REVIEW

December 2004

A systematic review of the impact of the Daily Mathematics Lesson in enhancing pupil confidence and competence in early mathematics

Review conducted by the Mathematics Education Review Group

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

AUTHORS

The main authors of this report are Chris Kyriacou and Maria Goulding (University of York, Department of Educational Studies). They drew extensively on the work of the Review Group as a whole.

REVIEW TEAM MEMBERSHIP

This Review Group was co-ordinated by Chris Kyriacou and Maria Goulding.

Patti Barber	Department of Primary Education, Institute of Education, University of London
Peter Bland	Huntington Comprehensive School, York
Dr Robert Coe	Curriculum Evaluation and Management Centre, University of Durham
Dr Ann Dowker	Department of Psychology, University of Oxford
Ann Gannon	Department of Educational Studies, University of York
Maria Goulding	Department of Educational Studies, University of York
Gillian Hatch	Institute of Education, Manchester Metropolitan University
Professor Charles Hulme	Department of Psychology, University of York
Qaimah Ismail	Department of Educational Studies, University of York
Professor Barbara Jaworski	Department of Mathematics, Agder University College, Norway
Dr Chris Kyriacou	Department of Educational Studies, University of York
Paul Lazenby	Department of Physics, University of York
Ann MacNamara	National Numeracy Strategy, London
Alison Robinson	Department of Educational Studies, University of York
Dr Tim Rowland	Department of Education, University of Cambridge
Dr Sue Sanders	Department of Education, University of Swansea
Sally Sutton	St Lawrence Church of England Primary School, York
Professor Peter Tymms	Curriculum Evaluation and Management Centre, University of Durham

ADVISORY GROUP MEMBERSHIP

The membership of the Advisory Group is the same as the Review Group. However, other individuals (teachers, researchers, policy-makers) with an interest in the review question were invited to comment of the work of the Review Group at appropriate times. This was largely done through email and through informal conversations at conferences.

CONFLICTS OF INTEREST

There were no conflicts of interest for any members of the Review Group.

ACKNOWLEDGEMENTS

The Mathematics Education Review Group and this review are part of the initiative on evidence-informed policy and practice at the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, funded by the Department for Education and Skills (DfES). The Review Group acknowledges the financial support from the DfES via the EPPI-Centre.

We wish to thank staff at the EPPI-Centre for their advice, support and guidance throughout the various stages of carrying out this review, with particular thanks to Diana Elbourne and Mark Newman.

Our thanks also go to members of the Review Group and the institutions to which they belong for their support. We are particularly grateful to those members of the Review Group who attended the various meetings in London and York and were involved in the data extraction stage of the review.

We would also like to thank numerous individuals who kindly sent us copies of their published and unpublished papers, together with their thoughts and comments on our work at various stages in the review process, and also kindly suggested references that we should consider for inclusion.

LIST OF ABBREVIATIONS

BEI	British Education Index
DfEE	Department for Education and Employment
DfES	Department for Education and Skills
EPPI-Centre	Evidence for Policy and Practice Information and Co-ordinating Centre
ERIC	Educational Resources and Information Centre
KS 1	Key Stage 1 (aged 5 to 7 years)
KS 2	Key Stage 2 (aged 7 to 11 years)
LEA	Local education authority
NLS	National Literacy Strategy
NNP	National Numeracy Project
NNS	National Numeracy Strategy
Ofsted	Office for Standards in Education
PGCE	Post Graduate Certificate in Education
REEL	Research Evidence in Education Library
SAT	Standard Assessment Test
SSCI	Social Science Citation Index
TIMSS	Third International Mathematics and Science Study

This report should be cited as: Kyriacou C, Goulding M (2004) A systematic review of the impact of the Daily Mathematics Lesson in enhancing pupil confidence and competence in early mathematics. In *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

TABLE OF CONTENTS

SUMMARY

Background

In September 1999, the Department for Education and Skills (DfES) introduced a National Numeracy Strategy (NNS) for primary schools in England. The NNS included the following key elements:

- a daily mathematics lesson in primary schools lasting between 45 and 60 minutes
- a three part structure to these lessons comprising (i) an oral/mental starter (ii) the main teaching and pupil activities and (iii) a plenary
- an emphasis on the use of interactive whole class teaching

Aims

The aim of this review is to consider the research evidence that bears upon the ways in which the approach adopted by teachers during the Daily Mathematics Lesson in delivering the NNS has impacted on pupils' confidence and competence in early mathematics.

Review question

Has the Daily Mathematics Lesson, in the context of the National Numeracy Strategy for primary schools in England, helped pupils to develop confidence and competence in early mathematics?

Methods

Identifying relevant studies involved carrying out an electronic search using keywords with bibliographic databases, handsearching through key journals and conference proceedings, citations and publications recommended by contacts. This resulted in 18 studies being identified for the in-depth analysis.

Results

- The key features of the Daily Mathematics Lesson have been well received by teachers and widely implemented.
- There is some evidence that this has enhanced pupil confidence and competence in early mathematics.

- A closer examination of the situation as evidenced by the studies included in this systematic review has highlighted a number of problematic issues.

- The intention that whole class teaching needs to be 'interactive' and promote higher quality dialogue, discussion and strategic thinking, has not been realised. Indeed, there is some evidence to indicate that the increased use of 'traditional' whole class teaching with 'pace', is in fact undermining the development of a more reflective and strategic approach to thinking about mathematics, and may be creating problems for lower attaining pupils.

- There is evidence that the stricter time management involved may pose particular problems for lower attaining pupils.

- The overall enhanced gains in pupil competence may be a reflection of a closer match between what is being taught and what is being tested, rather than greater pupil gains in their understanding of mathematics.

Conclusions

The main strengths of this review have been that the review process has followed a publicly visible procedure, and has benefited from the collaboration involved between the Review Group, the EPPI-Centre, and many other individuals who offered comment, help and advice. The close scrutiny of the procedures involved means that each stage of the review process involved discussion and justification.

The main limitations of the review are that the constraints involved in terms of time, cost and access to relevant papers, inevitably means that decisions about the focus of the review question and the conduct of the review process had to be taken in the context of keeping the review manageable.

Our main conclusions are as follows:

- Policy-makers need to consider the extent to which the apparent success of the NNS may in part be a reflection of greater teaching for the test, whether the approach of interactive whole class teaching with pace may be inculcating bad learning habits, and whether the needs of low attaining pupils are being well served by the NNS. There is also a need to consider how the national assessment of pupil progress in mathematics can occur without constraining time and pedagogy in ways that are undermining the development of pupils' mathematical understanding.

- There is a major need for in-service training for primary teachers to highlight the purpose and nature of 'interactive' teaching in fostering higher quality dialogue, discussion and strategic thinking in order to ensure that teachers better understand the notion of 'interactive' in interactive whole class teaching, and to ensure that teachers adopt the type of classroom practice that can effectively aid the development of pupils' understanding of the mathematics of the topics they are covering. In particular, there is a need for in-service training to strengthen teachers' subject matter knowledge of mathematics, so that in the classroom context they can take better advantages of opportunities to enhance pupils' understanding of the mathematics they are engaged in.

- There is a need for researchers to make greater use of measures and indicators of pupil confidence in order to draw firmer conclusions about how the features of the Daily Mathematics Lessons impact on pupils' feelings of self-efficacy concerning the learning of mathematics.

1. BACKGROUND

1.1 Aims and rationale for current review

The aim of the review is to consider the research evidence that bears upon the ways in which the approach adopted by teachers during the Daily Mathematics Lesson in delivering the National Numeracy Strategy (NNS) has impacted on pupils' confidence and competence in early mathematics.

The biggest challenge facing mathematics education is to ensure that each new generation of pupils makes the best possible start in the development of their knowledge and understanding of mathematics. Just as important as early progress in pupils' mathematics attainment is the development of pupils' self-confidence in themselves as learners of mathematics. This is because attainment is always normative – some pupils must inevitably do less well than others – and it is this sense of doing less well than one's peers that can all too easily lead to disaffection and anxiety in mathematics, and create a vicious circle of low self-confidence and underachievement. The key importance of developing pupils' self-confidence in mathematics education (sometimes referred to as 'self-efficacy beliefs') has been recognised by a host of writers (e.g. Coben, 2003; Dorman *et al.,* 2003; Jones and Smart, 1995) and the importance of developing and sustaining this during the early years of schooling has also been regarded as paramount (e.g. Gifford, 2003; Harries and Spooner, 2000).

As such, any approach that can help develop and sustain a sense amongst pupils that, regardless of their relative attainment, mathematics is understandable and doable, can play a key role in preventing such a vicious circle developing, and in contrast can help generate a virtuous circle of positive attitudes and good progress.

1.2 Definitional and conceptual issues

The Department for Education and Employment (DfEE) (1999) defined numeracy within the NNS as follows:

> Numeracy at Key Stages 1 and 2 is a proficiency that involves a **confidence and competence** with numbers and measures. It requires an understanding of the number system, a repertoire of computational skills and an inclination and ability to solve number problems in a variety of contexts. Numeracy also demands practical understanding of the ways in which information is gathered by counting and measuring, and is presented in graphs, diagrams, charts and tables. (p 11) [The emphasis on the words *confidence and competence* is ours.]

After much discussion the Review Group decided to incorporate the term 'mathematics' rather than 'numeracy' in the review question, as it was the contribution that the NNS makes to the development of pupils' mathematical education that was felt to be the main concern here. There has been much debate about the difference in meaning between mathematics and numeracy (Askew and

Brown, 2001; Maclellan, 2001). This debate need not concern us here except to say that numeracy can be regarded as being a subset of mathematics broadly along the lines of that indicated by the definition of numeracy given by the DfEE above.

Confidence refers to pupils' belief in their abilities. It is evidenced by positive attitudes pupils hold towards learning mathematics; behavioural indicators that they find learning mathematics interesting and enjoyable; and, most importantly, that when faced with a mathematics task to undertake, they approach that task with a feeling and expectation that they will be able to succeed.

Competence refers to what pupils know, understand and can do, as evidenced by their performance in a range of appropriate mathematics tasks.

1.3 Policy and practice background

The NNS was introduced into primary schools in England in September 1999. Other approaches to the teaching of mathematics have been adopted in Wales, Scotland and Northern Ireland, which may in time allow for some very interesting comparative research to be undertaken on pupils' progress in mathematics across these four nations of the UK (Jones, 2002; Maclellan *et al.,* 2003). The extent to which the design of the NNS was grounded in research evidence on the teaching and learning of mathematics has been widely debated (e.g. Brown *et al.,* 1998; Thompson, 2000).

The NNS makes provision for *a daily mathematics lesson* of between 45 and 60 minutes (often referred to as 'the numeracy hour') in which teachers use whole class teaching methods for a high proportion of the time, and in which oral and mental mathematics are given a high priority (Hughes, 1999). Key features of the Daily Mathematics Lesson are described by the DfEE (1999) as follows:

- The teaching programme is based on identified learning objectives, and is planned thoroughly, to ensure high expectations, consistent approaches and good progression throughout the school.
- The foundations of mental calculation and recall of number facts are established thoroughly before standard written methods are introduced.
- Assessments are used to identify pupils' strengths and difficulties, to set group and individual targets for them to achieve and to plan the next stage of work.
- Assessments include informal observations and oral questioning, regular mental tests, and half-termly planned activities designed to judge progress.
- Recording systems give teachers the information that they need to plan and report successfully, but are not too time-consuming to maintain.

In addition, teachers undertake the following:

- They structure their mathematics lessons and maintain a good pace.

- They provide daily oral and mental work to develop and secure pupils' calculation strategies and rapid recall skills.
- They devote a high proportion of lesson time to direct teaching of whole classes and groups, making judicious use of textbooks, worksheets and Information and Communication Technology (ICT) resources to support teaching, not to replace it.
- They demonstrate, explain and illustrate mathematical ideas, making links between different topics in mathematics and between mathematics and other subjects.
- They use and give pupils access to number lines and other resources, including ICT, to model mathematical ideas and methods.
- They use and expect pupils to use correct mathematical vocabulary and notation.
- They question pupils effectively, including as many of them as possible, giving them time to think before answering, targeting individuals to take account of their attainment and needs, asking them to demonstrate and explain their methods and reasoning, and exploring reasons for any wrong answers.
- They involve pupils and maintain their interest through appropriately demanding work, including some non-routine problems that require them to think for themselves.
- They ensure that differentiation is manageable and centred around work common to all the pupils in a class, with targeted, positive support to help those who have difficulties with mathematics to keep up with their peers.

The structure of the Daily Mathematics Lesson is also prescribed as having three parts (DfEE, 1998):

(i) *Clear start to lesson* (about 5 to 10 minutes)

- Whole class mental and oral work to rehearse and sharpen skills

(ii) *Main teaching and pupil activities* (about 30 to 40 minutes)

- Whole class/groups/pairs/individuals
- Clear objectives shared with pupils
- Interactive/direct teaching input
- Pupils clear about what to do
- Work on the same theme for all the class
- If group work, usually differentiated at no more than three levels of difficulty, with focused teaching of one or two groups for part of the time
- Continued interaction and intervention

- Misconceptions identified

(iii) *Plenary* (about 5 to 10 minutes)

- Whole class feedback from children to identify progress and sort out misconceptions
- Summary of key ideas
- Links made to other work
- Next steps discussed
- Work set to do at home

The NNS also requires that the Daily Mathematics Lesson be taught in special schools and in special or inclusion units in mainstream schools. Furthermore, it makes provision for the early identification of pupils who are falling behind their peers in their progress to receive additional support, such as through the provision of booster classes, summer schools and other types of numeracy recovery programmes. The review here will focus exclusively on the Daily Mathematics Lesson being taught to mainstream pupils in mainstream schools. This is because teaching pupils in these other contexts is very different and will raise different issues of pedagogical interpretation and generalisability. However, the impact of the NNS on low attaining pupils who require such additional support is an extremely important area that could usefully form the focus for a separate systematic review.

1.4 Research background

The NNS was adopted throughout England following the promising early results of the National Numeracy Project (NNP) which had been running in certain inner city areas since 1996. The NNP had been created to address concerns about low standards of numeracy, an overemphasis in the teaching schemes on standard written rather than mental methods, and a lack of direct teaching by the classroom teacher (Straker, 1999). Many of these concerns had arisen because of the relatively poor mathematical performance of English primary pupils in international comparisons such as the Third International Mathematics and Science Study (TIMSS) (Keys *et al.,* 1996), stimulating interest in the teaching methods of higher achieving countries. Primary pupils in mainland Europe, and in particular Swiss pupils, were taught as a whole class and there was virtually no grouping by ability within the class (Prais and Luxton, 1998). The high levels of achievement of primary Hungarian pupils were attributed to high expectations, consistent teaching, ongoing assessment in class, better continuity and progression in the texts used, and differentiated teaching with a high degree of support and challenge enabling pupils of different abilities to move forward together (Graham *et al.,* 1999). In the recommendations for teaching methods in the NNS, English primary teachers, whose pupils during the 1980s had largely worked individually from published schemes, were encouraged to incorporate the above features in daily mathematics lessons (Reynolds, 1998) using ‘whole class interactive teaching’ (Graham *et al.,* 1999).

A number of authors have pointed out how the emphasis on whole class teaching within the NNS requires teachers to develop a form of interactive teaching that provides a supportive environment for pupils in which their self-confidence in mathematics can develop (Moyles *et al.,* 2003; Wright *et al.,* 2002).

Despite the wealth of studies on the NNS that have emerged in the short period since it was introduced, these studies have largely focused on cognitive outcomes and levels of pupil attainment in mathematics (e.g. Aubrey and Godfrey, 2003; Brown and Millett, 2003). There is a desperate need to identify those studies which provide evidence of the extent to which the NNS has impacted on affective outcomes, which may be particularly evident in studies which draw on lesson observations and interviews with pupils and teachers (e.g. Brown *et al.,* 2003a; Butlin, 2003). Ofsted (2002) in its report on the first three years of the NNS notes that:

> Pupils' confidence, enjoyment and involvement have improved since the strategy began. They respond positively to the routines and clear structure of the daily mathematics lesson and they are motivated by the direct teaching which it requires. Many pupils understand their strengths and weaknesses in mathematics better, as well as the progress they are making. (p 2)

However, the evidence base for Ofsted's statements about affective outcomes is not made explicit in their report. In Ofsted's (2003) report, there is no mention at all regarding pupil confidence. As such, the Review Group intends to focus on the impact of the NNS on both pupils' *confidence* and *competence* in mathematics.

The NNS itself is in fact a wide ranging initiative that goes well beyond classroom practice and aims to provide a broad framework for the development and support for the effective learning of mathematics (DfEE, 1999; Earl *et al.,* 2003). In addition to guidelines on the content and approach to be adopted in the Daily Mathematics Lesson, the NNS also covers wider issues, including how the teaching of mathematics in schools can be effectively managed; early intervention programmes for pupils falling behind; the continuing professional development needs of teachers; the role of national testing; and the setting of attainment targets to raise standards. Each of these separate facets of the NNS may have a distinctive influence on the success or otherwise of the NNS in promoting mathematics.

However, it is the Daily Mathematics Lesson that lies at the heart of the NNS. It is the mathematical content of these lessons and the approach to teaching adopted there that have the most direct impact on pupils' learning. The focus of the review undertaken here will thus deal specifically with the teaching and learning that occurs in these lessons within the context of the NNS, although aspects of the broader framework will need to be considered to the extent that they clearly bear upon the effectiveness of these lessons. In focusing on the Daily Mathematics Lesson, we also need to consider how well what happens in these lessons matches what we know about the effective teaching of mathematics (Anghileri, 2001; Askew and Brown, 2001, 2003; Macrae, 2003).

1.5 Authors, funders and other users of the review

The Review Group comprises individuals from the key groups involved in mathematics education: teacher educators, academic researchers, primary and secondary school teachers, school governors, parents and policy-makers.

All members of the group were involved in all stages of the review process:

- identifying the review question
- outlining the scope and method for the review
- identifying studies to establish the main review database
- identifying studies based on paper titles and abstract (first-stage inclusions)
- identifying studies based on full papers (second-stage inclusions)
- mapping the second-stage inclusions
- extracting data from and analysing the papers selected for the in-depth analysis
- writing the report

However, a core group was established to undertake the bulk of the work involved in the in-depth stage and writing the report.

The main audience for this review comprises student teachers, teachers, teacher educators, school governors, researchers and policy makers, although parents of school-aged children and other members of general public will also have an interest in this review question.

This review is timely in the light of the Smith Report (2004), which has recognised both the importance of the NNS and the importance of systematic reviews in mathematics education to provide an evidence base to inform policy and practice.

It is intended to disseminate the findings of the review through internet access to the review report, publication in an academic journal and conference papers. A conference paper on this review was presented at the British Society for Research into Learning Mathematics (BSRLM) day conference held at Leeds (12 June 2004) and at the International Group for Research on the Psychology of Mathematics Education (PME) annual conference held at Bergen, Norway (14–18 July 2004).

1.6 Review questions

Our review question is as follows:

Has the Daily Mathematics Lesson, in the context of the National Numeracy Strategy for primary schools in England, helped pupils to develop confidence and competence in early mathematics?

The objectives of the review are as follows:

(i) to identify relevant studies carried out in English primary (including infant and/or junior) schools in the period from the introduction of the NNS in September 1999 until the end of 2003

(ii) to undertake a descriptive mapping of the relevant studies

(iii) to undertake an in-depth analysis of the relevant studies

(iv) to draw conclusions from these studies on the extent to which the Daily Mathematics Lesson has helped pupils to develop confidence and competence in early mathematics

In order to address this review question, a number of more specific questions need to be considered. These can be sub-divided in terms of (i) the approach to teaching adopted and (ii) pupil outcomes, as follows:

(i) The approach to teaching adopted

What teaching methods and learning activities are being used and are they being used with respect to each of the three phases of the lessons: the mental starter phase, the main teaching phase, and the plenary phase?

(ii) Pupil outcomes

Do pupils display confidence in early mathematics? Do pupils display competence in early mathematics?

It also important to note that implicit in the use of the word 'helped' in the review question is the notion of whether the development of pupil confidence and competence during KS 1 since the NNS was introduced has been enhanced compared with what would have been the case if the NNS had not been implemented.

It quickly became obvious in undertaking this systematic review that, in order to address the review question, we could not assume that the Daily Mathematics Lesson in the context of the NNS was being uniformly implemented in classrooms across England in the way intended by the NNS. As such, our review question does not simply or primarily have a straightforward focus on measures of effectiveness using data on narrowly conceived pupil learning outcomes. Rather, it is also concerned with the classroom practice that has taken place following the implementation of the NNS, and on what we can infer from such practice about pupils' learning of mathematics, the confidence they show in the classroom setting and their performance on classroom tasks as well as on national assessments. This consideration of both classroom practice and pupils' experience of learning mathematics will be reflected in our analysis of the included studies.

2. METHODS USED IN THE REVIEW

2.1 User involvement

2.1.1 Approach and rationale

User group involvement is reflected in the composition of the Review Group itself, which includes teacher educators, academic researchers, primary and secondary school teachers, school governors, parents and policy-makers.

2.1.2 Methods used

User perspectives on the review process and the provisional report were sought and perspectives written by users are included in the final report. Details of this review have been circulated to a number of professional organisations, teacher educators, numeracy coordinators, researchers and policy-makers.

2.2 Identifying and describing studies

2.2.1 Defining relevant studies: inclusion and exclusion criteria

For a paper to be included in the systematic map, it had to satisfy the following five criteria:

(i) It is an academic paper written in English published in an academic journal or presented at an academic conference during the period 1999 to 2003.

(ii) It reports a study presenting original data collected by the author(s).

(iii) The study deals with a mainstream class in England.

(iv) The study deals with KS 1 pupils.

(v) The study deals with the Daily Mathematics Lessons in terms of the teaching and learning activities which occur in the lessons and/or the effects of the lessons on pupils' confidence and/or competence in early mathematics.

These inclusion criteria were reformulated as six exclusion criteria (see Appendix 2.1) and placed in the hierarchical order for ease of exclusion and, importantly, to act as a system of gradual filtering, so that the papers that are excluded at each stage can be readily identified in the future as a useful list of references that could be drawn upon for other purposes by readers of the review report, or may indeed be of use in subsequent systematic reviews undertaken by this Review Group. For example, studies dealing with KS 2 pupils were only excluded by the final exclusion code.

2.2.2 Identification of potential studies: search strategy

Initially, the review intended to have a primary focus on papers published in journals. This decision to focus the review on journal papers (rather than giving equal weight in the review to other sources, such as conference papers, books and reports) was taken because publication in a journal offers a recognised degree of quality control, as such papers are normally (but not necessarily) peer reviewed 'blind' by at least two referees with expertise in the topic area, and submissions to a journal normally contain the author's most polished and carefully considered presentation of the empirical data and its interpretation, which can often also have benefited from revisions required by the referees prior to its acceptance for publication. On the other hand, other types of publication do not benefit from such a process of external evaluation, whether 'blind' or not. In addition, journal papers are unequivocally in the public domain and can be more easily accessed as a result, and the use made of a particular journal paper in a systematic review of the literature can therefore be more easily scrutinised and verified.

However, the arguments specifically concerning the publication bias which can occur if unpublished studies are not included in a systematic review have been particularly well rehearsed (e.g. Thomas and Harden, 2003; Torgerson, 2003). For this reason and because the NNS was a recent initiative, it was important not to exclude studies reported in the form of a conference paper. As such, the initial search for relevant publications also included conference papers which were relevant to the review question and which, in particular, had appeared as full length papers in edited conference proceedings. These papers were carefully considered and, where they met the inclusion criteria for the review, were included in the in-depth analysis.

Preliminary searches helped to establish the key sources for both electronic searching and hand-held searching which were likely to identify references relevant for this review, as well as other sources which were likely to be of limited value. For example, reports available on the 'teachernet' website (www.teachernet.gov.uk) arising from the Best Practice Research Scholarship Programme funded by the Department for Education and Skills (DfES) identified 95 papers for the search term 'numeracy', but an examination of these papers indicated that, although some of these were relevant to the review question, the lack of detail in the presentation of the research findings in these reports meant they were not suitable for consideration as potential papers for an in-depth analysis. In addition, once the preliminary searches of BEI and ERIC were completed, preliminary searches of Education-Line, PsycInfo and SSCI proved to be unnecessary, since the latter search was much less effective and identified few additional papers of potential relevance. It was thus evident that an electronic search of any databases in addition to BEI and ERIC would be very unlikely to yield any includable studies once a search of BEI and ERIC plus handsearches had been completed.

Preliminary searches also indicated that it was an easy matter to access the archives of major journals in order to look at the titles of every paper published in the period 1999–2003, and it is often possible to also consult the abstract and/or a full-copy of the paper online. Given the possibility that an electronic search of titles alone using even a very comprehensive set of keywords can still miss relevant papers, it was felt that a combination of handsearching and electronic searching of key journals in this way was important in adding to the list of

potentially relevant papers identified by the electronic search of BEI and ERIC. The importance of carrying out an extensive handsearch has been noted by Black (2004), and for this purpose 26 key journals were identified.

The focus on early mathematics reflects the concern stated earlier that the school experiences that pupils in the early years of primary schooling concerning the teaching of mathematics can have a crucial impact on their future view of themselves as learners of mathematics. This focus will also enable the review to focus on teaching which deals with the early understanding of mathematics (e.g. the early understanding of number facts and number counting processes). While the primary attention here is on pupils in KS 1, studies of pupils towards the end of KS 2 may well be making a reference to pupils based on their experience of the NNS in KS 1. As such, studies with an exclusive focus on KS 2 pupils were also examined to consider whether they needed to be included in the in-depth stage of this review.

The period 1999 to 2003 was chosen for this review to cover the period following the introduction of the NNS in September 1999 (Askew, 2002; Tanner and Jones, 2000). This enabled the research evidence obtained during this period to be more readily applicable to this specific pedagogical context of teaching mathematics that pertained at the time that this review was taking place. The NNS provided a framework within which a particular notion of numeracy was established together with a particular way of teaching it. Research studies of the teaching of mathematics prior to 1999 deal with a different pedagogical context. By focusing on studies which deal with the NNS published in journals (supplemented by appropriate conference papers) from 1999 to 2003, any implications for policy and practice in schools drawing upon such studies could be made with greater confidence than research drawing on studies conducted when a different pedagogical milieu in schools was in operation. Given the NNS was implemented in September 1999, it is clearly unlikely that journal papers published in 1999 itself will deal directly with pupils being taught in the period after September. However, papers published in 1999 may well contain some useful data concerning the preparations and expectations for the teaching of mathematics in the context of the NNS which foreshadowed the onset of the NNS and identify issues which are relevant in addressing the review question.

Relevant studies were drawn from papers published in journals or conference papers presented during the period 1999 to 2003. Five strategies were used.

(i) Electronic search of BEI and ERIC

The key search terms to be used will be: numeracy, primary mathematics and daily mathematics lesson. (Full details of this search strategy can be found in Appendix 2.2.)

Subject: National Numeracy Strategy

Population: Key Stage 1 and Key Stage 2 pupils in mainstream classes

Limits: English Language, 1999 to 2003

(ii) Electronic search and/or handsearch of nine key journals in Mathematics Education (1999–2003), looking at every title and where appropriate and available the abstract and/or the full-paper (880 papers) (Appendix 2.3)

Electronic searches and/or handsearching issues of 17 selected key UK journals in Educational Research (1999–2003), looking at every title and where appropriate and available the abstract and/or the full-paper (2,007 papers) (Appendix 2.3)

(iiii) Handsearch and/or electronic search of key recent conference proceedings, looking at every title and where appropriate and available the abstract and/or the full-paper (3,043 papers) (Appendix 2.3)

(iv) Searching the list of references at the end of papers identified as relevant

Searching the list of references in recent books or chapters in an edited book dealing with primary mathematics

(v) Contacting researchers working in this field for their recommendations

2.2.3 Screening studies: applying inclusion and exclusion criteria

We applied the inclusion/exclusion criteria at three points:

(i) to the title and abstract of papers from searching electronic databases;

(ii) to a full-paper copy of papers not previously excluded on the basis of the title and abstract

(iii) to additional papers identified by handsearching, citations and personal contacts

2.2.4 Characterising included studies

The included studies were mapped (characterised) using the EPPI-Centre's Educational Core Keywording Sheet (EPPI-Centre, 2002a), which comprises generic keywords for this process (Appendix 2.4). There were no review-specific keywords for the systematic map.

Keywording of papers included in the map was undertaken using the EPPI-Centre's online software, EPPI-Reviewer. All keyworded studies were then uploaded from EPPI-Reviewer to the EPPI-Centre database REEL (Research Evidence in Education Library) for others to access via the website.

2.2.5 Identifying and describing studies: quality-assurance process

Application of the exclusion criteria to titles (and, where available, abstracts) was carried out by one member of the Review Group. For quality-assurance purposes a ten percent random sample of these papers was also screened by a second member of the Review Group, and any differences between the judgements made by the two Review Group members were discussed and resolved. In addition, a member of the EPPI-Centre also applied the exclusion criteria to a random sample of studies in the first stage of screening, and any differences with the member of the EPPI-Centre were discussed and resolved.

Application of the exclusion criteria to a full copy of the paper was conducted by pairs of Review Group members, working first independently and then comparing their decisions and coming to a consensus. This stage included papers that had been identified by handsearching. For quality-assurance purposes, a member of the EPPI-Centre also applied the exclusion criteria to a random sample of these studies, and any differences between the Review Group members and the member of the EPPI-Centre were discussed and resolved.

Keywording of the included studies was conducted by pairs of Review Group members working first independently and then comparing their decisions and coming to a consensus. For quality-assurance purposes, a member of the EPPI-Centre also keyworded a random sample of the included studies and any differences between the Review Group members and the member of the EPPI-Centre were discussed and resolved.

2.3 In-depth review

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

All the studies included in the systematic map were also included in the in-depth analysis.

2.3.2 Detailed description of studies in the in-depth review

Studies identified as meeting the inclusion criteria were analysed in depth using the EPPI-Centre's data-extraction guidelines (EPPI-Centre, 2002b) together with its data- extraction software, EPPI-Reviewer. Two categories of review-specific questions were added: the first dealing with the review-specific weight of evidence (see below), and the second with the scope of the study.

The three review-specific 'scope of the study' questions asked for each study are as follows:

(i) Does the study address classroom practice?

(ii) Does the study address pupil confidence?

(iii) Does the study address pupil competence?

There were three options given for the response: 'explicit', 'implicit' and 'no', each option then providing a free text box in which appropriate details could be entered.

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

Components were identified to help in make explicit the process of apportioning different weights to the findings and conclusions of different studies. Such weights of evidence were based on the following:

A Soundness of studies (internal methodological coherence) based upon the study only

B Appropriateness of the research design and analysis used, for answering the review question

C Relevance of the study topic focus (from the sample, measures, scenario, or other indicator of the focus of the study) to the review question

D An overall weight taking account of A, B and C

Each of these three components (A, B and C) was assessed as low, medium or high (scored 1 to 3 respectively) and an overall weighting for the study (composite D) was arrived at by taking the arithmetic mean of the three component assessments (rounded to the nearest whole number), so that a mean of 1, 2 and 3 yielded an overall weighting of low, medium and high respectively. A discussion of this method of arriving at the overall weighting can be found in section 4.1.

2.3.4 Synthesis of evidence

The responses to these three questions, the weight of evidence assessments, together with the responses to the generic questions used for the data extraction, were then used as a basis for producing a narrative synthesis to address the review question. Tables summarising the included studies are presented in Appendix 4.2 to provide readers with details of the included studies; further details of the included studies can also be found by consulting the uploaded data extraction, which is available on the EPPI-Centre database REEL which can be accessed via the EPPI-Centre website.

2.3.5 In-depth review: quality-assurance process

Data extraction and assessment of the weight of evidence brought by the study to address the review question were conducted by two people, working first independently and then comparing their decisions and coming to a consensus. As part of the quality-assurance process, a member of the EPPI-Centre data extracted a sample of the included studies. Any differences between the judgements made by the Review Group and the member of the EPPI-Centre were discussed and resolved.

3. IDENTIFYING AND DESCRIBING STUDIES: RESULTS

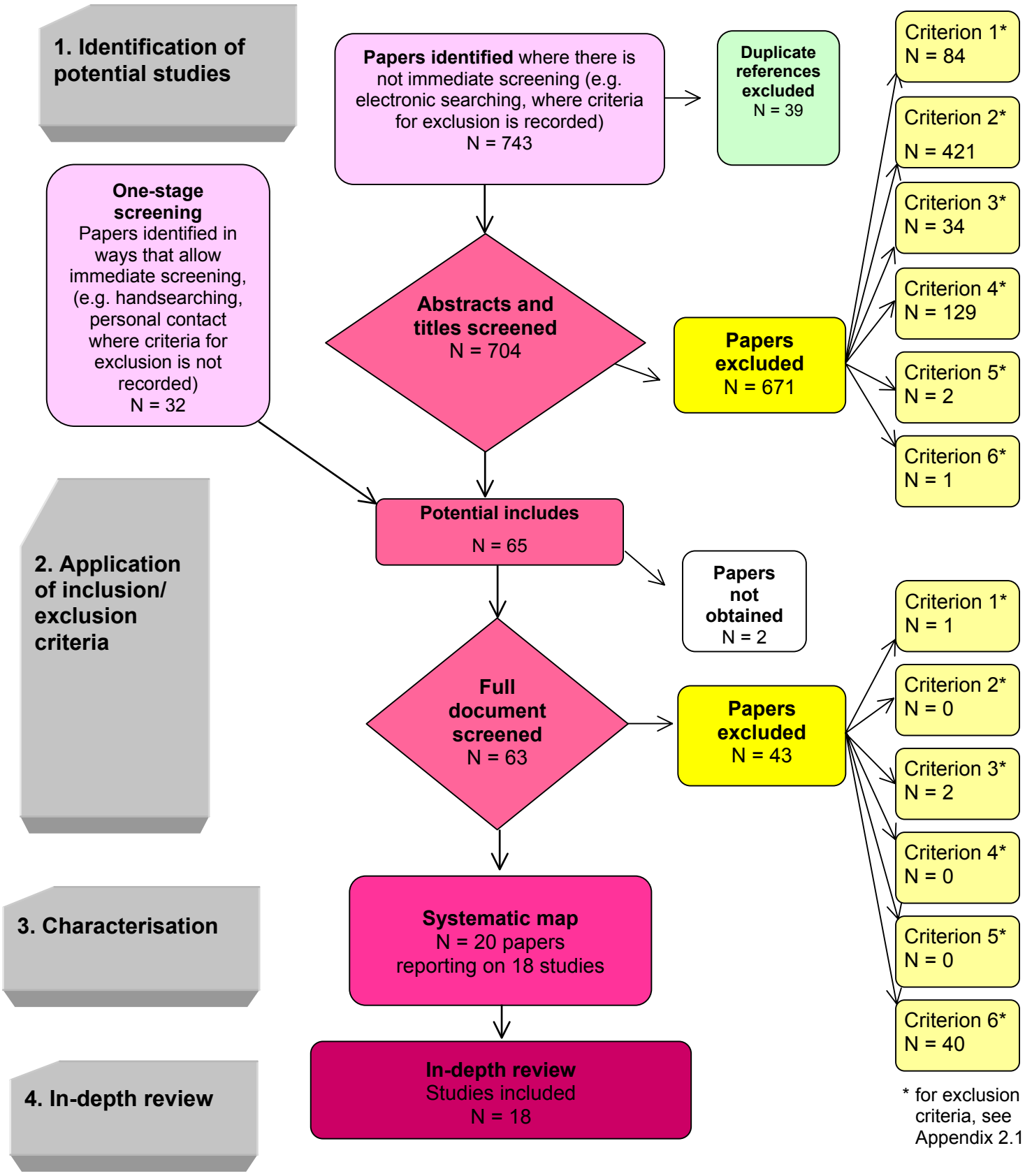
3.1 Studies included from searching and screening

After excluding duplicates, 704 papers were identified by an electronic search using the specified search strategy (the main review database). In the first stage of screening on titles and abstracts, the six exclusion codes were applied to these by a member of the Review Group resulting in 671 exclusions. The exclusion codes applied to each of these excluded papers are shown in Figure 3.1.

Full copies of the remaining 33 papers were then screened using the inclusion/exclusion criteria (Appendix 3.2). In addition, a further 32 papers were identified as a result of handsearching (Appendix 3.3) and were added to the main review database.

The six exclusion codes were applied to a full-length copy of 63 of the 65 papers identified. This resulted in a further 43 papers being excluded (Figure 3.1). As can be seen, almost all of these further exclusions were the result of applying exclusion code 6. However, one paper was clearly written for a practitioner audience and was excluded in terms of exclusion code 1, and consideration of a full length copy of two papers revealed that they did not contain any original data and were thus excluded in terms of exclusion code 3. A full-length copy of two papers was not available; these were both conference papers that had been presented orally without a written version. This resulted in 20 papers reporting 18 studies being identified for the systematic map (Appendix 3.1).

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



3.2 Characteristics of the included studies

The coding classification in the keywording provided a means for developing the systematic map for the 18 included studies. Tables based on the keywords are shown in Appendix 3.1. Eight studies were identified as a result of the electronic search strategy of BEI and ERIC. All eight studies were identified in BEI, none in ERIC. This is not surprising, given that the overwhelming majority of studies dealing with the NNS have not yet found their way into US-based publications and databases. It is interesting to note, however, that over half of the included studies came from other sources, including three studies identified through personal contact with the authors. This supports Black's (2004) observation that an over-reliance on an electronic search strategy based on keywords will almost certainly miss a number of important papers.

Only 10 of the 18 studies were published. This reinforces the view that, when the focus of a systematic review is on a recent initiative, recent conference papers will provide an importance source of research data relevant to the initiative.

All the studies were written in English. All the studies included data collected in England, although two studies had a comparative element: one study presented data from Wales (Jones, 2003) and another from Slovenia (Aubrey *et al.,* 2003).

All the studies had a curriculum focus on mathematics, but three studies also looked at the National Literacy Strategy (Earl *et al.,* 2002; Hardman *et al.,* 2003; Myhill, 2002).

Seven studies focused exclusively on pupils, but most studies had one or more other foci instead of, or in addition to, pupils: in eight cases, this was on teaching staff (Bibby *et al.,* 2003; Brown *et al.,* 2001; Earl *et al.,* 2002; Evans, 2001; Hardman *et al.,* 2003; Jones, 2003; Pinel, 2002; Raiker, 2002); in two cases, this was on student teachers (Basit, 2003; Huckstep *et al.,* 2002); and, in one case, this concerned parents (Baker and Street, 2003).

Half the studies were categorised in terms of study type as an 'evaluation'; all but one of these were evaluations of naturally occurring interventions (Bibby *et al.,* 2003; Earl *et al.,* 2002; Evans, 2001; Hardman *et al.,* 2003; Jones, 2003; McSherry and Ollerton, 2002; Pinel, 2002; Raiker, 2002) rather than researcher-manipulated (Hopkins and Pope, 2000).

The remaining nine studies were divided between the study types of 'description' (two studies: Basit, 2003; Denvir and Askew, 2001) and 'exploration of relationships' (seven studies: Aubrey *et al.,* 2003; Baker and Street, 2003; Bills, 2003; Brown *et al.,* 2001; Brown *et al.,* 2003b; Huckstep *et al.,* 2002; Myhill, 2002).

3.3 Identifying and describing studies: quality-assurance results

Quality assurance for the first stage of screening

A ten percent sample of the 704 papers (70 papers) identified in the first stage of screening was screened by a second member of the Review Group. There were six cases where the code that had been applied was queried and, on each occasion, the reason given by the first reviewer for the code used was agreed. In every case, this was because the first reviewer had more knowledge of the paper. For example, the first reviewer might recognise the name of the author of the paper and the study which was reported, and know this study was conducted in the USA, even though this information was not given in the title and/or abstract of paper, and as such was able to apply the most appropriate exclusion code at this stage, without needing to obtain a full-length copy of the paper.

A random sample of 30 papers out of the 704 papers identified by electronic searching was then screened by a member of the EPPI-Centre in London. All the 30 papers in this random sample had been excluded by the first reviewer, while the member of the EPPI-Centre in London included 16 of these. In every case this discrepancy was due to two main factors.

Firstly, the first reviewer had greater knowledge of these papers. For example, the title of one paper gave the name of a school, which the first reviewer was able to recognise as a secondary school where the study reported in the paper had been conducted, and not a primary school which would have warranted its further consideration.

Secondly, the first reviewer often had a full-copy version of potential relevant papers readily to hand and could thus, in a matter of seconds, identify a key feature of the paper that warranted an appropriate exclusion code. For example, the details available on the databases for 15 of these 30 papers did not include an abstract, which meant that the EPPI-Centre reviewer was having to reach a decision based solely on name of the author, title and publication details, and as such had to ‘include’ several papers as there was ‘not enough information to exclude’. In contrast, the first reviewer was able to consult an abstract and/or a full copy of the paper.

Ideally, in the first-stage screening, the reviewer should have available the name of author, the title, the publication details and the abstract for each paper, but a number of databases unfortunately do not provide an abstract and, when they do, this is often only a condensed version written by the database provider of an original and much longer abstract written by the author.

This quality-assurance check clearly indicates how a reviewer experienced in the field with ready access to copies of the papers is able to screen out much more efficiently a number of papers at this first stage, although it is accepted that in some cases the papers screened out at this first stage by the first reviewer could strictly speaking be regarded as a second-stage exclusion (i.e. excluded after consideration of a full-length copy of the paper rather than just on the basis of the title and/or abstract which was obtained by the electronic searching using the specified keywords). The difference between the first reviewer and the EPPI-

Centre reviewer was also markedly enhanced by the fact that this review included a substantial amount of handsearching which had largely been completed prior to the first reviewer undertaking the first stage screening. This often meant that the first reviewer had recently considered a full length copy of a paper just prior to coming across the same paper in its summary form (that is, author and title with or sometimes without a short abstract) in the main database drawn up from the electronic search using the keywords.

In terms of quality assurance, however, what is important is whether the procedure followed resulted in any papers being excluded during the first-stage screening process that should and would have been included in the descriptive mapping after consideration of a full length copy of the paper. In the case of these 30 papers, none of these papers fall into the category of being an erroneous exclusion in terms of the identification of papers to be included in the descriptive mapping.

Quality assurance for the second stage of screening

The six exclusion codes were applied to a full-length copy of these 65 papers by two members of the Review Group working independently and then comparing their codes. There were only three cases of disagreement, which were resolved after further consideration of the papers involved. A random sample of 7 of these 65 papers identified for second-stage consideration was sent to the EPPI-Centre reviewer for quality-assurance purposes. In six cases, the code applied by the EPPI-Centre reviewer agreed with the code applied by the two internal reviewers. In the seventh case, after further clarification of one particular detail in this paper, the EPPI-Centre reviewer agreed with the code applied by the two internal reviewers.

Quality assurance for keywording

Keywording was carried out by two members of the Review Group, working independently on each of the 20 studies. Their versions were then compared and an agreed version was then uploaded. A sample of three studies was also keyworded by a member of the EPPI-Centre staff in order to act as a further check on the process. Any differences were discussed and resolved. No major problems were encountered here. However, most discussion centred on describing the type of study reported (‘description’, ‘exploration of relationships’, and ‘evaluation’). It was agreed to interpret these three categories as hierarchical, and to code each study using the highest category where appropriate: that is, if a study involved both ‘exploration of relationships’ and ‘evaluation’, it would be coded for keywording purposes as ‘evaluation’.

4. IN-DEPTH REVIEW: RESULTS

4.1 The in-depth studies

Twenty reports of 18 studies met the inclusion criteria for the in-depth review. The weight of evidence assessments for each study are shown in Appendix 4.1 (a). The scope of the study assessments for each study are shown in Appendix 4.1 (b). Details of these studies are shown in Appendix 4.2.

Key findings of the included studies

As noted in section 1.6, it quickly became obvious in undertaking this systematic review that in order to address the review question we could not assume that the Daily Mathematics Lesson in the context of the NNS was being uniformly implemented in classrooms across England in the way intended by the NNS. As such, our review question does not simply or primarily have a straightforward focus on measures of effectiveness using data on narrowly conceived pupil learning outcomes. Rather, it is also concerned with the classroom practice that has taken place following the implementation of the NNS, and what we can infer from such practice about pupils' learning of mathematics, the confidence they show in the classroom setting and their performance on classroom tasks as well as on national assessments. This consideration of both classroom practice and pupils' experience of learning mathematics is reflected in our analysis of the included studies.

This section presents key details of the nature of the findings and conclusions presented in each of the 18 studies. In this section and the following section (the synthesis), we have decided not to use subheadings as the inter-relatedness of the three elements of our review (classroom practice, pupil confidence and pupil competence) simply did not allow for the sort of sign-posting that might be appropriate for other review questions. As such, we have adopted a continuous narrative.

A summary of included studies tables is shown in Appendix 4.2. The summary tables deal with four key aspects of each of the 18 studies:

(i) the focus of the study

(ii) the data collected

(iii) key claims/evidence regarding the classroom teaching of early mathematics

(iv) key claims/evidence regarding pupil confidence and competence in early mathematics

Aubrey *et al.* (2003) compared the mathematical performance of English and Slovenian 5–6 year-old pupils using the Utrecht Early Numeracy Competence Test. This test comprised eight sub-topics each with five items. Three versions of the test were each taken by approximately 100 pupils at three points in time between the ages of 5 and 6. At this age, the Slovenian pupils were in pre-primary school provision. A multilevel analysis was used which incorporated the

hierarchical structure of the data collected, with pupils nested within institutions, within the two countries. The dependent variable was test score and the explanatory variables were age, sex, country and moment of testing. The analysis presented focused primarily on the effects of age on the overall score in the test and the relative performance in the relational and numerical tasks. They found that a model regressing score on age reduced the variance at school level by 42.2%, indicating that much of the difference between scores for a single child was explicable in terms of improvement with age. A further model allowed for different fixed effects in each country at each moment of testing, which showed that Slovene and English pupils performed equally well overall. Further analysis showed that English pupils were more successful on numerical tasks and less successful in relational tasks compared with the Slovene pupils. The authors concluded that: (i) since the Slovenian pupils did as well as the English pupils, their carefully prepared pre-primary provision is at least as effective as the NNS in English primary schools; and (ii) the curriculum goals and pedagogical practices of the NNS may account for the English pupils' superior performance on numerical items but this may be short lived. The Slovenian pupils' better performance on relational items is not accounted for by the authors, although it is stated that their pre-school preparation stresses conceptual understanding of time, space, size and quantity.

Baker and Street (2003) explored the links between home and school numeracy practices. They report an ethnographic case study of a primary school pupil (Seth) based on home and school visits over a three-year period (from reception year to Year 2). He is described as being from a white, working-class home. His mother did not have a job but attended adult education classes. Seth is described as being one of three 'demanding' children. The authors found that: (i) school numeracy practices were very different from home numeracy practices; (ii) the Year 1 teacher had control over choice of activities, knew the answers and how pupils were to do the tasks; and (iii) school practices had a clear educational focus (e.g. to learn subtraction), while home practice was embedded in real-life purposes. Seth is reported as making mistakes with finger-counting and with moving off-task in class. He and his peer group found a change giving task difficult. The authors express concern that differences in home and school practices can undermine transfer of learning from school to home and can undermine pupil confidence if the school tasks are not embedded in a purposeful and meaningful context which the pupil has experienced and can relate to. The authors conclude that the difference between home and school numeracy practices may explain Seth's disengagement from numeracy at school; he did not identify with school in general and numeracy in particular.

Basit (2003) explored the effects of the NNS on primary student-teachers by conducting two in-depth interviews with 30 final year BEd students who were non-mathematics specialists. This group of trainees are interesting because they had completed three years of training prior to the introduction of the NNS and then a fourth-year school experience during the year in which in which the NNS was applied. Her findings are based on the perceptions reported to her by these students; the findings are thus presented in the form of supporting quotations from her interview data. She found that (i) the NNS framework has boosted trainees' confidence; (ii) some experienced teachers do not follow the three part lesson to the letter; (iii) the three part lesson forces teachers to think about the lesson as a whole; (iv) the NNS framework is easy to use but is sometimes over ambitious regarding what can be covered in the time allocated for each topic; (v) on the one

hand, restricted time for activities does not allow an activity to be continued over several consecutive lessons, but, on the other hand, it helps with time management of topics that need to be covered; (vi) the mental starter gets pupils involved and excited about mathematics; (vii) restricted time for activities can disadvantage low attainers as teachers have to move on to another topic before they have adequately grasped the current one; and (viii) the NNS assumes pupils have reached a certain level at the end of the year and have to be taken on from there at the beginning of the next year when in fact many pupils have considerable gaps in their understanding. Basit concludes that: (i) the NNS was viewed positively by the trainee teachers; (ii) they embraced the NNS whole heartedly; and (iii) we need to wait and see if the NNS has a genuine impact on pupils' mathematical ability.

Bibby *et al.* (2003) explored primary school teachers' interpretations and implementations of the teaching of mental calculations in the Daily Mathematics Lesson. Fifty-nine lessons were observed (19 KS 1, 40 KS 2), each followed by brief discussion with the class teacher, involving six schools from three LEAs. They found that: (i) the three part lesson was being widely used, with a tendency towards rather long mental/oral starters with brief (or omitted) plenaries; (ii) there was little evidence of strategic thinking (developing a repertoire of mental and written calculation strategies and informed decision making about the use of these) being encouraged; (iii) the majority of the lesson involved recall (e.g. learning multiplication facts) or procedures (formal, written procedures and teacher led procedures with informal methods); (iv) there were very few instances of teachers responding to and building on pupils' answers and questions; (v) there was little evidence that pupils are better able to think strategically; and (vi) pupils were primarily learning mental arithmetic (in the traditional sense of an emphasis on recall of answers and methods) rather than mental calculation in the spirit of the NNS. The findings of this study are particularly important as, when the NNS was introduced, the development of strategic thinking was flagged up as a major change from previous teaching. The authors conclude that: (i) teachers have increased the amount of time and emphasis on mental work in mathematics but the emphasis is on recall and procedures rather than strategy; (ii) there was very limited evidence of good practice; and (iii) teachers need more support in understanding the principles behind the emphasis on mental calculation.

Bills (2003) explored the language pupils used when asked 'What is in your head?' after they had performed a mental calculation. Data collection involved lesson observations and interviews with Year 2 to Year 4 pupils, leading to an analysis of over 1,000 responses given by 26 pupils in six (group) interviews. Bills found that: (i) teachers can take account of pupils' use of language to indicate pupils' level of confidence with taught rules in their classroom teaching; (ii) pupils' adoption of a classroom speech style involving the use of 'you' and the present tense is an indication that they have internalised the rules and follow them when making mental calculations; (iii) pupils' use of explanatory expressions indicates a confidence in and familiarity with procedures; and (iv) both the above indicators correlated with accuracy in calculations. Bills concludes that pupils' language provides a good indicator of pupil confidence and competence and that it can be used by teachers to inform the appropriate teaching strategy to adopt in the light of such indicators. This study provides evidence of how pupils' use of language can indicate their confidence in following taught rules, but no evidence was provided of the extent to which teachers are doing this and/or the impact that doing so had on pupil confidence and competence.

Brown *et al.* (2001) explored whether a particular type of teacher or way of teaching produces higher learning gains in numeracy. This study was part of the Leverhulme Numeracy Project, which was a five-year longitudinal study (1997–2002) involving one core project and various focus projects designed to explore factors leading to low attainment in primary numeracy and testing out ways of raising attainment. The core project tracked two cohorts of pupils, one starting in reception year and the other starting in Year 4 (c. 1,600 pupils in each), and monitored their performance on a sequence of numeracy tests. Each school was visited once a year and a lesson observation was undertaken for each class, interviews took place with teachers, heads and coordinators and a teacher questionnaire was conducted. Part of this study focused on data for Year 2 pupils tested at the start and end of 1999/2000 (the first year of the NNS); it was this part of the study which has high relevance for our review question. Teaching was measured in terms of Saxe lesson observation evaluation scores based on four main parameters: *task* (e.g. mathematical challenge); *talk* (e.g. extent to which teacher talk focuses on mathematical meanings; *tools* (e.g. range of modes); and *relationships and norms* (e.g., extent to which there is a community of learners). The measures of pupil learning gains used numeracy tests developed by the research programme. Pupils were tested in a sequence of tests twice a year, and the characteristics of teachers and teaching were correlated with attainment gains. These data were drawn from 10 schools in four LEAs and involved 75 classes. They found that: (i) teachers were already using a high proportion of whole class teaching at the start of the study; (ii) no teachers were using individualised work patterns; (iii) there was no benefit associated with a higher proportion of whole class teaching; (iv) the three different lesson evaluation models were not reliable in predicting numeracy gains: the correlation between Year 2 gain and the Saxe score was low ($r = 0.18$) (this correlation was only marginally higher than the one obtained in the previous year,1998/1999 when the pupils were in Year 1, $r = 0.08$); and (v) it may be very hard to adequately characterise effective teaching of numeracy and to construct a reliable lesson observation based on such a characterisation. Pupil gain scores for Year 1 pupils at the end of 1998/1999 and for Year 2 at the end of 1999/2000 are presented, but no comparison is made here to indicate whether this provides any evidence that Year 2 pupils made a greater gain (in the first year of the NNS) than Year 1 pupils did (in the year prior to the introduction on the NNS); however, this is addressed in Brown *et al.* (2003b). The authors conclude that: (i) the effectiveness of whole class teaching is much less strong that claimed; (ii) the effects of certain teacher behaviour is small and variable; and (iii) the rise in standards, if there is one, is unlikely to result from changing pedagogy but from the curriculum taught.

Brown *et al.* (2003b) in a subsequent study as part of the Leverhulme Numeracy Project (see above) explored pupil progress in numeracy and the links between test data to lesson observation data for a sample of 30 case study pupils. The study looked at pupil numeracy test data (progression in facility values on test items) and related this to classroom observation data for the case-study pupils. They found that the data from the lesson observations indicated some variability between teachers and the effects this has on outcomes; for example, a teacher employing a more relaxed and supportive approach towards pupils having problems helped lower attainers to recover in terms of their confidence and competence, but this was at the expense of the rest of the class which was not being challenged enough and the class as a whole made less progress overall during the year. Data were also presented on common items that appeared in the

numeracy tests to graph their items facility scores at the beginning and end of each year for the first cohort (Year 1 to Year 4), and for the second cohort (Year 4 to Year 7). They found that: (i) the mean increase in facility for Year 1 pupils at the end of 1998/1999 (the year before the NNS was introduced) and for Year 2 at the end of 1999/2000 was the same (20 percent increase in both cases) and this was higher than for older year groups; and (ii) the progress of the case-study pupils display idiosyncratic trajectories. The authors concluded that pupil progress, whether looked at as a whole or in terms of individual case studies, does not follow a simple pattern and reflects a complex mix of factors at work. It is particularly interesting to note here that the authors argue that the improvement in numeracy standards observed just prior to the introduction of the NNS is likely to be a reflection of the impact of the regime of national testing and the way what is taught more closely matches what it tested; this will be a more enduring factor in influencing test results than the introduction of the three part lesson and its associated features (most importantly the use of interactive whole class teaching).

Denvir and Askew (2001) explored the ways in which pupils manage their participation during whole class interactive teaching based on 150 lesson observations, with a particular focus on the mental and oral starter. The data presented here deal with three case-study pupils. They found that the three part lesson included a substantial element of public answering of questions and that this involved a strong 'performative' element which stresses speed and correctness. The authors conclude that the emphasis on whole class interactive teaching encourages pupils to participate but they may often be doing so in a way that undermines the quality of their mathematical thinking. This finding is extremely important as it indicates that, in whole class interactive teaching, pupils may not be participating in the mathematical thinking intended. Indeed, the strong performative element prompts pupils to adopt behaviour which militates against them developing good habits as learners; they may be *participating* in the activities rather than *engaging* in the mathematical thinking, and the pace of the lesson may disguise this distinction from the teacher. This study is very relevant to the review question as it shows how type of participation may mask whether competence and confidence are being acquired. When taken in conjunction with those studies which rely on teacher report (e.g. Basit, 2003; Jones, 2003), this study highlights the difficulty of inferring understanding from the appearance of participation. Although this study was too small to be assigned high weight of evidence, it may be unique in the relevance of the questions which it is attempting to address which are crucial to the review.

Earl *et al.* (2002) explored what can be learnt from the national evaluations of the National Literacy Strategy (NLS) and the NNS which they conducted. In the second year of the national evaluation they surveyed 500 schools, visited 10 schools and conducted 200 interviews. They found that: (i) teachers and headteachers were very positive about the NNS; (ii) the three part mathematics lesson provided more focus, is easier, more disciplined and progression is known; (iii) teachers were motivated to improve their teaching, and feel their skills were now sharper and have better techniques; (iv) teachers were making use of NNS materials; and (v) pupils' learning is enhanced because the need for differentiation in the middle phase of the three part lesson and the plenary session both help teachers to link new learning to what pupils already know. The authors conclude that the NNS has made a significant change in primary education throughout England in a very short time and the change has been pervasive with substantial early success. However, they also note that there is a danger that rigid

compliance with key features of the three part lesson and a lack of subject and pedagogical knowledge can lead to ineffective teaching. When considered in conjunction with the study by Jones (2003), the concern is that a superficial adherence to the NNS, an externally imposed intervention, may undermine the sustainability of change.

Evans (2001) explored how small rural schools were implementing the NNS, with a focus on mixed age classes with three or more year groups (e.g. containing reception year to Year 2 pupils). Seventeen teachers were interviewed about their strategies for organising and carrying out daily mathematics lessons. The study found that: (i) teachers reported a combination of positive and negative comments about the NNS; (ii) the NNS framework sometimes did not provide enough time for pupils to grasp a topic adequately, but it does help teachers to avoid spending far too long on a topic; (iii) whole class teaching, particularly mental/oral starters and plenaries, was problematic; (iv) able older pupils and least able younger pupils can get easily bored and find the pace is inappropriate for them; (v) about half the teachers did not always ensure they carried out a plenary session, with many teachers feeling it was better to ascertain what had been learnt during the main activity; (vi) the use of teaching assistants was particularly helpful, and made it easier to use segregated activities; (vii) several teachers felt the emphasis on mental mathematics had increased pupils' confidence, particularly for the less able pupils; (viii) teachers felt the NNS had improved the ability of pupils to spot different ways of working out calculations; (ix) pupils in mixed aged classes often spent too much time waiting or on holding activities; and (x) the needs of able older pupils may not be adequately met, but able younger pupils may be stimulated by the presence of older pupils. Evans concludes that: (i) implementing the NNS in small rural schools is problematic but not unmanageable; (ii) the study provides useful practical information that will be of help to other teachers and to student teachers; and (iii) the study raises two issues that require further attention: catering for the needs of the most able oldest pupils in a mixed age class and reconciling the philosophy of the NNS with that of the foundation stage curriculum.

Hardman *et al.* (2003) explored patterns of whole class interaction in the NLS and NNS, and whether effective teachers of literacy and numeracy differ in their discourse strategies. Data were collected on the discourse strategies used by 72 teachers (35 literacy, 37 numeracy) based on systematic classroom observation, discourse analysis of lesson transcripts and teacher questionnaires; half the teachers were selected as highly effective and the other half as average, based on value-added data on their pupils' progress. The main focus of the data dealt with the discourse moves of teachers and pupils. However, the study also collected data on teachers' understanding of interactive whole class teaching, their perceptions of the discourse strategies they use, and their views about the quality of training they had received for interactive whole class teaching. They found that: (i) traditional patterns of whole class interaction have not been dramatically transformed by the NLS and NNS; (ii) whole class teaching in the NNS was intended to be more interactive and to promote high quality of dialogue and discussion, but most of the discourse in numeracy lessons involved teachers asking closed questions, evaluating, explaining and directing, and pupils answering a question; (iii) teachers had no clear concept of what interactive whole class teaching meant and, although they reported that they valued it and frequently invited pupils to elaborate on their answers, this was not supported by the lesson transcripts; (iv) effective teachers had a more interactive style in terms

of quantity of interactions, but not in terms of quality of interactions; and (v) pupils' answers were usually short (three words or less) and pupil contributions were rarely sustained or extended to encourage higher cognitive interactions; this is unlikely to foster greater confidence and competence in their ability to articulate their ideas regarding their understanding of mathematics. The authors conclude that (i) the NNS has not transformed traditional patterns of whole class interaction; (ii) far from encouraging and extending pupil contributions to promote higher levels of interaction and cognitive engagement, most of the questions were designed to funnel pupils' responses towards a required answer and were only rarely used to assist pupils to more complete or elaborated ideas; (iii) this was more marked in the numeracy than in the literacy lessons; (iv) teachers had no clear concept of what interactive whole class teaching meant or shared language to discuss it, and had been given little practical guidance; and (v) although most teachers reported that they had valued it and frequently invited pupils to elaborate on their answers, this was not supported by the observations. This study is extremely important as it deals with one of the key features underpinning the NNS: the NNS framework states that 'high quality direct-teaching is oral, interactive and lively ... in which pupils are expected to play an active part by answering questions, contributing points to discussion, and explaining and demonstrating their methods to the class' (DfEE, 1999, p 11). This study indicates that the type of interaction occurring is not as intended by the NNS. Considered in conjunction with the study by Bibby *et al.* (2003), this study raises serious concerns about classroom practice in the Daily Mathematics Lesson.

Hopkins and Pope (2000) explored whether the use of educational television programmes can improve numeracy. They used a researcher-manipulated evaluation based on an experimental design (the only included study which did so) to compare pre-test and post-test assessment task scores for experimental groups (four year 1 classes) and control groups (three Year 1 classes) involving 157 pupils in four schools and classroom visits to obtain background information. They found that: (i) the use of five television programmes (each about 10 minutes in length) over a two-week period based on key NNS objectives (e.g. counting, reciting numbers, odds and even) coupled with associated worksheets can be incorporated into the three part lesson format successfully; (ii) school visits indicated that the pupils in the experimental group showed enthusiasm for the television programmes and were active watchers; and (iii) overall, the evidence that the use of the television programme led to greater gains was inconclusive – the main gain was higher for the TV groups than control groups in two schools, lower in one school and, in a fourth school, the post-test was not taken because of teacher absence – but there was some evidence to indicate that some low attaining pupils in the experimental classes did improve substantially (for all the TV groups but not the control groups, small groups with low pre-test scores showed substantial improvement). The authors conclude that the evidence that the television programmes had enhanced gains in numeracy was inconclusive but it did suggest it might be helpful for low attaining pupils. However, an additional and interesting finding was that watching the TV programmes broadened pupil gains across a wider spectrum of assessment items. This is an important finding, as it suggests that a series of well balanced TV programmes (i.e. one that covers the full range of mathematical objectives in the National Curriculum) can enrich the mathematical learning of pupils in a way that can counteract any tendency of teachers to gear a disproportionate amount of curriculum time to those mathematical objectives that will be assessed in national tests.

Huckstep *et al.* (2002) explored whether student teachers' subject matter knowledge, or lack of it, is evident in their teaching, with a focus on *connecting*, *responding* and *exemplifying*. Their study was based on videotapes of 24 mathematics lessons, plus a *descriptive synopsis* of the lesson written by an observer, which was then elaborated into an *analytic account* after viewing the videotape. The person observing the lesson was a post-graduate certificate in education (PGCE) course tutor, and the descriptive synopsis of the lesson was 400 to 500 words in length. They found that: (i) the three part lesson structure has been adopted; (ii) examples of both taking or missing opportunities for making connections between mathematical ideas appears to reflect the strength or weakness, respectively, of certain mathematical subject matter or pedagogical content knowledge; (iii) positive and negative examples of responding to pupils' articulated ideas indicate that these are enabled (or otherwise) by their mathematical content knowledge for teaching; and (iv) the wisdom (or otherwise) of trainees' choice of examples when teaching was a strong indicator of their mathematical content knowledge for teaching. The findings were evidenced in illustrative examples which were given to indicate how more effective teaching is grounded in teachers' subject matter knowledge, with lesson vignettes and extracts from teacher-pupil interaction which demonstrate how more effective teaching in terms of *connecting*, *responding* and *exemplifying* can aid the development of pupil confidence and competence. However, the effects on pupil outcomes in these examples are implicit rather than explicit. The authors conclude that the data provides useful evidence of how trainees' subject knowledge, or the lack of it, was evident in their teaching. This study is particularly important because it indicates how effective interactive whole class teaching of the type advocated in the NNS crucially depends on the quality of teachers' subject matter knowledge.

Jones (2003) compared the impact of the NNS in England with the numeracy initiative in Wales. This study was based on questionnaires completed by 243 and 249 headteachers (or a delegated member of staff) in England and Wales respectively. Jones found that: (i) the overwhelming majority of headteachers in England reported that the NNS had improved teacher motivation and the quality of teaching (the pattern of responses from Welsh headteachers regarding the Welsh initiative was similar); (ii) more frequent use was reported in England (compared with Wales) of the three part lesson, whole class teaching, pupils working as groups, and teaching of mental mathematics; and (iii) the overwhelming majority of head teachers in England reported that the NNS had improved pupil motivation and pupil attainment in numeracy (the pattern of responses from Welsh headteachers regarding the Welsh initiative was similar). Jones concludes that the NNS in England was more prescriptive than the Welsh approach and, as a result, has had a more uniform impact on classroom practice, but both initiatives seem to have had a similar effect in enhancing pupil motivation and attainment (as reported by headteachers). The fact that this study indicates that, in England, teachers were making greater use of the three part lesson, whole class teaching, pupils working as groups, and teaching of mental mathematics, compared with Wales whereas both initiatives appeared to have been equally successful begs the question of whether it is these features of NNS classroom practice in England that have brought about these gains, or whether this study indirectly provides evidence that other factors with the NNS being adopted in England account for the gains which have occurred.

McSherry and Ollerton (2002) explored school policies towards pupil grouping for numeracy lessons based on questionnaires completed by 192 primary schools. They found that: (i) the overwhelming majority of schools group pupils by ability for all or part of the time: by ability within the class (55%), setting across different classes (20%), by ability for part of the time (15%) and mixed ability (10%); (ii) a marked increase in ability grouping occurred in the two years prior to and in the year following September 1999 (when the NNS was introduced) in order to implement the NNS and/or to improve SAT scores; (iii) grouping by ability was seen to make whole class teaching and provide differentiated work more manageable; and (iv) grouping by ability is occurring at an increasingly younger age (in Year 2, 14% of schools set and 47% employ within-class ability groups). The authors conclude that grouping by 'ability' is happening at an increasingly younger age and expressed a concern that some pupils could spend seven years in 'low' ability groups in mathematics in their primary school years and another five years in 'low' ability groups in their secondary school years, thereby living under the label of being less able in mathematics and are likely to become mathematically unconfident and disaffected. The authors note that the study raises questions about the appropriateness of grouping and about effective issues of grouping.

Myhill (2002) explored whether patterns of classroom interaction for boys and girls can explain boys' under achievement. The study was based on teacher interviews, pupil interviews and classroom observation, using structured observation schedules in 36 classes (six classes from each of Years 1, 4, 5, 8, 9 and 10). The 144 pupils (72 boys and 72 girls) in the study were made up of one high-achieving boy and girl, and one low-achieving boy and girl selected from each class. The observation schedule made a distinction between whole class teaching episodes and teaching episodes where pupils are engaged in individual, pair on group work; both literacy and numeracy hours in primary schools are included in the data. The data presented in the paper focuses on the whole class teaching episodes based in 106 teaching sessions. She found that: (i) whole class teaching episodes included 'joins in collective response', 'puts hand up' and 'answers question after invitation' (these were taken to indicate positive interactions); (ii) for the category 'answers question after invitation', the Year 1 underachieving boys participated more often than other pupils, which was interpreted as indicating that teachers use this strategy during whole class teaching to increase the participation of underachieving boys; (iii) underachievers were less likely to participate positively in the classroom and more likely to engage in off-task interactions; (iv) this was evident for underachieving boys in Year 1 for the category of 'joins in a collective response' (which is a frequent feature of the teaching repertoire); and (v) for the category 'puts hand up', there was no gender difference amongst year 1 underachieving pupils. Myhill concludes that: (i) there is a significant relationship between underachievement and level of interaction and response in whole class teaching episodes; and (ii) it is important not to look at gender alone, but to take account of other variables at work: not all boys and girls have similar needs. Taken in conjunction with the study by Denvir and Askew (2001), we can infer that, if underachieving pupils tend not to participate, then they are even less likely to engage in mathematical thinking.

Pinel (2002) explored the occurrence of catechetics (question-and-answer interactions) in mathematics lessons based on lesson observations in 12 schools (pupils' ages ranged from 4 to 14 years) together with teacher self-evaluation of the lesson. The study set out to categorise teachers' question and answer

interactions in terms of their richness. Rich interactions are characterised by higher order questions involving pupils using reasoning, considering structures and reflecting on mathematics. Lower order questions require recall of known methods, specific answers or facts. The study was also interested in the incidence of pupil-initiated episodes. He found that: (i) teacher-directed catechetics during the oral/mental starter occurred very frequently in some lessons and hardly at all in other lessons; (ii) very few lessons involved inverse catechetics (i.e. pupils' use of question-and-answer with their teacher or with peers); (iii) lessons also differed in the extent to which they involved higher order, middle order and lower order questions; and (iv) teachers appear to think more interactions of this type are occurring in their lessons than the lesson observations bear out (this finding is in line with other studies which report that teachers overestimate how much interaction is occurring in their lessons). Pinel concludes that teachers' prior practices seemed to be more enduring than the more interactive approaches, such as drawing out learners' ideas and methods and allowing space for their questions, as advocated in the NNS. Pinel, however, notes that the teachers expressed their allegiance to the NNS and evaluated their lessons as following the NNS approach.

Raiker (2002) explored the extent to which the mathematical vocabulary used in three part mathematics lessons can cause problems in the teaching and learning of mathematical concepts. The data were collected during one numeracy lesson for each of six classes – one Year 2 class, two mixed Year 3/4 classes, and three Year 4 classes – and discourse was collected from teacher-class interactions during the mental/oral starter and the introduction to the main teaching part of the lessons. The findings are based on the discourse analysis of six groups of teachers and pupils. She found that: (i) teachers use mathematical vocabulary more than pupils do; (ii) problems can arise in lessons when the meaning of mathematical words is not established; (iii) teachers did not appear to be aware of the importance of the mathematical vocabulary they were using or to have planned for its introduction, explanation of meaning, and repetition; and (iv) teachers did not distinguish between differing purposes of language in a mathematics lesson and therefore did not alert pupils when essential key vocabulary was being taught. She concludes that problems in the teaching and learning of mathematical concepts in part arise from and are compounded by the spoken language involved. These demonstrate that mathematical language plays a crucial role in the building of sound concepts and the subsequent development of mathematical thinking. She also notes that the data indicates that pupils' understanding may be more insecure than is evident from their success in completing worksheets because they may not have grasped the meaning of the mathematical vocabulary that is being used.

Characteristics of the included studies

As noted in Chapter 3, seven studies focused exclusively on pupils, but most studies had one or more other foci instead of or in addition to pupils. Nine of the studies were categorised in terms of study type as an 'evaluation'; the remaining nine studies were divided between the study types of 'description' (two studies) and 'exploration of relationships' (seven studies).

The data extraction and the details included in the summary of included studies tables show that the 18 studies used a range of different methods for data collection. Fifteen of the studies were dominated by or exclusively used one

particular type of data-collection method: questionnaire data (Jones, 2003; McSherry and Ollerton, 2002); interview data (Basit, 2003; Evans, 2001); classroom observation data (Bibby *et al.,* 2003; Bills, 2003; Denvir and Askew, 2001; Hardman *et al.,* 2003; Huckstep *et al.,* 2002; Pinel, 2002; Raiker, 2002); case study data (Baker and Street, 2003); and mathematics test data (Aubrey *et al.,* 2003; Brown *et al.,* 2001; Hopkins and Pope, 2000). Three studies employed a combination of data-collection methods (Brown *et al.,* 2003b; Earl *et al.,* 2002; Myhill, 2002).

What was also evident from the data extraction and from the summary of included studies tables is that only nine of the 18 studies explicitly address all three key aspects of the review: classroom practice, pupil confidence and pupil competence (see Appendix 4.1 (b)). In the case of classroom practice, 17 studies addressed this explicitly. However, as far as pupil outcomes are concerned, there were considerably more data pertaining to pupil competence than to pupil confidence. Fourteen studies address pupil competence explicitly, while only nine studies address pupil confidence explicitly. Moreover, much of the evidence regarding pupil confidence is inferred from observation of or comments about pupils' general behaviour and attitudes rather than through the use of robust measures of pupil confidence *per se.*

Weight of evidence results

Eleven studies received an overall weight of evidence rating of 'high', and the remaining seven studies were rated 'medium' (see Appendix 4.1 (a)). It needs to be borne in mind, however, that the use of a three-point rating scale (high, medium and low) for each of the components A, B and C, means that each band is fairly broad. Moreover, a simple algorithm of taking an average of the three component grades (scored 3, 2 and 1 respectively) was used to arrive at the composite overall weight of evidence grade. As such, it is not perhaps surprising that, as these 18 studies have already 'survived' a careful and rigorous selection process based on applying explicit inclusion/exclusion criteria, more than half of them have received an overall weight of evidence rating of 'high'. However, one of these 11 studies, rated as 'high' overall (Jones, 2003), had a medium rating for the appropriateness of the design and analysis in answering the review question (component B), but was rated high overall using this algorithm. As such, this study could perhaps be regarded as a 'qualified' high overall, in terms of answering the review question.

On the issue of weight of evidence, we are aware that some authors of systematic reviews take the view that, for a systematic review dealing with 'impact', the grading of the weight of evidence for component B (appropriateness of the research design and analysis for the review question) can only be rated as 'high' for impact when an unbiased investigation using a high quality experimental research design was used. The use of systematic reviews within social sciences, however, begs a number of important questions about what type of studies can provide good research evidence of impact. We have taken the view that a study which adopts a research design and analysis that is capable of making a significant contribution to informing the reader of the nature of the impact (e.g. questionnaire, interview or classroom observation data which address the extent to which teachers are making use of the type of interaction with pupils intended by the NNS) can provide a trustworthy and insightful evidence base. Where our

answers to the data-extraction questions have led us to this conclusion, we feel a rating of ‘high’ for component B is valid.

If the weight of evidence assessment for component B addressed whether the research design and analysis adopted by the study provides unequivocal evidence of cause and effect in relation to our review question solely in terms of pupil learning outcomes, then none of these studies would have been rated high on this component; and some researchers would argue that, in a systematic review concerned with impact, a study that is not rated high on component B should not be rated high overall.

However, our assessment of component B was based on the study’s appropriateness and usefulness in terms of how well the research design and analysis can provide data which will inform our understanding and judgement of whether the classroom practice taking place and pupils’ learning experiences occurring in the context of the Daily Mathematics Lesson are likely to have enhanced their competence and confidence. As such, our view is that, in principle, there is no reason why a questionnaire survey or a case study cannot be regarded as providing a basis for a highly appropriate research design and analysis for our review question. The reader thus needs to bear in mind that an overall rating of high adopted here for component B will not equate with a rating of high recorded in those systematic reviews of impact where a strict cause and effect model utilising an experimental design has been used as the measure of appropriateness and usefulness.

Moreover, and as indicated in section 1.6, our review question is as much about classroom practice as it is about learning outcomes. As such, the findings presented here concerning the development of pupils’ confidence and competence must be treated with some caution and regarded as tentative as inferences are drawn from evidence concerning classroom practice and pupils’ learning experiences. However, we do feel that those studies which attained an overall weight of ‘high’ do make a highly significant contribution to the evidence base for our review question.

4.2 Synthesis of evidence

A narrative synthesis of data from the included studies is adopted to address the review question based on the data-extraction process. This section presents an overall synthesis of the included studies which addresses our review question in the light of the analysis of the included studies. For the reasons noted in section 4.1 above, we have not used subheadings here.

An analysis of these 18 studies identified a number of key themes and issues. Importantly, it is clear from these studies that the impact of the pedagogy espoused by the NNS differs for different groups of pupils and that variables such as gender and ability add to the difficulty of producing simple answers to the review question. Moreover, all but one of these studies include explicit evidence of certain classroom practices which differ from those recommended in the NNS, so any conclusions on impact have to take into account what is commonly happening in classrooms rather than what is intended in the policy.

Most teachers welcomed the three part lesson format (oral/mental starter, main teaching and pupil activities, and plenary), and felt this structure gave them greater clarity and confidence in their planning and teaching. The three part lesson forces teachers to think about the lesson as a whole. Surveys and interviews with headteachers, teachers and student teachers all indicated that the Daily Mathematics Lesson appears to have helped to raise general standards in mathematics and to have raised pupil motivation (Basit, 2003; Earl *et al.,* 2002; Huckstep *et al.,* 2002; Jones, 2003). However, this is not enough to conclude that the impact on confidence and competence has been positive, since the data comprise the perceptions of people who have some investment in the implementation and who may or may not have conducted any systematic investigation to support their perceptions. We simply do not know on what evidence, if any, these perceptions are based, although school assessment data are likely to have made a contribution.

Although the three part lesson has been widely adopted, not all experienced teachers follow the three part lesson 'to the letter' (Basit, 2003) and there is a tendency towards long mental/oral starters, with brief (or omitted) plenaries (Bibby *et al.,* 2003). Indeed, there is some evidence that an over rigid compliance with key features of the three part lesson can lead to ineffective teaching (Earl *et al.,* 2002). Given that the guidance on pedagogy is not being followed to the letter, it is difficult to conclude that any impact on confidence and competence is a direct result of the implementation of the three part lesson as intended by policy-makers.

One drawback of the approach as advocated, however, is that topics are rarely developed and extended over several lessons. The NNS framework seems to be over ambitious regarding what can be covered in the time allocated for each topic (Basit, 2003). This can have the disadvantage for low attaining pupils that they have to move on to another topic before they adequately grasp the topic in hand. However, an advantage of this is that it helps teachers maintain better time management, and avoids 'doing a topic to death'. Another concern is that the NNS assumes pupils have reached a certain level at the end of the year and have to be taken on from there at the beginning on the next year, when in fact many pupils have considerable gaps in their understanding (Basit, 2003).

There is an important distinction to be made between traditional whole class teaching and the notion of interactive whole class teaching. The former relies heavily on a teacher-centred didactic approach, making heavy use of explaining and demonstrating; teacher questions and pupil answers are based on the teacher asking a high proportion of closed questions or questions requiring a simple recall of facts or procedures, and pupil answers are often short. The latter (interactive whole class teaching), which is what is advocated in the NNS, is an approach that is intended actively to involve pupils in the lesson through the use of more searching, higher-order questions which seek to challenge and extend pupils' thinking, in which pupils' answers are probed, built upon and elaborated, and which encourage pupils to ask questions and to interact with peers. Many teachers were not familiar with this distinction and those who were often claimed to be using a greater degree of interactive teaching in their classrooms than is borne out by lesson observation data (Bibby *et al.,* 2003; Hardman *et al.,* 2003; Pinel, 2002).

Looking at these last two issues together (the coverage of each topic within the timeframe available for it and the use of interactive whole class teaching), the evidence of strict adherence to some aspects of the policy on the one hand and

weak implementation of other aspects of the policy on the other hand further reinforces the earlier points about the difficulty in measuring impact. What can be said is that the fragmentation produced by the objectives-led approach may be negatively affecting the confidence and competence of lower attaining pupils, and that the current classroom practice evidenced here is not helping pupils to develop competence and confidence in higher order thinking since it is so rarely being encouraged.

The use of whole class teaching adopted in the mental/oral starter phase of the lesson was regarded by many teachers as stimulating pupils and getting the lesson off with a 'buzz', and such activities were generally regarded by pupils as fun and enjoyable. Nevertheless, there was some concern that this phase of the lesson publicly exposes low attaining pupils and can generate levels of anxiety amongst some pupils that can undermine their confidence; and there is some evidence that boys in Years 1 and 2 may more vulnerable to such exposure than girls (Myhill, 2002). Here the problems surface again of treating pupils as an undifferentiated group in claims about impact. What we have from the observational data here is further evidence of differential impact, particularly on the confidence of low attaining pupils, with a gender factor coming into play in this particular study.

In addition, a concern has also been expressed that the style of whole class teaching typically observed in fact generates bad habits regarding the way pupils are expected to perform. A substantial element of the three part lesson involves the public answering of questions which stresses speed and correctness, such that pupils are participating in the activities rather than engaging in the mathematical thinking (Denvir and Askew, 2001). This can undermine the development of a more reflective approach and the ability of pupils to think strategically. Strategic thinking refers to developing a repertoire of mental and written calculation strategies and informed decision making about the use of these. When the NNS was introduced, the development of strategic thinking was flagged up as a major change from previous teaching, but there is little evidence that pupils are better able to think strategically (Bibby *et al.,* 2003). This links with the earlier point about higher order thinking. Pupils may appear to be displaying confidence and competence, but this may be more about being able to play by the classroom rules and expectations of performance, rather than genuine engagement with the mathematics.

Hard evidence of greater confidence and competence of pupils in Years 1 and 2 is equivocal. Test data on pupils' performance in numeracy tests collected over Year 1 and/or Year 2 and compared with comparative data, such as the performance of pupils in other countries, or with the performance of Year 1 and 2 pupils just prior to the introduction of the NNS in September 1999 with that immediately after, do not provide clear evidence that performance has improved. One problem is that, in the period immediately before the introduction of the NNS, test data indicate that performance in mathematics had improved, in part because teachers were addressing more of their teaching towards meeting the pupil performance targets that were being set, and in part because the content being covered was matching more closely the content of what was being assessed and the amount of time being devoted to this content was also increasing (Brown *et al.,* 2001, 2003b). As such, by the time the NNS was introduced in September 1999, much of the gain that the NNS was expected to have promoted had already been achieved. Nevertheless, there is some evidence that pupils are now more confident and competent in their mental mathematics, and in numerical tasks

(tasks requiring the counting and manipulation of number), perhaps at the expense of relational tasks (tasks requiring the understanding of relations in space, size, quantity and order) (Aubrey *et al.,* 2003). In terms of the review question, which relates to the impact of the NNS specifically, we have to be very guarded. Conclusions about competence gains may only refer to some aspects of mathematical understanding; the NNS curriculum and earlier curricular reform may be the reasons for these gains rather than the NNS pedagogy. In this respect, we are cautious about what would have been the case if the NNS had not been implemented (see section 1.6), although it seems plausible that gains made before September 1999 would have been consolidated.

Some of the evidence for impact on pupil confidence points to the role played by teachers' expertise, particularly in the area of language. The in-depth analysis indicated that whole class teaching was placing greater demands on the teachers' use of mathematical vocabulary and that some teachers were not taking enough care to explain the meaning of the terms they were introducing and this is likely to undermine the development of pupil confidence, particularly amongst low attaining pupils (Raiker, 2002). In addition, teachers need to make use of the type of language used by pupils to indicate the extent to which this reflects their mathematical understanding. Such language can be elicited by the use of appropriate teacher questions for this purpose, which can usefully indicate whether pupils have internalised and are trying to follow a rule in order to generate the answer to a problem that has been set (Bills, 2003).

In addition, whole class teaching also makes great demands on teachers' mathematical subject matter knowledge and, where this is insecure, the quality of teaching and learning that occurs can be less effective (Earl *et al.,* 2002; Huckstep *et al.,* 2002). This is particularly evident in the extent to which teachers are able to use their subject matter knowledge in connecting, responding and exemplifying in order to aid the development of pupil confidence and competence.

Another study also indicates that there was still a gap between the nature of how mathematics was being used at home in real-life contexts and how it was being used in school. While much of school mathematics is embedded in real-life contexts, it is still taught and used in a way that emphasises this context as a means of developing and sustaining the use and understanding of mathematical operations (e.g. subtraction) rather than how the mathematics was used in the real-life context to solve meaningful and purposeful problems. This difference inhibits transfer of learning, so that pupils often do not see the connection between what they are doing in a mathematical operation and how that relates to what happens in the real-life context (Baker and Street, 2003). In this case study, there was convincing evidence that the case study child, for whom there was a big difference in home and school numeracy practices, was disengaged and lacking in confidence in the school classroom. This may be associated with the NNS curriculum and classroom practice which stresses conceptual understanding and number skills, rather than problem solving and application. This is not a problem confined to the NNS, however, but emerging research on the social practices of home numeracies is still very recent and did not inform the NNS policy despite the stated good intentions of involving parents and communities.

The way the three part lesson is organised in terms of an increased use of ability groupings (both in terms of setting across classes and in terms of within-class ability groups) has been noted in a number of studies in this review. Many teachers have attributed this increase to the introduction of the three part lesson

and the use of more whole class teaching. Some concern has been expressed that this will result in some low attaining pupils spending several years in a low set or ability group, which is likely to lead to disaffection with mathematics (McSherry and Ollerton, 2002). This inference is of course not based on primary evidence and thus needs to be treated with some caution. A concern has also been expressed that the style and approach of teaching adopted may have a different effect on different levels of ability; for example, there is evidence that a more relaxed and supportive approach to teaching helps low attaining pupils to recover in terms of confidence and competence, but this may be at the expense of the rest of the class which will not be challenged enough by such an approach, leading to less progress being made by the class as a whole (Brown *et al.,* 2003b). Organising the three part lessons with mixed age classes, which commonly occurs in small rural schools, poses a number of problems for teachers, who have largely had to adapt the format, by either shortening or omitting plenaries, or by making use of classroom assistants to run lessons in parallel for the different age or ability groups within the class (Evans, 2001). In this situation, able older pupils' needs may not be met but able younger pupils may be stimulated. In all these studies, a differential impact in terms of ability and age of pupils in mixed age classes can be inferred and the observational evidence strengthens the case for the concern already raised about the confidence of low attainers. The latter two studies (Brown *et al.,* 2003b; Evans, 2001) are the only studies in our review which also begin to raise concerns for pupils at the other (high) end of the ability continuum. These do, however, suggest the possibility that pupils of average ability may be experiencing the most positive impact.

The three part lesson has also led to a number of classroom teaching initiatives, such as an evaluation of whether educational television programmes can usefully contribute to whole class teaching in stimulating dialogue and discussion, and there is some evidence that pupils do display an enthusiasm for the television programmes and are active watchers, but the evidence that they enhance gains in numeracy is equivocal (Hopkins and Pope, 2000)**.** The range of opportunities offered by the programmes and the scope of resulting numeracy gains points to the teacher as a potentially weak link between the strategy and its implementation. This links with some of the issues already raised about the salience of the teacher's use of language, subject matter knowledge and expertise in conducting interactive whole class teaching, all of which may have a potential effect on pupils' confidence and competence.

In conclusion, it is difficult to give a clear-cut answer to our review question. There is some evidence of positive impact, but it seems to depend particularly on the ability of pupils and on the teacher's expertise. There is strong evidence suggesting that the pedagogy associated with the NNS is disadvantaging some pupils in terms of confidence and competence, that some important aspects of mathematics are being neglected, and that weaknesses in implementation coupled with a lack of ownership cast doubt on the sustainability of the policy.

4.3 In-depth review: quality-assurance results

Data extraction and assessment of the weight of evidence brought by the study to address the review question was conducted by two people, working first independently and then comparing their decisions and coming to a consensus. Two papers were data extracted by a member of the EPPI-Centre, with whom

there was a broad measure of agreement. Any differences were discussed and resolved. Most discussion centred on questions dealing with the reliability and validity of the studies, the generalisability of findings and the weight of evidence assessments, particularly the rating of 'high' for weight of evidence component B (see weight of evidence results in section 4.1).

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

The membership of the Review Group includes a variety of user groups, although the data extraction was undertaken by academics and researchers. Other user-group involvement was largely through email and informal contacts at conferences, and through publicising the work of the Review Group through subject and professional associations, organisations and societies. In addition, papers based on this systematic review have been and will be presented at a variety of conferences. Digests of the key findings and implications for policy and practice will be drawn to the attention of different user groups. The initial stage of dissemination has largely been directed at academics, teacher educators, researchers and policy-makers, but it is intended to widen the dissemination through the use of websites and articles in magazines and newspapers. It is too early to comment on the likely impact that this review will have on policy and practice.

5. FINDINGS AND IMPLICATIONS

5.1 Summary of principal findings

This review set out to answer the question:

Has the Daily Mathematics Lesson, in the context of the National Numeracy Strategy for primary schools in England, helped pupils to develop confidence and competence in early mathematics?

5.1.1 Identification of studies

The review identified 20 reports of 18 studies which met the inclusion criteria for producing the systematic map.

5.1.2 Mapping of all included studies

The mapping of these studies indicates the following:

- Just under half the studies were identified in the BEI electronic database and a number of important papers were identified through handsearching and personal contacts.
- Just over half the studies are published.
- Half the studies were classified as involving an evaluation.

5.1.3 Nature of studies selected for in-depth review

No studies in the map were excluded from the in-depth review.

5.1.4 Synthesis of findings from studies in in-depth review

The synthesis of the findings of these papers highlighted a number of problematic issues concerning the impact of the Daily Mathematics Lesson in the context of the NNS on pupil confidence and competence in early mathematics. The main findings are as follows:

- The key features of the Daily Mathematics Lesson have been well received by teachers and widely implemented.
- There is some evidence that this has enhanced pupil confidence and competence in early mathematics.
- A closer examination of the situation as evidenced by the studies included in this systematic review have highlighted a number of problematic issues.

- The intention that whole class teaching needs to be 'interactive' and promote higher quality dialogue, discussion and strategic thinking, has not been realised. Indeed, there is some evidence to indicate that the increased use of 'traditional' whole class teaching with 'pace' is in fact undermining the development of a more reflective and strategic approach to thinking about mathematics, and may be creating problems for lower attaining pupils.

- There is evidence that the stricter time management involved may pose particular problems for lower attaining pupils.

- The overall enhanced gains in pupil competence may in large measure be a reflection of a closer match between what is being taught and what is being tested, rather than greater pupil gains in their understanding of mathematics.

5.2 Strengths and limitations of this systematic review

The main strengths of this review have been that the review process has followed a publicly visible procedure and has benefited from the collaboration involved between the Review Group, the EPPI-Centre and many other individuals who offered comment, help and advice. The close scrutiny of the procedures involved means that each stage of the review involved discussion and justification.

The main limitations of the review are that the constraints involved in terms of time, cost and access to relevant papers, inevitably means that decisions about the focus of the review question and the conduct of the review process have to be taken in the context of keeping the review manageable. Such decisions involve trade-offs. For example, the decision to focus on early mathematics defined in terms of the teaching of Year 1 and 2 pupils, meant that a useful coverage of issues could be incorporated into the view, but this was at the expense of exploring how the issues highlighted here might change when considering the teaching of older pupils in KS 2. Perhaps the main limitation of the review is that a number of conference papers have not been written in a polished and full-length form, so the material presented in the conference paper sometimes omits details that would normally be required if the paper were to be accepted for publication in a journal. As such, the fact that half of the papers considered in this review were unpublished means that some details concerning the studies that would have been helpful to the review were not readily available. This meant that during data extraction, a number of questions about the paper had to be coded as unclear or not stated in the paper.

In addition, none of the studies employed a research design which would be ideal for addressing the review question, and hence many findings need to be regarded as tentative.

5.3 Implications

5.3.1 Policy

Policy-makers need to consider the extent to which the apparent success of the NNS may in part be a reflection of greater teaching for the test, whether the approach of interactive whole class teaching with pace may be inculcating bad learning habits, and whether the needs of low attaining pupils are being well served by the NNS. There is also a need to consider how the national assessment of pupil progress in mathematics can occur without constraining time and pedagogy in ways that undermine the development of pupils' mathematical understanding.

5.3.2 Practice

There is a major need for in-service training for primary teachers to highlight the purpose and nature of 'interactive' teaching in fostering higher quality dialogue, discussion and strategic thinking in order to ensure that teachers (i) understand better the notion of 'interactive' in interactive whole class teaching and (ii) adopt the type of classroom practice that can effectively aid the development of pupils' understanding of the mathematics of the topics they are covering. In particular, there is a need for in-service training to strengthen teachers' subject matter knowledge of mathematics, so that in the classroom context they can take better advantage of opportunities to enhance pupils' understanding of the mathematics in which they are engaged.

5.3.3 Research

There is a need for researchers to make greater use of measures and indicators of pupil confidence in order to draw firmer conclusions about how the features of the Daily Mathematics Lesson impacts on pupils' feelings of self-efficacy for the learning of mathematics.

Large-scale studies could measure self-efficacy in relation to performance on specific test items as in the Assessment of Performance Unit's study into attitudes and gender differences in mathematics (Joffe and Foxman, 1988), although the use of tests might not be appropriate for young children. It would be better to conduct a similar investigation through interviews with individual or pairs of young children, possibly using computer software for the presentation of tasks and input of responses. Naturalistic observation studies could also be used to probe further pupils' behaviour in daily mathematics lessons, while pupil interviews in a naturalistic setting could probe their self-assessment of confidence.

We do, however, recognise that there are difficulties in measuring or identifying increased pupil confidence in mathematics and in establishing a relationship between an increase in pupil confidence and pupils' understanding in mathematics and performance in national tests.

The roll-out of the NNS into KS 2 and 3 also means that there is a need for a systematic review to be carried out to explore the impact of aspects of the NNS, including the use of the three part lesson format and the use of interactive whole

class teaching, on pupil confidence and competence as they progress beyond KS 1 to the end of KS 4.

6. REFERENCES

6.1 Studies included in map and synthesis

The reference list includes all 20 papers that reported details of the 18 studies included in the review. Papers that are linked are highlighted using [ ] at the end of the reference.

Aubrey C, Godfrey R, Kavkler M, Magajna L, Tancig S (2003) The development of early numeracy and its relationship to curriculum goals and pedagogical practices in the European context. Paper presented at the European Early Childhood Educational Research Conference, University of Strathclyde: 3–6 September.

Baker D, Street B (2003) Researching home and school numeracy practices in the early years. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September.

Basit TN (2003) Changing practice through policy: trainee teachers and the National Numeracy Strategy. *Research Papers in Education* **18**: 61–74.

Bibby T, Askew M, Hodgen J (2003) Strategic thinking and the National Numeracy Strategy: an oxymoron? Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September.

Bills C (2003) What can teachers learn from the language that children use? *Education 3-13* **31**: 66–71.

Brown M, Askew M, Rhodes V, Denvir H, Ranson E, Wiliam D (2001) Magic bullets or chimeras? Searching for factors characterising effective teachers and effective teaching in numeracy. Paper presented at the British Educational Research Association Annual Conference, University of Leeds: 13–15 September.

Brown M, Askew M, Rhodes V, Denvir H, Ranson E, Wiliam D (2003b) Characterising individual and cohort progression in learning numeracy: results from the Leverhulme five-year longitudinal study. Paper presented at the American Educational Research Association Annual Conference, Chicago: 21–25 April. Available from: http://www.kcl.ac.uk/depsta/education/papers/AERA2003.pdf

Denvir H, Askew M (2001) Pupils' participation in the classroom examined in relation to 'interactive whole class teaching'. In: Rowland T (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conference held at Manchester Metropolitan University, 3 March,* Volume 21(1). London: British Society for Research into Learning Mathematics, pages 25–30.

Earl L, Watson N, Torrance N (2002) Front row seats: what we've learned from the national literacy and numeracy strategies in England. *Journal of Educational Change* **3**: 35–53.

Evans R (2001) Implementing the National Numeracy Strategy in small rural schools. *Mathematics Education Review* **13**: 55–74.

Hardman F (2003) Interactive whole class teaching in the national literacy and numeracy strategies. Paper presented at the Learning Conference 2003: What Learning Means. University of London, Institute of Education, London, UK: July 15-18. [See also Hardman *et al.,* 2003.]

Hardman F, Smith F, Mroz M, Wall K (2003) Interactive whole class teaching in the national literacy and numeracy strategies. Paper presented at the British Educational Research Association Annual Conference. Heriot-Watt University, Edinburgh: 11-13 September. Available from: http://www.leeds.ac.uk/educol/documents/00003267.htm [see also Hardman, 2003.]

Hopkins C, Pope S (2000) Children's numeracy skills and the use of educational television programmes. In: Jaworski B (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conferences held at the University of Exeter, 26 February, and the University of Loughborough, 6 May,* Volume 20 (1 & 2). London: British Society for Research into Learning Mathematics, pages 11–17.

Huckstep P, Rowland T, Thwaites A (2002) Primary teachers' mathematics content knowledge: what does it look like in the classroom? Paper presented at the British Educational Research Association Annual Conference, University of Exeter: 12–14 September. Available from: http://www.leeds.ac.uk/educol/documents/00002534.htm

Jones DV (2002) A comparative analysis of national numeracy initiatives in England and Wales. Paper presented at the British Educational Research Association Annual Conference, University of Exeter: 12–14 September. [See also Jones, 2003.]

Jones DV (2003) Learning from recent numeracy initiatives in England and Wales. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [see also Jones, 2002.]

McSherry K, Ollerton M (2002) Grouping patterns in primary schools. *Mathematics in School* **31**: 2–7.

Myhill D (2002) Bad boys and good girls? Patterns of interaction and responses in whole class teaching. *British Educational Research Journal* **28**: 339–352.

Pinel AJ (2002) Catechetics observed in mathematics classrooms. In: Cockburn AD and Nardi E (eds) *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education: 21-26 July,* Volume 1. Norwich: University of East Anglia, page 312.

Raiker A (2002) Spoken language and mathematics. *Cambridge Journal of Education* **32**: 45–60.

6.2 Other references used in the text of the report

Anghileri J (ed.) (2001) *Principles and Practices in Arithmetic Teaching: Innovative Approaches for the Primary Classroom.* Buckingham: Open University Press.

Askew M (2002) The changing primary mathematics classroom - the challenge of the National Numeracy Strategy. In: Haggarty L (ed.) *Aspects of Teaching Secondary Mathematics: Perspectives on Practice.* London: Routledge Falmer, pages 3–17.

Askew M, Brown M (eds) (2001) *Teaching and Learning Primary Numeracy: Policy, Practice and Effectiveness.* Southwell: British Educational Research Association. Available from: http://www.bera.ac.uk/beradev2002/root/publications/pdfs/numeracyreview.pdf

Askew M, Brown M (2003) *How Do We Teach Children to be Numerate?* Southwell: British Educational Research Association. Available from: http://www.bera.ac.uk/publications/pdfs/520668_Num.pdf

Aubrey C, Godfrey R (2003) The development of children's early numeracy through Key Stage 1. *British Educational Research Journal* **29**: 821–840.

Black P (2004) Assessment and learning. Keynote speech presented at the day conference of the British Society for Research into Learning Mathematics, King's College London: February 28.

Brown M, Askew M, Baker D, Denvir H, Millett A (1998) Is the National Numeracy Strategy research-based? *British Journal of Educational Studies* **46**: 362–385.

Brown M, Askew M, Millett, A. and Rhodes, V (2003a) The key role of educational research in the development and evaluation of the National Numeracy Strategy, *British Educational Research Journal* **29**: 655–672.

Brown M, Millett A (2003) Has the National Numeracy Strategy raised standards? In: Thompson I (ed.) *Enhancing Primary Mathematics Teaching.* Maidenhead: Open University Press, pages 199–209.

Butlin J (2003) Implementation of the National Numeracy Strategy in Birmingham LEA. *Mathematics Education Review* **16**: 17–34.

Coben, D (ed.) (2003) *Adult Numeracy: Review of Research and Related Literature.* London: National Research and Development Centre for Adult Literacy and Numeracy.

Department for Education and Employment (DfEE) (1998) *The Implementation of the National Numeracy Strategy: The final report of the Numeracy Task Force.* London: DfEE.

Department for Education and Employment (DfEE) (1999) *The National Numeracy Strategy: Framework for Teaching Mathematics from Reception to Year 6.* London: DfEE.

Dorman JP, Adams JE, Ferguson JM (2003) A cross-national investigation of students' perceptions of mathematics classroom environments and academic

efficacy in secondary schools. *International Journal of Mathematics Teaching and Learning* [e-journal]. 15 April. Available from: http://www.ex.ac.uk/cimt/ijmtl/dormanj.pdf

Earl E, Watson N, Levin B, Leithwood K, Fullan M, Torrance N (2003) *Watching and Learning 3: Final report of the External Evaluation of England's National Literacy and Numeracy Strategies.* London: DfES.

EPPI-Centre (2002a) *Core Keywording Strategy: Data Collection for a Register of Educational Research. Version 0.9.7.* London: EPPI-Centre, Social Science Research Unit.

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

Gifford S (2003) How should we teach mathematics to 3-year and 4-year olds? Pedagogical principles and practice for the foundation stage. *Mathematics Teaching* **184**: 33–38.

Graham T, Rowlands S, Jennings S, English J (1999) Towards whole-class interactive teaching. *Teaching Mathematics and its Applications* **18**: 50–60.

Harries T, Spooner M (2000) *Mental Mathematics for the Numeracy Hour.* London: David Fulton.

Hughes M (1999) The National Numeracy Strategy: are we getting it right? *Psychology of Education Review* **23**: 3–7.

Joffe L, Foxman D (1988) *Attitudes and Gender Differences: Mathematics at Age 11 and 15.* Slough: NFER-Nelson.

Jones DV (2002) National numeracy initiatives in England and Wales: a comparative study of policy. *Curriculum Journal* **13**: 5–23.

Jones L, Smart T (1995) Confidence and mathematics: a gender issue? *Gender and Education* **7** 157–166.

Keys W, Harris S, Fernandes C (1996) *Third International Study, First National Report, Part One.* Slough: National Foundation for Educational Research.

Maclellan E (2001) What counts as numeracy. *Scottish Educational Review* **33**:157–168.

Maclellan E, Munn P, Quinn V (2003) *Thinking about Maths: A Review of Issues in Teaching Number from 5 to 14 years.* Dundee: Teaching and Learning Scotland.

Macrae S (2003) The National Numeracy Strategy in primary schools. In: Coben D (ed.) *Adult Numeracy: Review of Research and Related Literature*. London: National Research and Development Centre for Adult Literacy and Numeracy, pages 85–90.

Moyles J, Merry R, Paterson F, Esarte-Sarries V (2003) *Interactive Teaching in the Primary School: Digging Deeper into Meanings.* Maidenhead: Open University Press.

Ofsted (2002) *The National Numeracy Strategy: The First Three Years 1999–2002.* London: Ofsted.

Ofsted (2003) *The National Literacy and Numeracy Strategies and the Primary Curriculum.* London: Ofsted.

Prais SJ, Luxton R (1998) Are the proposed reforms of numeracy teaching sufficient for success? *Teaching Mathematics and its Applications* **17**: 145-151.

Reynolds D (1998) *Numeracy Matters: The Preliminary Report of the Numeracy Task Force.* London: DfEE.

Smith A (2004) *Making Mathematics Count: The Report of Professor Adrian Smith's Inquiry into Post-14 Mathematics Education*. London: Stationery Office.

Straker A (1999) The National Numeracy Strategy: 1996–1999. In: Thompson I (ed.) *Issues in Teaching Numeracy in Primary Schools.* Buckingham: Open University Press, pages 39–48.

Tanner H, Jones S (2000) *Becoming a Successful Teacher of Mathematics*. London: RoutledgeFalmer.

Thomas J, Harden A (2003) Practical systems for systematic reviews of research to inform policy and practice in education. In: Anderson L, Bennett N (eds) *Developing Educational Leadership: Using Evidence for Policy and Practice*. London: Sage, pages 39–54.

Thompson I (2000) Is the National Numeracy Strategy evidence-based? *Mathematics Teaching.* **171**: 23–27.

Torgerson C (2003) *Systematic Reviews.* London: Continuum.

Wright RJ, Martland J, Stafford AK, Stanger G (2002) *Teaching Number: Advancing Children's Skills and Strategies.* London: Paul Chapman**.**

APPENDIX 1.1: Advisory Group membership

The membership of the Advisory Group is the same as the Review Group. However, other individuals (teachers, researchers, policy-makers) with an interest in the review question were invited to comment of the work of the Review Group at appropriate times. This was done through email and informal conversations at conferences. We are particularly indebted here to comments and advice we received from Carol Aubrey, Mike Askew, Margaret Brown, Jeremy Hodgen, Jenny Houssart, Ian Thompson and Julian Williams.

APPENDIX 2.1: Inclusion and exclusion criteria

For a paper to be included in the systematic map, it had to satisfy the following five criteria:

(i) It is an academic paper published in English in an academic journal or presented at an academic conference during the period 1999 to 2003.

(ii) It reports a study presenting original data collected by the author(s).

(iii) The study deals with a mainstream class in England.

(iv) The study deals with KS 1 pupils.

(v) The study deals with the Daily Mathematics Lesson in terms of the teaching and learning activities which occur in the lessons and/or the effects of the lessons on pupils' confidence and/or competence in early mathematics.

These inclusion criteria were reformulated as six exclusion criteria and placed in the hierarchical order, as indicated below, for ease of exclusion and, importantly, to act as a system of gradual filtering, so that the papers that were excluded at each stage could be readily identified in the future as a useful list of references that could be drawn upon for other purposes by readers of the review report, or might indeed be of use in subsequent systematic reviews undertaken by this Review Group. For example, criterion 2 below allows for studies dealing solely with KS 2 to be included initially, while such papers were excluded later when they were considered in terms of criterion 6. Such papers, however, can now be identified and used in a subsequent analysis where KS 2 appears within the focus of the review.

Exclusion codes: criteria for excluding a paper

EXC1: Not an academic paper published in an academic journal or presented at an academic conference during the period 1999 to 2003 in English

Examples of exclusion: a paper which is a descriptive outline of an approach; or a descriptive summary of a study which is intended for a practitioner audience; or a brief descriptive introduction of papers comprising a symposium

EXC2: Not about the classroom-based teaching and learning of primary school mathematics (years 1 to 6) in England

Examples of exclusion: a report of a study based on data collected in another country; or a study which only has tangential or contextual relevance to the teaching and learning of mathematics in the classroom; or a study which looks at pre-school numeracy; or a study looking at the numeracy skills tests for primary school student teachers; or a study looking at organisational aspects of implementing the NNS; or a study analysing the curriculum content of the mathematics programme of study; or a study looking at the design and implementation of national tests

EXC3: Not a report of a research study presenting original data collected by the author(s)

Examples of exclusion: a review of the literature; or a paper which offers a critique of policy and practice

EXC4: Does not mention the Daily Mathematics Lesson (or equivalent, e.g. numeracy hour, three part lesson, mental starter, whole-class interactive teaching, plenary)

Examples of exclusion: a paper dealing with the development of pupils' understanding of a specific topic, such as place value or division, without contextualising or making links in any way with how the Daily Mathematics Lesson approach has or may impact on this; or a paper dealing with pupils' use of language without contextualising or making links in any way with how the Daily Mathematics Lesson approach has or may impact on this

EXC5: Not about mainstream pupils in mainstream classes

Examples of exclusion: a study which looks at booster classes, numeracy recovery sessions, or one-to-one individual teaching for a statemented pupil

EXC6: Does not deal with the impact of the Daily Mathematics Lessons on *either* teaching and learning activities during the Daily Mathematics Lessons *or* the development of pupil confidence and/or competence in early mathematics during the Daily Mathematics Lessons with reference to one or more of the following types of data collected since September 1999:

(i) the perceptions of KS 1 (i.e. Year 1 and 2) pupils

(ii) the perceptions of teachers with reference to teaching KS 1 (i.e. Year 1 and 2) pupils

(iii) direct classroom observation of KS 1 (i.e. Year 1 and 2) pupils

(iv) assessment tasks undertaken by KS 1 (i.e. Year 1 and 2) pupils during or at the end of KS 1

Examples of exclusion: interviews with Year 4 pupils, lesson observations of Year 6 pupils; attainment tests taken at the end of KS 2

APPENDIX 2.2: Search strategy for electronic databases

The British Education Index (BEI) and the Educational Resources Information Center (ERIC) were searched on 6th January and 9th January 2004 respectively. The databases were accessed via the Dialog interface and the searches constructed as follows:

Database: ERIC – CIJE & RIE 1990 – September 2003

1. MATHEMATICS AND PRIMARY
2. NUMERACY
3. DAILY MATHEMATICS LESSON
4. 1 OR 2 OR 3
5. Limit 4 to Language=("ENGLISH")
6. Limit 5 to: Publication Year=("1999" OR "2000" OR "2001" OR "2002" OR "2003") and Document Type=("COLLECTED WORKS--PROCEEDINGS (021)" OR "JOURNAL ARTICLE")

Database: British Education Index 1976 – September 2003

1. MATHEMATICS AND PRIMARY
2. NUMERACY
3. DAILY MATHEMATICS LESSON
4. 1 OR 2 OR 3
5. Limit 4 to: Publication Year=("1999" OR "2000" OR "2001" OR "2002" OR "2003")
6. Limit 5 to: Non-theses (BEI) Records Only

APPENDIX 2.3: Journals handsearched

(i) Electronic search and/or handsearch of nine key journals in Mathematics Education (1999–2003) looking at every tilte and where appropriate and available the abstract and/or the full-paper (880 papers):

Educational Studies in Mathematics
For the Learning of Mathematics
Journal of Mathematics Teacher Education
Journal for Research in Mathematics Education
International Journal of Mathematics Teaching and Learning
Mathematics Education Review
Mathematics in Schools
Mathematics Teaching
Teaching Mathematics and its Applications

(ii) Electronic searches and/or handsearching issues of the following 17 selected key UK journals in Educational Research (1999–2003) looking at every title and where appropriate and available the abstract and/or the full paper (2,007 papers):

British Educational Research Journal
British Journal of Educational Psychology
British Journal of Educational Studies
Cambridge Journal of Education
Curriculum Journal
Education 3 to 13
Educational Psychology
Educational Research
Educational Review
Educational Studies
Evaluation and Research in Education
Journal of Education Policy
Oxford Review of Education
Research in Education
Research Papers in Education
Scottish Educational Review
Welsh Journal of Education

(iii) Handsearch and/or electronic search of key recent conference proceedings looking at every title and where appropriate and available the abstract and/or the full paper (3,043 papers):

- British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: September 11–13, 2003
- British Society for Research into Learning Mathematics, Day Conferences, 1999–2003
- European Conference for Research on Learning and Instruction, University of Padua: August 26–30, 2003
- European Conference on Educational Research, University of Hamburg: September 17–20, 2003

- International Group for the Psychology of Mathematics Education Annual Conference, University of East Anglia: July 21-26, 2002.
- Scottish Educational Research Association Annual Conference, Perth: November 27–29, 2003

APPENDIX 2.4: EPPI-Centre Keyword sheet

A1. Identification of report
Citation
Contact
Handsearch
Unknown
Electronic database
(Please specify.) .....

A2. Status
Published
In press
Unpublished

A3. 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.)
.....
.....
.....
.....

A4. Language (Please specify.)
.....

A5. In which country/countries was the study carried out? (Please specify.)
.....
.....
.....

A6. 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.).....

A7. Curriculum
Art
Business studies
Citizenship
Cross-curricular
Design and technology
Environment
General
Geography
Hidden
History
ICT
Literacy – first language
Literacy further languages
Literature
Maths
Music
PSE
Physical education
Religious education
Science
Vocational
Other (Please specify.).....

A8. Programme name (Please specify.)
.....

A9. 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.).....

A10. Age of learners (years)
0-4
5-10
11-16
17-20
21 and over

A11. Sex of learners
Female only
Male only
Mixed sex

A12. 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.).....

A13. Which type(s) of study does this report describe?
A. Description
B. Exploration of relationships
C. Evaluation
   a. naturally-occurring
   b. researcher-manipulated
D. Development of methodology
E. Review
   a. Systematic review
   b. Other review

APPENDIX 3.1: Details of studies included in the systematic map

Table 3.1 Identification of report (18 studies)

Electronic search strategy	8
Handsearching journals	2
Handsearching conferences	4
Citation	1
Personal contact	3

Table 3.2 Status (18 studies)

Published	10
Unpublished	8

Table 3.3 Sample focus (18 studies)

Exclusive focus on pupils	9
One or more other foci instead of or in addition to pupils	9

Table 3.4 Type of study (18 studies)

Description	2
Exploration of relationships	7
Evaluation (naturally occurring)	8
Evaluation (researcher-manipulated)	1

APPENDIX 3.2: Possible inclusions identified by electronic searching strategy (this refers to using specified keywords with BEI and ERIC) (33 papers)

Anderson JA, Boylan M (2000) The National Numeracy Strategy: teacher questions and pupil anxiety. In: Jaworski B (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conferences held at the University of Exeter, 26 February and University of Loughborough, 6 May*, Volume 20 (1 & 2). London: British Society for Research into Learning Mathematics, pages 49–54. [excluded]

Aubrey C, Tancig S, Magajna L, Kavkler M (2000) The development of numeracy in England and Slovenia. *Education 3–13* **28**: 13–23. [excluded]

Basit TN (2003) Changing practice through policy: trainee teachers and the National Numeracy Strategy. *Research Papers in Education* **18**: 61–74. [included]

Bills C (2001) Indicators of abstraction in young children's description of mental calculations. In: Winter J (ed.) *Proceedings of the Fifth British Congress of Mathematics Education held at Keele University, 5–7 July*, Volume 21 (2). London: British Society for Research into Learning Mathematics, pages 75–85. [excluded]

Bills C (2002) The use of pronouns, causal connectives and tense in young children's descriptions of their mental mathematics. In: Winter J and Pope S (eds) *Research in Mathematics Education: Papers of the British Society for Research into Learning Mathematics,* Volume 4. London: British Society for Research into Learning Mathematics, pages 3–20. [excluded]

Bills C (2003) What can teachers learn from the language that children use? *Education 3–13* **31:** 66–71. [included]

Black L (2002) 'She's not in my head or in my body': constructing pupil identities of exclusion and full participation in classroom learning processes. Paper presented at the British Educational Research Association Annual Conference, University of Exeter: 12–14 September. [excluded]

Bold C (2000) Researching resources for teaching and learning: the counting stick. In: Jaworski B (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conferences held at the University of Exeter, 26 February, and the University of Loughborough, 6 May*, Volume 20 (1 & 2). London: British Society for Research into Learning Mathematics, pages 61–66. [excluded]

Borthwick A (2001) Capturing or missing moments: an analysis of how teachers respond to mathematical 'moments' in the classroom. In: Winter J (ed.) *Proceedings of the Fifth British Congress of Mathematics Education held at Keele University, 5-7 July,* Volume 21 (2). London: British Society for Research into Learning Mathematics, pages 86–96. [excluded]

Bruce B, Martin J (2002) The teaching of numeracy: what constitutes an effective approach? *Education 3–13,* **30**: 9–13. [excluded]

Campbell RJ, Kyriakides L (2000) The National Curriculum and standards in primary schools: a comparative perspective. *Comparative Education,* **36**: 383-395. [excluded]

Davison I (2003) Using interactive whiteboard to facilitate pupil understanding of quadrilateral definitions. In: Pope S (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conference held at Sheffield Hallam University: 22 February*, Volume 23 (1). London: British Society for Research into Learning Mathematics, pages 13–18. [excluded]

Denvir H, Askew M (2001) Pupils' participation in the classroom examined in relation to 'interactive whole class teaching'. In: Rowland T (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conference held at Manchester Metropolitan University, 3 March,* Volume 21 (1). London: British Society for Research into Learning Mathematics, pages 25–30. [included]

Evans J (2002) Talking about maths, *Education 3–13* **30**: 66–71. [excluded]

Felgate R, Minnis M, Schagen I (2000) Some results from the analysis of data from the National Numeracy Project. *Research Papers in Education* **15**: 163–184. [excluded]

Hardman F, Smith F, Mroz M, Wall K (2003) Interactive whole class teaching in the national literacy and numeracy strategies. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. Available from: http://www.leeds.ac.uk/educol/documents/00003267.htm [included]

Hopkins C, Pope S (2000) Children's numeracy skills and the use of educational television programmes. In: Jaworski B (ed.) *Proceedings of the British Society for research into learning mathematics: proceedings of the day conferences held at the University of Exeter, 26 February, and the University of Loughborough, 6 May*, Volume 20 (1 & 2). London: British Society for Research into Learning Mathematics, pages 11–17. [included]

Houssart J (2000) The role of number resources in the daily mathematics lesson. In: Rowland T (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conferences held at the University of Surrey Roehampton, 17-18 November*, Volume 20 (3). London: British Society for Research into Learning Mathematics, pages 37–42. [excluded]

Houssart J (2000) Nice and easy does it: pupil responses to non-challenging tasks. In: Jaworkski B (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conferences held at University of Exeter, 26 February and University of Loughborough, 6 May*, Volume 20 (1 & 2). London: British Society for Research into Learning Mathematics, pages 91–96. [excluded]

Houssart J (2002) Count me out: task refusal in primary mathematics. *Support for Learning* **17:** 75–79. [excluded]

Howat H (2002) Improving children's flexibility in elementary arithmetic. In: Goodchild S (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conferences held at the University of Durham, 23 February and the University of Bristol, 18 May*, Volume 22 (1 & 2), London: British Society for Research into Learning Mathematics, pages 25–30. [excluded]

Huckstep P, Rowland T, Thwaites A (2002) Primary teachers' mathematics content knowledge: what does it look like in the classroom? Paper presented at the British Educational Research Association Annual Conference, University of Exeter: 12–14 September. [included]

MacIntyre H, Ireson J (2002) Within-class ability grouping: placement of pupils in groups and self-concept. *British Educational Research Journal* **28:** 249–263. [excluded]

McSherry K, Ollerton M (2002) Grouping patterns in primary schools. *Mathematics in School* **31:** 2–7. [included]

Muijs D, Reynolds D (2000) School effectiveness and teacher effectiveness in mathematics: some preliminary findings from the evaluation of the Mathematics Enhancement Programme (primary). *School Effectiveness and School Improvement* **11**: 273–303. [excluded]

Muijs D, Reynolds D (2003) Student background and teacher effects on achievement and attainment in mathematics: a longitudinal study. *Educational Research and Evaluation* **9**: 289–314. [excluded]

Patilla P (2002) Interactive and participatory mathematics in the primary classroom. *Education Review* **15**: 66–71. [excluded]

Raiker A (2002) Spoken language and mathematics. *Cambridge Journal of Education* **32**: 45–60. [included]

Ryan J, Williams J (2002) Charting argumentation space in conceptual locales: tools at the boundary between research and practice. In: Winter J and Pope S (eds) *Research in Mathematics Education: Papers of the British Society for Research into Learning Mathematics*, Volume 4. London: British Society for Research into Learning Mathematics, pages 89–111. [excluded]

Sutherland R, Godwin S, Olivero F, Peel P (2002) Design initiatives for learning: ICT and geometry in the primary school. In: Goodchild S (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conferences held at the University of Durham, 23 February and the University of Bristol, 18 May,* Volume 22 (1 & 2). London: British Society for Research into Learning Mathematics, pages 73–78. [excluded]

Whitburn J (2001) Effective classroom organisation in primary schools: mathematics. *Oxford Review of Education* **27**: 411–428. [excluded]

Williams J, Ryan J (2000) National testing and the improvement of classroom teaching: can they coexist? *British Educational Research Journal* **26**: 49–73. [excluded]

Wilson L, Andrew C, Sourikova S (2001) Shape and structure in primary mathematics lessons: a comparative study in the North-east of England, and St Petersburg, Russia – some implications for the daily mathematics lesson. *British Educational Research Journal* **27**: 29–58. [excluded]

APPENDIX 3.3: Possible inclusions identified by handsearching (this refers to handsearching and electronic searching of key journals and conferences proceedings, citations and personal contacts) (32 papers)

Adhami M, Johnson DC, Shayer M (1999) Tensions in planned mathematics lessons. In: Bills L (ed.) *Proceedings of the British Society for Research into Learning Mathematics: Proceedings of the Day Conferences held at University of Warwick, 12–13 November*. London: British Society for Research into Learning Mathematics, pages 35–41. [excluded]

Askew M, Brown M (2003) The impact of the National Numeracy Strategy on mathematics attainment and learning in year 4. Paper presented as part of a symposium on changing learning in mathematics at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [excluded]

Askew M, Brown M (2003) Pupils developing social identities as learners of primary mathematics: the interplay of the situational and the personal. Paper presented as part of a symposium on changing learning in mathematics at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [full paper not available]

Aubrey C, Godfrey R, Kavkler M, Magajna L, Tancig S (2003) The development of early numeracy and its relationship to curriculum goals and pedagogical practices in the European context. Paper presented at the European Early Childhood Educational Research Conference, University of Strathclyde: 3–6 September. [included]

Baker D, Street B (2003) Researching home and school numeracy practices in the early years. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [included]

Bibby T, Askew M, Hodgen J (2003) Strategic thinking and the national numeracy strategy: an oxymoron? Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [included]

Borthwick A (2002) What do pupils think? Pupils' perceptions of their mathematics lessons. In: Cockburn AD, Nardi E (eds) *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education,* Volume 2. Norwich: University of East Anglia, pages 145–152. [excluded]

Brown M (2002) Researching primary numeracy. In: Cockburn AD, Nardi E (eds) *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education,* Volume 1. Norwich: University of East Anglia, pages 15–30. [excluded]

Brown M (2003) Successes and failures of an imposed systemic reform: the National Numeracy Strategy in England. Paper presented at the 10th European Conference for Research on Learning and Instruction, University of Padua, Italy: 26–30 August. [excluded]

Brown M, Askew M (2003) Changing teaching in year 4: the impact of the National Numeracy Strategy. Paper presented as part of a symposium on changing teaching in mathematics at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [excluded]

Brown M, Askew M, Hodgen J (2003) Characterising individual and cohort progression in learning numeracy. Paper presented as part of a symposium on changing learning in primary mathematics at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [excluded]

Brown M, Askew M, Hodgen J (2003) The relative contribution of different factors in the learning of mathematics in primary schools. Paper presented as part of a symposium on changing learning in primary mathematics at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [full paper not available]

Brown M, Askew M, Millett A, Rhodes V (2003) The key role of educational research in the development and evaluation of the National Numeracy Strategy. *British Educational Research Journal* **29**: 655–672. [excluded]

Brown M, Askew M, Rhodes V, Denvir H, Ranson E, Wiliam D (2001) Magic bullets or chimeras? Searching for factors characterising effective teachers and effective teaching in numeracy. Paper presented at the British Educational Research Association Annual Conference, University of Leeds: 13–15 September. [included]

Brown M, Askew M, Rhodes V, Denvir H, Ranson E, Wiliam, D (2003) Characterising individual and cohort progression in learning numeracy: results from the Leverhulme five-year longitudinal study. Paper presented at the American Educational Research Association Annual Conference, Chicago, 21–25 April. Available from: http://www.kcl.ac.uk/depsta/education/papers/AERA2003.pdf [included]

Butlin J (2003) Implementation of the National Numeracy Strategy in Birmingham LEA. *Mathematics Education Review* **16**:17–34. [excluded]

Donaldson G (2000) Evaluation of the National Numeracy Strategy. *Primary Mathematics* **4**: 20–22. [excluded]

Earl L, Watson N, Torrance N (2002) Front row seats: what we've learned from the National Literacy and Numeracy Strategies in England. *Journal of Educational Change* **3**: 35–53. [included]

Evans R (2001) Implementing the National Numeracy Strategy in small rural schools. *Mathematics Education Review* **13**: 55–74. [included]

Gray E, Howat H, Pitta-Pantazi D (2003) Have all the children benefited? A perspective on curriculum initiatives and low achievers. In: Cockburn AD, Nardi E

(eds) *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education,* Volume 3. Norwich: University of East Anglia, pages 113–120. [excluded]

Hardman F (2003) Interactive whole class teaching in the national literacy and numeracy strategies. Paper presented at the Learning Conference 2003: What learning means. University of London, Institute of Education: 15–18 July. [included]

Haylock D (2002) Primary student-teachers teaching numeracy. *Mathematics Education Review* **15**: 21–31. [excluded]

Houssart J (2001) Counting difficulties at Key Stage 2. *Support for Learning* **16**: 11–16. [excluded]

Jones DV (2002) A comparative analysis of national numeracy initiatives in England and Wales. Paper presented at the British Educational Research Association Annual Conference, University of Exeter: 12–14 September. [included]

Jones DV (2003) Learning from recent numeracy initiatives in England and Wales. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 1–13 September. [included]

McCallum B, Gipps C, Hargreaves E (2001) The structure of lessons. *Education 3–13* **29**: 33–37. [excluded]

Murphy C (2003) A study of the formalised teaching of a mental calculation strategy and children's construction of meaning. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. [excluded]

Murphy C (2003) A theoretical comparison of the teaching of mental calculation strategies in England and the Netherlands. In: Pope S, McNamara O (eds) *Research in Mathematics Education: papers of the British Society for Research into Learning Mathematics*, Volume 5, London: British Society for Research into Learning Mathematics**,** pages 123–137. [excluded]

Myhill D (2002) Bad boys and good girls?: patterns of interaction and responses in whole class teaching. *British Educational Research Journal* **28**: 339–352. [included]

Ogden L (2000) Collaborative tasks, collaborative children: an analysis of reciprocity during peer interactions at Key Stage 1. *British Educational Research Journal* **26**: 211–226. [excluded]

Pinel AJ (2002) Catechetics observed in mathematics classrooms. In: Cockburn AD, Nardi E (eds) *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education,* Volume 1. Norwich: University of East Anglia, page 312. [included]

Sahin C, Bullock K, Stables A (2002) Teachers' beliefs and practices in relation to their beliefs about questioning at Key Stage 2. *Educational Studies* **28**: 371–384. [excluded]

APPENDIX 4.1: Details of studies included in the in-depth review

(a) Weight of evidence

Paper	Component A: Trustworthiness of the study in answering the study's questions	Component B: Appropriateness of design and analysis for the review question	Component C: Relevance of the focus of the study for the review question	Composite D: Overall weight taking account of A, B and C
Aubrey *et al.*, 2003	High	High	High	High
Baker and Street, 2003	High	High	High	High
Basit, 2003	High	High	High	High
Bibby *et al.*, 2003	High	High	High	High
Bills, 2003	High	High	High	High
Brown *et al.,* 2001	High	High	High	High
Brown *et al.*, 2003b	High	Medium	Medium	Medium
Denvir and Askew, 2001	Medium	Medium	High	Medium
Earl *et al.*, 2002	High	High	High	High
Evans, 2001	High	High	High	High
Hardman *et al.*, 2003	High	Medium	Medium	Medium
Hopkins and Pope, 2000	High	Medium	Medium	Medium
Huckstep *et al.*, 2002	High	Medium	Medium	Medium
Jones, 2003	High	Medium	High	High
McSherry and Ollerton, 2002	High	High	High	High
Myhill, 2002	High	Medium	Medium	Medium
Pinel, 2002	Medium	Medium	High	Medium
Raiker, 2002	High	High	High	High

(b) Scope of study

Paper	Addresses classroom practice	Addresses pupil confidence	Addresses pupil competence
Aubrey *et al.*, 2003	No	No	Explicit
Baker and Street, 2003	Explicit	Implicit	Explicit
Basit, 2003	Explicit	Explicit	Explicit
Bibby *et al.*, 2003	Explicit	Implicit	Explicit
Bills, 2003	Explicit	Explicit	Explicit
Brown *et al.,* 2001	Explicit	No	Explicit
Brown *et al.*, 2003b	Explicit	Explicit	Explicit
Denvir and Askew, 2001	Explicit	Explicit	Explicit
Earl *et al.*, 2002	Explicit	Implicit	Explicit
Evans, 2001	Explicit	Explicit	Explicit
Hardman *et al.*, 2003	Explicit	Implicit	Implicit
Hopkins and Pope, 2000	Explicit	Explicit	Explicit
Huckstep *et al.*, 2002	Explicit	Explicit	Explicit
Jones, 2003	Explicit	Explicit	Explicit
McSherry and Ollerton, 2002	Explicit	Implicit	Implicit
Myhill, 2002	Explicit	Implicit	Implicit
Pinel, 2002	Explicit	Implicit	Implicit
Raiker, 2002	Explicit	Explicit	Explicit

APPENDIX 4.2: Summary of included studies

(1) Aubrey C, Godfrey R, Kavkler M, Magajna L, Tancig S (2003) The development of early numeracy and its relationship to curriculum goals and pedagogical practices in the European context. Paper presented at the European Early Childhood Educational Research Conference, University of Strathclyde: 3–6 September.

Overall weight of evidence score = high

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Compares overall performance in numeracy tasks for English and Slovenian pupils	• Administered a numeracy tests to 300 5-6 year-olds in England and Slovenia	• Data do not address classroom teaching	• English pupils more successful in numerical tasks and less successful in relational tasks compared with Slovenian children*

* **Note:** One might speculate that the mental starter may be having a more effective influence on numerical tasks (tasks requiring the counting and manipulation of number) than on relational tasks (tasks requiring the understanding of relations in space, size, quantity and order) and/or that relatively more time is being devoted to numerical tasks.

(2) Baker D, Street B (2003) Researching home and school numeracy practices in the early years. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September.

Overall weight of evidence score = high

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Explores links between home and school numeracy practices	• Ethnographic case study of a primary school pupil based on home and school visits over a three-year period (from reception year to year 2)	• School numeracy practices are very different from home numeracy practices • The Year 1 teacher had control over choice of activities, knew the answers and knew how pupils were to do the tasks • School practices have a clear educational focus (e.g. to learn subtraction) while home practice is embedded in real-life purposes	• Differences in home and school practices can undermine transfer of learning from school to home and can undermine pupil confidence if the school tasks are not embedded in a purposeful and meaningful context which the pupil has experienced and can relate to.

(3) Basit TN (2003) Changing practice through policy: trainee teachers and the National Numeracy Strategy. *Research Papers in Education* **18**: 61–74.

Overall weight of evidence score = high

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Explores effects of the NNS on primary student-teachers	• Two in-depth interviews with 30 final year BEd students who were non-maths specialists	• NNS framework has boosted trainees' confidence • Some experienced teachers do not follow the three-part lesson to the letter • The three part lesson forces teachers to think about the lesson as a whole • NNS framework is easy to use but is sometimes over ambitious regarding what can be covered in the time allocated for each topic • On the one hand, restricted time for activities does not allow activity to be continued over several consecutive lessons, but on the other hand it helps with time management of topics that need to be covered	• Mental starter gets pupils involved and excited about maths • Restricted time for activities can disadvantage low attainers as you have to move on to another topic before they have adequately grasped the current one • NNS assumes pupils have reached a certain level at the end of the year and have to be taken on from there at the beginning of the next year when in fact many pupils have considerable gaps in their understanding

(4) Bibby T, Askew M, Hodgen J (2003) Strategic thinking and the national numeracy strategy: an oxymoron? Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh, 11–13 September.

Overall weight of evidence score = high

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Explores primary school teachers' interpretations and implementations of the teaching of mental calculations in the daily maths lesson	• 59 lessons observed (19 KS 1, 40 KS 2), each followed by brief discussion with the class teacher	• Three part lesson is being widely used, with a tendency towards rather long mental/oral starters with brief (or omitted) plenaries • Little evidence appears of strategic thinking (developing a repertoire of mental and written calculation strategies and informed decision making about the use of these) being encouraged • The overwhelming majority of the lesson involved recall (e.g. learning multiplication facts) or procedures (formal written procedures and teacher-led procedures with informal methods) • Very few instances exist of teachers responding to and building on pupils' answers and questions	• Little evidence is shown that pupils are better able to think strategically* • Pupils were primarily learning mental arithmetic (in the traditional sense of an emphasis on recall of answers and methods) rather than mental calculation in the spirit of the NNS

***note:** when the NNS was introduced, the development of strategic thinking was 'flagged up' as a major change from previous teaching.

(5) Bills C (2003) What can teachers learn from the language that children use? *Education 3–13* **31**: 66–71.

Overall weight of evidence score = high

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Explores the language pupils used when asked 'What is in your head?' after they had performed a mental calculation	• Lesson observations and interviews with Year 2 to Year 4 pupils, leading to an analysis of over 1,000 responses given by 26 pupils in six (group) interviews	• Teachers can take account of pupils' use of language to indicate pupils' level of confidence with taught rules in their classroom teaching	• Pupils' adoption of a classroom speech style involving the use of 'you' and the present tense is an indication that they have internalised the rules and follow them when making mental calculations • Pupils' use of explanatory expressions indicates a confidence in and familiarity with procedures • Both the above indicators correlated with accuracy in calculations • Evidence is provided here of how pupils' use of language can indicate their confidence in following taught rules, but no evidence is provided of the extent to which teachers are doing this and/or the impact that doing so has on pupil confidence and competence

(6) Brown M, Askew M, Rhodes V, Denvir H, Ranson E, Wiliam D (2001) Magic bullets or chimeras? Searching for factors characterising effective teachers and effective teaching in numeracy. Paper presented at the British Educational Research Association Annual Conference, University of Leeds, 13–15 September.

Overall weight of evidence score = high

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Explores whether a particular type of teacher or way of teaching produces higher learning gains in numeracy	• Leverhulme Numeracy Project is a five-year longitudinal study (1997-2002) involving five projects (one core project and five focus projects) designed to explore factors leading to low attainment in primary numeracy and testing out ways of raising attainment • The core project tracked two cohorts of pupils, one starting in reception and the other starting in Year 4 (c. 1,600 pupils in each) and monitored their performance on a sequence of numeracy tests; each school was visited once a year and there was a lesson observation of each class, interviews with teachers, headteachers and coordinators, and a teacher questionnaire • The five focus projects mainly took the form of qualitative accounts relating to a particular factor which affects the learning of primary numeracy projects:	• Data are presented for Year 2 pupils tested at the start and end of 1999/2000 (the first year of the NNS) and the correlation between Year 2 gain and Saxe lesson observation evaluation score, r = 0.18) (This correlation was only marginally higher than the one obtained in the previous year, 1998/1999, when the pupils were in Year 1, r = 0.08) • Data indicate that it may be very hard to characterise adequately effective teaching of numeracy and to construct a reliable lesson observation based on such a characterisation	• Pupil gain scores for Year 1 pupils at the end of 1998/1999 and for Year 2 at the end of 1999/2000 are presented, but no comparison is made here to indicate whether this provides any evidence that Year 2 pupils made a greater gain (in the first year of the NNS) than Year 1 pupils did (in the year prior to the introduction on the NNS); however, this is addressed in Brown *et al.* (2003)

	(i) classroom experiences (ii) teachers' conceptions and practices (iii) school leadership and action (iv) home and school numeracy practices (v) cognitive acceleration intervention		

(7) Brown M, Askew M, Rhodes V, Denvir H, Ranson E, Wiliam D (2003b) Characterising individual and cohort progression in learning numeracy: results from the Leverhulme five-year longitudinal study. Paper presented at the American Educational Research Association Annual conference, Chicago: 21–25 April. http://www.kcl.ac.uk/depsta/education/papers/AERA2003.pdf

Overall weight of evidence score = medium

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Explores pupil progress in numeracy and the links between test data to lesson observation data for a sample of case study pupils	• Leverhulme Numeracy Project is a five-year longitudinal study (1997–2002) involving five projects (one core project and five focus projects) • The core project tracked two cohorts of pupils, one starting in reception, and the other starting in Year 4 (c. 1,600 pupils in each) and monitored their performance on a sequence of numeracy tests; each school was visited once a year and there was a lesson observation of each class,	• Data from lesson observations indicate some variability between teachers and the effects this has on outcomes: for example, a teacher employing a more relaxed and supportive approach towards pupils' having problems helped lower attainers to recover in terms of their confidence and competence, but this was at the expense of the rest of the class who were not being challenged enough and the class as a whole made less progress overall during	• The study presents data on common items that appeared in the numeracy tests to graph their items facility scores at the beginning and end of each year for the first cohort (Years 1 to 4), and for the second cohort (Years 4 to 7) • The mean increase in facility for Year 1 pupils at the end of 1998/1999 and for Year 2 at the end of 1999/2000 is the same (20% increase in both cases) and this is higher than for older Year

	interviews with teachers, head teachers and coordinators, and a teacher questionnaire • The five focus projects mainly took the form of qualitative accounts relating to a particular factor which affects the learning of primary numeracy projects: (i) classroom experiences (ii) teachers' conceptions and practices (iii) school leadership and action (iv) home and school numeracy practices (v) cognitive acceleration intervention • 30 case-study pupils (six pupils from one class in each of five schools) were interviewed about their experiences of learning mathematics and were also asked to talk through a set of test items; their teachers were also interviewed about the case-study pupils	the year.	groups • The progress of the 30 case-study pupils display idiosyncratic trajectories, and the explanation of the impact of the teaching on their competence and competence reflect a complex mix of factors

(8) Denvir H, Askew M (2001) Pupils' participation in the classroom examined in relation to 'interactive whole class teaching'. In Rowland T (ed.) *Proceedings of the British Society for research into learning mathematics: proceedings of the day conference held at Manchester Metropolitan University, 3 March*, Volume 21(1), London: British Society for Research into Learning Mathematics, pages 25–30.

Overall weight of evidence score = medium

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Explores the ways in which pupils manage their participation during whole class interactive teaching	• 150 lesson observations	• Data indicate that the three part lesson includes a substantial element of public answering of questions • This involves a strong 'performative' element which stresses speed and correctness	• When pupils participate in whole class interactive teaching, they may not be participating in the mathematical thinking intended • The strong performative element prompts pupils to adopt behaviour which militates against them developing good habits as learners • They may be *participating* in the activities rather than *engaging* in the mathematical thinking

(9) Earl L, Watson N, Torrance N (2002) Front row seats: what we've learned from the national literacy and numeracy strategies in England, *Journal of Educational Change* **3:** 35–53.

Overall weight of evidence score = high

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Explores what can be learnt from the national evaluations of the NLS and the NNS	• In Year 2 of the evaluation surveyed 500 schools, visited 10 schools, conducted 200 interviews	• Data indicate teachers and headteachers are very positive about the NNS • The three part mathematics lesson provides more focus, is easier, more disciplined and progression is known • Teachers are motivated to improve their teaching, and feel their skills are now sharper and they have better techniques • Teachers are making use of NNS materials • There is the danger that rigid compliance with key features of the three part lesson and a lack of subject and pedagogical knowledge can lead to ineffective teaching	• Pupils' learning is enhanced because the need for differentiation in the middle phase of the three part lesson and the plenary session both help teachers to link new learning to what pupils already know

(10) Evans R (2001) Implementing the National Numeracy Strategy in small rural schools. *Mathematics Education Review* **13:** 55–74.

Overall weight of evidence score = high

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Explores how small rural schools are implementing the NNS, focusing on mixed age classes with three or more year groups (e.g. containing reception to Year 2 pupils)	• Interviewed 17 teachers about their strategies for organising and carrying out daily maths lessons	• A combination of positive and negative comments are reported about the NNS • The NNS framework sometimes does not provide enough time for pupils to grasp a topic adequately, but it does avoid spending far too long on a topic • Whole class teaching, particularly mental/oral starters and plenaries, is problematic • Able older pupils and least able younger pupils can get easily bored and find the pace is inappropriate for them • About half the teachers did not always ensure they carried out a plenary session; many teachers felt it was better to ascertain what had been learnt during the main activity • The use of teaching assistants was particularly helpful and made it easier to use segregated activities	• Several teachers felt the emphasis on mental maths had increased pupils' confidence, particularly for the less able pupils • Teachers felt the NNS had improved the ability of pupils to spot different ways of working out calculations • Pupils in mixed-aged classes often spent too much time waiting or on holding activities • The needs of able older pupils may not be met adequately, but able younger pupils may be stimulated by the presence of older pupils

(11) Hardman F (2003) Interactive whole class teaching in the national literacy and numeracy strategies. Paper presented at the Learning Conference 2003: What learning means. University of London, Institute of Education, 15–18 July.

See number 12.

(12) Hardman F, Smith F, Mroz M, Wall K (2003) Interactive whole class teaching in the national literacy and numeracy strategies. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September. http://www.leeds.ac.uk/educol/documents/00003267.htm

Overall weight of evidence score = medium

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Explores patterns of whole class interaction in the NLS and NNS, and whether effective teachers of literacy and numeracy differ in their discourse strategies	• Discourse strategies used by 72 teachers (35 literacy, 37 numeracy) based on systematic classroom observation, discourse analysis of lesson transcripts and teacher questionnaire; half the teachers selected as highly effective, other half as average, based on value-added data on their pupils' progress	• Data indicate that traditional patterns of whole class interaction have not been dramatically transformed by the NLS and NNS • Whole class teaching in the NNS* was intended to be more interactive and to promote high quality of dialogue and discussion, but most of the discourse in numeracy lessons involved teachers asking closed questions, evaluating, explaining and directing, and pupils answering a question • Teachers had no clear concept of what interactive whole class teaching meant and, although they reported that they valued and frequently invited pupils to elaborate on their answers, this was not supported by the lesson	• Pupils' answers were usually short (three words or less) and pupils' contributions were rarely sustained or extended to encourage higher cognitive interactions; this is unlikely to foster greater confidence and competence in their ability to articulate their ideas regarding their understanding of maths

		transcripts • Effective teachers had a more interactive style in terms of quantity, but not in terms of quality	

* **Note:** The NNS framework states that 'high quality direct-teaching is oral, interactive and lively ... in which pupils are expected to play an active part by answering questions, contributing points to discussion, and explaining and demonstrating their methods to the class' (DfEE, 1999, p 11).

(13) Hopkins C, Pope S (2000) Children's numeracy skills and the use of educational television programmes. In Jaworski B (ed.) *Proceedings of the British Society for research into learning mathematics: proceedings of the day conferences held at the University of Exeter, 26 February, and the University of Loughborough, 6 May,* Volume 20(1&2). London: British Society for Research into Learning Mathematics, pages 11–17.

Overall weight of evidence score = medium

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Explores whether the use of educational television programmes can improve numeracy	• Pre-test and post-test assessment task scores for an experimental group (four Year 1 classes) and a control group (three Year 1 classes) involving 157 pupils, and classroom visits to obtain background information	• Use of five television programmes (each about 10 minutes in length) over a two-week period based on key NNS objectives (e.g. counting, reciting numbers, odds and even) coupled with associated worksheets can be incorporated into the three-part lesson format successfully • School visits indicate that the pupils in the experimental group showed enthusiasm for the television programmes and were active watchers	• Overall, the evidence that the use of the television programme led to greater gains was inconclusive • However, there is some evidence to indicate that some low attaining pupils in the experimental group did improve substantially

(14) Huckstep P, Rowland T, Thwaites A (2002) Primary teachers' mathematics content knowledge: what does it look like in the classroom? Paper presented at the British Educational Research Association Annual Conference, University of Exeter, 12–14 September. Available from: http://www.leeds.ac.uk/educol/documents/00002534.htm

Overall weight of evidence score = medium

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Explores whether student teachers' subject knowledge, or lack of it, is evident in their teaching, with a focus on *connecting*, *responding* and *exemplifying*	• Videotapes of 24 mathematics lessons, plus a *descriptive synopsis* of the lesson written by an observer, which was then elaborated into an *analytic account* after viewing the videotape	• Data indicate the three part lesson structure has been adopted • Examples of both taking or missing opportunities for making connections between mathematical ideas appears to reflect the strength or weakness respectively of certain subject matter or pedagogical content knowledge • Positive and negative examples of responding to pupils' articulated ideas indicate these are enabled (or otherwise) by their content knowledge for teaching • The wisdom (or otherwise) of trainees' choice of examples when teaching is a strong indicator of their mathematical content knowledge for their teaching	• Illustrative examples are given to indicate how more effective teaching is grounded in teachers' subject knowledge • Lesson vignettes and extracts from teacher-pupil interaction demonstrate how more effective teaching in terms of *connecting*, *responding* and *exemplifying* can aid the development of pupil confidence and competence, although the effects on pupil outcomes in these examples are implicit rather than explicit

(15) Jones DV (2002) A comparative analysis of national numeracy initiatives in England and Wales. Paper presented at the British Educational Research Association Annual Conference, University of Exeter: 12–14 September.

See number16.

(16) Jones DV (2003) Learning from recent numeracy initiatives in England and Wales. Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh: 11–13 September.

Overall weight of evidence score = high

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Compares the impact of the NNS in England with the numeracy initiative in Wales	• Questionnaires completed by 243 and 249 headteachers (or a delegated member of staff) in England and Wales respectively	• The overwhelming majority of head teachers in England reported that the NNS had improved teacher motivation and the quality of teaching; the pattern of responses from Welsh head teachers was similar • A more frequent use was reported in England (compared with Wales) of the three part lesson, whole class teaching, pupils working as groups, and teaching of mental maths	• The overwhelming majority of head teachers in England reported that the NNS had improved pupil motivation and pupil attainment in numeracy; the pattern of responses from Welsh head teachers was similar

(17) McSherry K, Ollerton M (2002) Grouping patterns in primary schools. *Mathematics in School* **31:** 2–7.

Overall weight of evidence score = high

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Explores school policies towards pupil grouping for numeracy lessons	• Questionnaire completed by 192 primary schools	• The majority of schools group pupils by ability for all or part of the time: by ability within the class (55%), setting across different classes (20%), by ability for part of the time (15%) and mixed ability (10%) • A marked increase in ability grouping has occurred in the two years prior to and in the year following September 1999 (when the NNS was introduced) in order to implement the NNS and/or to improve SAT scores • Grouping by ability was seen to make whole class teaching and providing differentiated work more manageable	• Data indicate that grouping by ability is occurring at an increasingly younger age (in Year 2, 14% of schools set and 47% employ within-class ability groups) • Authors express the concern that pupils spending up to seven years in low ability groups or sets are likely to become mathematically unconfident and disaffected

(18) Myhill D (2002) Bad boys and good girls?: patterns of interaction and responses in whole class teaching. *British Educational Research Journal* **28:** 339–352.

Overall weight of evidence score = medium

Focus of study	**Data collected**	**Key claims/evidence regarding the classroom teaching of early mathematics**	**Key claims/evidence regarding pupil confidence and competence in early mathematics**
• Explores whether patterns of classroom interaction for boys and girls can explain boys' underachievement	• Teacher interviews, pupil interviews and classroom observation using structured observation schedules in 36 classes (six classes from each of Years 1, 4, 5, 8, 9 and 10); the observation schedule makes a distinction between whole class teaching episodes and teaching episodes where pupils are engaged in individual, pair or group work; both literacy and numeracy hours in primary schools are included in the data • Data presented in the paper focuses on the whole class teaching episodes based in 106 teaching sessions	• Whole class teaching episodes include 'joins in collective response', 'puts hand up', and 'answers question after invitation'; these are taken to indicate positive interactions • For the category 'answers question after invitation', the Year 1 underachieving boys participated more often than other pupils, which is interpreted as indicating that teachers using this strategy during whole class teaching increase the participation of underachieving boys	• Underachievers are less likely to participate positively in the classroom and more likely to engage in off-task interactions • This was evident for underachieving boys in Year 1 for the category of 'joins in a collective response' (which is a frequent feature of the teaching repertoire) • For the category 'puts hand up', there was no gender difference amongst Year 1 underachieving pupils

(19) Pinel AJ (2002) Catechetics observed in mathematics classrooms. In Cockburn AD and Nardi E (eds) *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education: 21–26 July*, Volume 1. Norwich: University of East Anglia, page 312.

Overall weight of evidence score = medium

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Explores the occurrence of catechetics (question-and-answer interactions) in mathematics lessons	• Lesson observations in 12 schools (pupils' ages ranged from 4 to 14 years) together with teacher self-evaluation of the lesson; analysis of 12 lessons are presented in this paper	• Teacher-directed catechetics during the oral/mental starter occurred very frequently in some lessons and hardly at all in other lessons • Very few lessons involved inverse catechetics (i.e. pupils' use of question-and-answer with their teacher or with peers) • Lessons also differed in the extent to which they involved higher order, middle order and lower order questions • Teachers appear to think more of such interactions are occurring in their lessons than the lesson observations bear out*	(None)

* **Note:** This finding is in line with other studies which report that teachers overestimate how much interaction is occurring in their lessons.

(20) Raiker A (2002) Spoken language and mathematics. *Cambridge Journal of Education* **32**: 45–60.

Overall weight of evidence score = high

Focus of study	Data collected	Key claims/evidence regarding the classroom teaching of early mathematics	Key claims/evidence regarding pupil confidence and competence in early mathematics
• Explores the extent to which the mathematical vocabulary used in three-part maths lessons can cause problems in the teaching and learning of mathematical concepts	• Data were collected during one numeracy lesson for each of six classes: one Year 2 class, two mixed Year 3/4 classes, and three Year 4 classes • Discourse was collected from teacher-class interactions during the mental/oral starter and the introduction to the main teaching part of the lessons; discourse analysis was undertaken for six groups of teachers and pupils	• Teachers use mathematical vocabulary more than pupils do • Problems can arise in lessons when the meaning of mathematical words is not established • Teachers did not appear to be aware of the importance of the mathematical vocabulary they were using or to have planned for its introduction, explanation of meaning, and repetition • Teachers did not distinguish differing purposes of language in a mathematics lesson and therefore did not alert pupils when essential key vocabulary was being taught	• Data demonstrate that mathematical language plays a crucial role in the building of sound concepts and the subsequent development of mathematical thinking • Data indicate that pupils' understanding may be more insecure than is evident from their success in completing worksheets because they may not have grasped the meaning of the mathematical vocabulary that is being used