine students in the early interviews of problem 1, 2 and 3 were 10.2, 87.5 and 85 (Table III, p211). 'A decrease in the total number of inferences during the later interviews indicates that students' investigations become more focused and systematic as they gain experience' (Table II, p210). 'Students were able to transfer their newly acquired reasoning strategies to a new problem taken from a new biological topic. They were also able to retain their newly acquired strategies across time, and to transfer from them to yet another biological topic 5 months after instruction took place' (p205).	'These findings show that the new learning environment was indeed effective in inducing change in student's reasoning strategies' (p210). 'The study shows that methods which were previously found to be effective in inducing change in laboratory conditions (one interviewer per subject and long-term treatment lasting numerous sessions) are also effective in classroom conditions (one teacher per 25-30 students over shorter periods)' (p216). 'Creating a learning environment that can induce conceptual change is not an easy task. Teaching scientific reasoning strategies is not an easier task than teaching scientific concepts.'

APPENDIX 4.4: Synthesis tables

Synthesis summary

EPPI ID Review ID #	Author(s)	WoE A (Study quality)	Review WoE D (Overall)	Quantitative impact curriculum	Quantitative impact non-curric	Retention	Transfer	Qualitative impact attainment
14198, #837	Adey and Shayer, 1990	H	H	=	+ (u[1])	=	+	
11895, #44	Chang and Barufaldi, 1999	H	H	+	+			+, =
13508, #935	Csapó, 1992	H	H		+ (u)			
13509, #909	De Corte *et al*., 2001	H	H	=	+		+	+
13480, #928	De Koning and Hamers, 1999	H	M	+	+		+	
13887, #802	Ritchie and Edwards, 1996	H	M	=	=. +			=, +
13612, #922	Blank, 2000	M	M	=		+		+
13471, #147	Cardelle-Elawar, 1992	M	M	+				+
11899, #314	Georghiades, 2000	M	M			+	+	
13879, #774	Haywood *et al*., 1988	M	M	+, -	+			+
13839, #431	Kramarski and Mevarech, 1997	M	M	+	+			+
13507, #575	Strang and Shayer, 1993	M	M	+ (u)				
11897, #87	Verschaffel *et al.*, 1997	M	M	=		+	+	
13466, #212	Ward and Traweek, 1993	M	M	+				
12365, #480	Wegerif, 1996	M	M		+			+
13838, #114	Zohar, 1996	M	M	+		+	+	
11910, #160	Hojnacki and Grover, 1992	M	L	=				+
13562, #235	Head and O'Neill, 1999	L	L	+	+		+	+
13531, #883	Bowdler *et al.,* 1992	L	L	+	+			+
13530, #925	Fields, 1995	L	L	=	+			+
13505, #485	Maltby, 1995	L	L		+			+
13885, #97	Mastropieri *et al.*, 1997	L	L	=	+			+
13613, #386	Mercer *et al.*, 1999	L	L		+			+

Key: = indicates *no significant difference* between intervention and comparison/controls; + indicates an *attainment gain* by the thinking skills intervention pupils; - *relative decline in attainment* by the thinking skills pupils; (u) indicates *uneven* impact.

[1] Impact was not even across all groups of pupils