



REVIEW

May 2005

The impact of population inclusivity in schools on student outcomes

Review conducted by the Inclusive Education Review Group

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

AUTHORS

Afroditi Kalambouka, Peter Farrell, Alan Dyson, Ian Kaplan

School of Education
University of Manchester
Oxford Road
Manchester
M13 9PL

REVIEW GROUP MEMBERSHIP

Afroditi Kalambouka	Research Associate
Peter Farrell	Professor of Special Needs and Educational Psychology
Alan Dyson	Professor of Special Needs Education
Ian Kaplan	Research Assistant

ADVISORY GROUP MEMBERSHIP

The Review Group was supported by an Advisory Group comprising academic and research staff from the Educational Support and Inclusion (ESI) Group in the School of Education, University of Manchester. Staff who had been involved in previous EPPI-Centre reviews, in particular Dr Andy Howes and Professor Mel Ainscow, were particularly helpful in providing guidance on every stage of the review process. All ESI staff were involved in the data-extraction stage for at least one study. Throughout the period of the review, members of the Advisory Group were invited to attend the formal Review Group meetings and additional informal meetings were held on a regular basis. Full membership of the Advisory Group is given in Appendix 1.1.

CONFLICTS OF INTEREST

At the time this review was being carried out, two members of the Review Group, Professors Dyson and Farrell, were also involved in a study funded by the DfES on the relationship between inclusion and pupil achievement. This is referred to in several places throughout the report. The potential conflict of interest was dealt with by ensuring that there were two distinct products. The literature review for the DfES report (Dyson *et al.*, 2004) draws extensively on the systematic review but is separate from it. Although the steering committee for the DfES study was aware of, and provided some advice on, the direction of the review, it only steered the production of its own report.

ACKNOWLEDGEMENTS

We are grateful for the funding assistance of the following organisations and bodies:

DfES through the EPPI-Centre

The University of Manchester

The DfES through funding provided for the research project on Inclusion and Pupil Achievement, published in October 2004

We are also grateful for the contributions of everyone who worked on the review, including members of the Advisory Group in the Educational Support and Inclusion Research and Teaching Group. We acknowledge the support of staff in the EPPI-Centre who assisted with advice and practical support at all stages of the review, in particular Zoe Garrett and her colleagues.

LIST OF ABBREVIATIONS

BEI	British Education Index
DfES	Department for Education and Skills
EAL	English as an additional language
EBD	emotional/behavioural difficulties
EPPI-Centre	Evidence for Policy and Practice Information and Co-ordinating Centre
ERIC	Educational Resources Information Centre
ESI	Educational Support and Inclusion
ESRC	Economic and Social Research Council
GES	general education student
GPE	general physical education
ICM	integrated classroom model
LEA	local education authority
NCE	normal curve equivalent
PE	physical education
PMLD	profound and multiple learning difficulties
SAT	Stanford achievement test
SEN	special educational needs
SLD	severe learning difficulties
UK	United Kingdom
USA	United States of America

This report should be cited as: Kalambouka A, Farrell P, Dyson A, Kaplan I (2005) The impact of population inclusivity in schools on student outcomes. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

The opinions expressed in this review are not necessarily those of the EPPI-Centre or the review funders. Responsibility for the views expressed remains solely with the review authors.

© Copyright

Authors of the systematic reviews on the EPPI-Centre website (<http://eppi.ioe.ac.uk/>) hold the copyright for the text of their reviews. The EPPI-Centre owns the copyright for all material on the website it has developed, including the contents of the databases, manuals, and keywording and data-extraction systems. The centre and authors give permission for users of the site to display and print the contents of the site for their own non-commercial use, providing that the materials are not modified, copyright and other proprietary notices contained in the materials are retained, and the source of the material is cited clearly following the citation details provided. Otherwise users are not permitted to duplicate, reproduce, re-publish, distribute, or store material from this website without express written permission.

TABLE OF CONTENTS

- SUMMARY 1
- Background..... 1
- Aims..... 1
- Methods 2
- Results..... 3
- Conclusions 4
- 1. BACKGROUND 6
 - 1.1 Aims and rationale for current review 6
 - 1.2 Definitional and conceptual issues 7
 - 1.3 Policy and practice background..... 8
 - 1.4 Research background..... 10
 - 1.5 Authors, funders and other users of the review 12
 - 1.6 Review questions..... 13
- 2. METHODS USED IN THE REVIEW 14
 - 2.1 User involvement..... 14
 - 2.2 Identifying and describing studies..... 15
 - 2.3 In-depth review 21
- 3. IDENTIFYING AND DESCRIBING STUDIES: RESULTS..... 25
 - 3.1 Studies included from searching and screening 25
 - 3.3 Identifying and describing studies: quality-assurance results..... 29
- 4. IN-DEPTH REVIEW: RESULTS 30
 - 4.1 Further details of studies included in the in-depth review 30
 - 4.2 Synthesis of evidence 33
 - 4.3 In-depth review: quality-assurance results 59
 - 4.4 Nature of actual involvement of users in the review and its impact 59
- 5. FINDINGS AND IMPLICATIONS..... 61
 - 5.1 Summary of principal findings..... 61
 - 5.2 Strengths and limitations of this systematic review 64
 - 5.3 Implications..... 65
- 6. REFERENCES 68
 - 6.1 Studies included in map and synthesis..... 68
 - 6.2 Other references used in the text of the report 70
- Appendix 1.1: Advisory Group membership 73
- Appendix 2.1: Inclusion and exclusion criteria..... 74
- Appendix 2.2: Search strategy for electronic databases 76
- Appendix 2.3: Journals handsearched 80

Appendix 2.4: EPPI-Centre keyword sheet, including review-specific keywords .81	
Appendix 3: Details of studies included in the in-depth review	83

SUMMARY

Background

National education policy in England (and more generally across the UK) pursues avowedly inclusive aims, but within the context of a highly demanding 'standards' agenda which focuses on meeting targets for raising the attainments of students to specified levels. Not surprisingly, this apparently twin-track approach has led to some concerns as to how schools can reconcile the imperatives to which they are subject and whether increasing their inclusiveness might not at the same time reduce their capacity to produce good outcomes for their students. Although there are substantial reviews of research on the impact of inclusion for pupils with special educational needs (SEN), to date little is known about the impact of inclusion on the academic and social outcomes for pupils without SEN. This is an important aspect of the inclusion debate at a time when questions are being raised about the viability of inclusion (Ofsted, 2000) and when some teachers are expressing concerns about the increased inclusion of particular groups of pupils, especially those with emotional and behavioural problems.

Aims

The aim of this review is to explore empirical evidence about the relationship between the inclusiveness of a school and the outcomes it produces for its student population, especially the population of students without special educational needs. This is a significant gap in the evidence-base which is currently being used by educators to inform the inclusion debate. The danger of leaving this gap unfilled is that policy and practice will be developed on the basis of an enthusiasm for inclusion or an antipathy towards it, neither of which is informed by robust evidence.

The aims of the review are closely linked to those of a major study that we recently completed for the DfES on the relationship between inclusion and pupil achievement in English schools (Dyson *et al.*, 2004) and findings from this study complement those of the systematic review. However, the findings of this major UK study have not been included in the present systematic review since the study was still in progress at the time of the review.

This review seeks to answer the following question:

What evidence is there that the inclusiveness of schools has impacts on outcomes for the students without special needs in those schools?

There is a large volume of literature on the impact of inclusion for the pupils included rather than those without SEN and this falls outside the scope of this review. 'Population inclusivity' is defined in this review as what the Audit Commission has recently called 'presence' (Audit Commission, 2002). This refers to the inclusion in a regular school population of students who in otherwise comparable schools might be placed outside the mainstream. It excludes evidence relating to inclusion of pupils with English as an additional language (EAL) or the inclusion of pupils from ethnic minorities. In addition, although very

important, the review ignores evidence of the impact of inclusivity on the following groups: teachers, headteachers, managers and other school-related staff, parents and carers of pupils with SEN and/or pupils without SEN. The review also only focuses on pupils aged 5–16.

Methods

Identifying and describing studies

Inclusion criteria

We reviewed studies which met the following criteria:

- They are in English (given limitations on available resources).
- They report on the results of empirical research rather than being purely theoretical or exhortatory.
- They are concerned with the phases of compulsory schooling.
- They report the outcomes of the ‘intervention’ (i.e. an increase in population inclusivity).
- They report these outcomes in relation to students without SEN or whole school populations (but not simply in relation to students with SEN).
- They report robust evidence of the impact of the intervention:
 - through a longitudinal study of one school or
 - by comparison with a similar but less inclusive school (with a lower level of population inclusivity) or
 - by comparison between different conditions within the same school (such as more and less inclusive classes) or
 - by some other equally robust means
- They are concerned with the impact of inclusion on pupils’ personal, social and/or academic outcomes.

Search strategy

Searches through electronic databases constituted the main strand of our search strategy. These covered journal articles, books and book chapters, conference papers and proceedings, theses, dissertations and reports. The strategy was refined in consultation with members of the Advisory Group and members of staff at the EPPI-Centre. A set of agreed search terms to guide the review was subsequently tested out in a particular database and new terms were added or existing terms were altered or removed. When the Review Group was satisfied that, by using the devised set of terms, all the available items stored in the databases could be identified, the search was run in all relevant databases.

Mapping

The inclusion/exclusion criteria were applied to all studies and, when there was uncertainty about the inclusion of a study, the criteria were refined and clarified further. This resulted in a set of very specific criteria and, consequently, the mapping exercise led to the identification of a small number of studies. All studies in the descriptive map were also included in the in-depth review.

Synthesis

The framework adopted for synthesising studies viewed the impact of population inclusivity according to the following variables:

- area of key impact (academic or social outcomes)
- type of school (primary versus secondary schools)
- nature of SEN (adopting four main categories)
 - cognition and learning
 - behavioural, emotional and social development
 - sensory and/or physical needs
 - communication and interaction

Results

After completing the searches and selecting publications that met the criteria, the number of studies included in the review was reduced to 26, all of which were subjected to the data-extraction process. All these were evaluation studies, 15 of which were 'naturally occurring' and 11 'researcher manipulated', involving some form of experimental design. The majority of these studies were carried out in the United States of America (USA) (N = 22); there were also two studies from Australia, one from Canada and one from Ireland. There was a slight preponderance of studies in which the included pupils experienced difficulties in the area of cognition and learning, although pupils with other types of difficulties were also mentioned. The majority of the studies (N = 21) focused on academic outcomes and these were measured in a wide variety of ways, including class tests, national examinations and teacher ratings.

Virtually all the studies (N = 21) focused on the outcomes of inclusion for primary aged pupils. The nature of the inclusion experienced by the pupils with SEN was described in different ways. In some studies (N = 16), this was described as the proportion of pupils with SEN in a mainstream class, whereas in others (N = 14) it was described as the number of hours per week (or day) that a child with SEN spent in a mainstream class. Some studies described inclusion in both ways.

Many of the 26 studies focused on more than one type of outcome. Although 12 focused solely on academic outcomes and five addressed social outcomes, there are nine that considered both academic and social outcomes. This means that the 26 studies yielded more than 26 findings. Indeed there are 40 findings from the studies that relate to one or other or both of these outcomes.

The majority of studies (N = 24) reported the results of including pupils with difficulties in the area of cognition and learning, although studies, more often than not, refer to groups of pupils with a variety of SEN. Therefore, it is difficult to provide direct conclusions regarding the impact of including pupils with a specific type of SEN on the academic and/or social or other outcomes of all school pupils. However, there seem to be more negative outcomes reported when pupils with emotional/behavioural difficulties (EBD), as a main or additional difficulty, were included compared with the other types of SEN. There were no studies reporting negative outcomes for the majority of pupils in school/class when pupils with physical and/or sensory and communication difficulties were included.

Of these 40 different findings, nine (23%) indicate that there was a positive academic and/or social impact on non-SEN pupils of including pupils with SEN. Six (15%) suggest a negative impact, 21 (53%) a neutral impact and four (10%) suggest mixed impacts. Taken as a whole, the findings indicate that placing children with SEN in mainstream schools is unlikely to have a negative impact on academic and social outcomes for pupils without SEN.

Further analysis indicated the following:

- Findings are slightly more positive for academic rather than social outcomes.
- At secondary level, where there are very few studies, outcomes are slightly more mixed.
- Some of the findings suggest that the inclusion of pupils with SEN in primary schools can have a positive impact on the achievement of their mainstream peers, particularly if the support offered to the pupil with SEN is well managed.

There is no evidence about the impact of placing pupils with SEN in mainstream schools on achievement in different curriculum areas. Hence none of the findings indicate that the 'inclusion effect' is more or less serious for any one particular curriculum area.

Conclusions

There is nothing in this review or in the DfES study in which we have been involved (Dyson *et al.*, 2004) to suggest that the commitment to inclusion in relation to pupils with SEN is likely to have a significant impact on overall levels of attainment in mainstream schools. This suggests that the government, LEAs and schools should have no concerns about pursuing the inclusion agenda. However, policymakers should pursue inclusion policies in an informed way, consulting with all relevant stakeholders at all times. In addition, the lack of studies in the secondary sector suggests that schools and LEAs should pursue the inclusion agenda with some caution and, where possible, commission research that can explore this complex area in more depth.

In relation to practice, this review suggests that schools, parents and LEA professionals should have no concerns about the impact of inclusion on achievement, especially in primary schools. This applies across all four categories of SEN. However, these studies and other research reviews (Harrower, 1999; Farrell, 2000) indicate that successful inclusion does not occur in a vacuum. Parents, teachers and pupils need to be fully committed to the idea; programmes

of work have to be carefully planned and reviewed regularly; and support staff need to work flexibly as a team and receive appropriate support and training.

There are a number of implications for further research that arise from this review. In particular, the lack of studies that focused on secondary schools indicates that this is a key area that future studies should address. In addition, further studies could focus on the impact of including larger groups of students with SEN. There is also a need for more longitudinal research that could trace the relationship between inclusion and the achievements of non-SEN pupils over time. Finally, more studies should focus on the views of pupils without SEN about inclusion. Given current interest in involving users in planning, carrying out and evaluating research, it is surprising that so few studies actually focus on the pupils' views.

1. BACKGROUND

1.1 Aims and rationale for current review

National education policy in England (and more generally across the UK) pursues avowedly inclusive aims, but within the context of a highly demanding ‘standards’ agenda which focuses on meeting targets for raising the attainments of students to specified levels. Not surprisingly, this apparently twin-track approach has led to some concerns as to how schools can reconcile the imperatives to which they are subject and whether increasing their inclusiveness might not at the same time reduce their capacity to produce good outcomes for their students.

Set against this concern are some powerful theoretical arguments suggesting that an inclusive approach by schools should enable them to generate better student outcomes (Ainscow, 1991; Lipsky and Gartner, 1997; Skrtic, 1991). There is also a good deal of empirical evidence which, whilst not supporting some of the more ambitious claims made for the effects of inclusion, suggests that some groups of students – particularly those with special educational needs – who are ‘included’ (i.e. not placed elsewhere when that might be the normal practice) in regular schools do no worse socially and academically than if they were placed outside the mainstream (see Farrell, 2000, and Lunt and Norwich, 1999, for recent reviews). In addition, the previous reviews undertaken by this Review Group (Dyson *et al.*, 2002; Howes *et al.*, 2003), were related to this general issue. The first (Dyson *et al.*, 2002), which focused essentially on process-oriented case studies, identified some which *suggested* that outcomes for all students might be better in inclusive schools – although these did not set out to demonstrate this claim unequivocally. The second (Howes *et al.*, 2003), on the impact of paid adult support in the classroom, found no evidence of the negative impact of support on the achievements of non-SEN pupils.

What have not yet been subject to a systematic review, however, are studies which set out to explore empirically the relationship between the inclusiveness of a school and the outcomes it produces for its student population, especially the population of students without special educational needs. This is a significant gap in the evidence-base which is currently being used by educators to inform the inclusion debate. The danger of leaving this gap unfilled is that policy and practice will be developed on the basis of an enthusiasm for inclusion or an antipathy towards it, neither of which is informed by robust evidence.

This review seeks to address this situation by identifying and synthesising such studies as exist. The review will therefore seek to provide an answer to the above issue by collating evidence surrounding the impact of inclusion on outcomes for all students in those schools. More specifically, the aims of this study are as follows:

- To identify and describe the research literature that has investigated the impact of inclusivity of schools/classes on outcomes for pupils
- To synthesise the data and results from these studies
- To summarise what is the impact of inclusion on pupils without SEN

- To consider these findings in terms of implications for practice, policy and research

These aims are closely linked to those of a major study that we have undertaken for the DfES on the relationship between inclusion and pupil achievement in English schools. This has involved a detailed statistical analysis of individual pupil data from the 2002 national pupil database that has sought the relationship between school and LEA inclusivity, and pupil attainment across all four key stages (Dyson *et al.*, 2004). We have also carried out a series of 16 case studies of high and moderately achieving inclusive schools. Findings from this study complement those of the systematic review.

1.2 Definitional and conceptual issues

The notion of 'inclusive schooling' is, of course, complex, ambiguous and contested (Dyson, 1999; Farrell and Ainscow, 2002; Farrell *et al.*, 2004). It can refer to many different aspects of schools' policies and practices in relation to different groups of students. For the purposes of this review, however, we will focus on one aspect only: that is, what we might call 'population inclusivity', or what the Audit Commission has recently called 'presence' (Audit Commission, 2002). This refers to the inclusion in a regular school population of students who in otherwise comparable schools might be placed outside the mainstream.

Again, there is a range of groups for whom provision outside the mainstream is sometimes made: children whose behaviour is disruptive and who elsewhere might be excluded from school; children from ethnic and/or linguistic groups who elsewhere might be out of school or in special provision; and young mothers who elsewhere might be educated in special units, and so on. However, we propose to focus on students identified as having special educational needs (or the equivalent in other national systems). This is partly for pragmatic reasons, so that the demands of the review can be managed within the available resource constraints. However, there are also good theoretical reasons for this focus. In many national systems, the 'inclusion' or otherwise of such students is a live issue, accompanied by a high level of research activity. Moreover, most countries have more-or-less formal systems of identification and assessment, which increases the probability that like populations can be compared across schools. Finally, students with special educational needs represent a diverse range of groups and it ought to be possible to investigate any differences in the impacts on outcomes attributable to the inclusion of these groups. Although further specific reviews of the literature may still be needed, it is reasonable to suppose that policy and practice implications from the current review can cautiously be extrapolated in the interim.

On this definition, a more inclusive school has in its population a greater number and/or a wider range of students with special educational needs who might otherwise be placed outside the mainstream than does a less inclusive school. This definition is, of course, context-specific. Different education systems routinely place different proportions of students in non-mainstream provision. This means that the synthesis of findings from across systems will have to be sensitive to this issue. However, it does not invalidate comparisons within systems where the effects of different levels of inclusiveness may be apparent.

1.3 Policy and practice background

As the above sections imply, the issue of inclusion has become a key feature of discussions about the development of education policy and practice in many countries. The movement has been strongly endorsed internationally by the Salamanca Statement (UNESCO, 1994) and reflects the United Nations' global strategy of 'Education for All'. Both have had a major impact on policy developments in many different countries. This is confirmed by recent accounts of trends in inclusion in different countries (see, for example, Egelund, 2000; Miejer, 1998; Norwich, 2002). There is also no shortage of books and articles that have extolled the values of inclusion and which have provided a whole range of accounts of 'good practice' in inclusive education (see, for example, Ainscow, 1999; Ballard 1999; Mittler, 2000; Thomas and Vaughn, 2004).

Despite these developments, however, inclusion remains a complex and controversial issue that tends to generate heated debates (e.g. Brantlinger, 1997). This is reflected in the ongoing discussions regarding the definition of inclusion (Farrell and Ainscow, 2002). We are focusing on studies that are solely concerned with 'population' inclusivity, but, to put this term into context, it is important to see how it relates to other conceptions of inclusion.

Up until the early 1990s, the term 'inclusion' was hardly employed. Instead the terms 'integration' or 'mainstreaming' were used and these referred exclusively to the placement of pupils with special needs in general education classes in mainstream schools (i.e. to 'population inclusivity'). There were different degrees of integration, that could vary from fulltime placement of a child with disabilities in a mainstream class in his/her local school (functional integration) to the placement of a pupil in a special class or unit attached to a mainstream school (locational integration) (Hegarty, 1991). However there was little difference between locational integration and a traditional segregated special school. Furthermore other arrangements also came under the banner of integration, notably the occasional visits by children in special schools into a mainstream school, which, again, is hardly radical in terms of signalling a major policy shift.

An obvious problem with defining integration solely in terms of the location of provision (i.e. the setting in which a pupil is placed) is that it indicates nothing about the quality of the education that is received. Are pupils placed in units attached to a mainstream school, for example, more 'integrated' than if they were taught in a special school? Jupp (1992) argues that such units can be just as segregating. Indeed, even pupils with special needs placed in a mainstream class may be isolated from the rest of the class and not truly 'integrated' within the group, particularly if they work with a support worker in one-to-one sessions for the majority of each day. Integrated placements, therefore, may still leave the pupil 'segregated' (Harrower, 1999).

Partly for these reasons, the term 'inclusion' has become a more usual way of describing the extent to which a pupil categorised as needing to receive special educational provision is truly 'integrated'. Used in this way, the term refers to the extent to which a school or community welcomes such pupils as full members of the group and values them for the contribution which they make. This implies that for inclusion to be seen to be 'effective', all pupils must actively belong to, be welcomed by and participate in a mainstream school and community. Their diversity of interests, abilities and attainment should be welcomed and be seen to

enrich the life of the school. In this sense, as Ballard (1999) argues, inclusion is about valuing diversity rather than mere assimilation.

In relation to pupils with SEN, this discussion has led us to build on the Audit Commission's (2002) definition of inclusion and relate the concept to outcomes for pupils in terms of their presence, acceptance, participation and achievement in mainstream settings:

- *Presence* refers to the extent to which pupils attend lessons in mainstream settings in local schools and committees. (This is similar to 'population inclusivity'.)
- *Acceptance* refers to the extent to which other staff and pupils welcome all pupils as full and active members of their community.
- *Participation* refers to the extent to which all pupils contribute actively in all the school's activities.
- *Achievement* refers to the extent to which pupils learn and develop positive views of themselves.

It is argued that, for a school to be truly inclusive, all four conditions should apply to all children in the schools regardless of their abilities and disabilities and of their ethnic origin, social class or gender. It is not, for example, sufficient for children to simply be *present* in a school. They need to be *accepted* by their peers and by staff, they need to *participate* in all the school's activities and they need to attain satisfactory levels of *achievement* in their work and behaviour. This formulation is proactive in the sense that it sets goals for schools, local authorities, communities and governments, and can act as a benchmark against which to judge the extent to which inclusive policies and practices are working.

There are, of course, considerable pressures in schools that have the potential to act as a barrier towards inclusion. In the UK, for example, schools are required to raise academic standards at the same time as being asked to develop more inclusive policies and practices. Similar pressures are exerted on schools in the USA. Many argue (e.g. Evans and Lunt, 2002) that these competing priorities can make it more difficult for schools to include fully children with disabilities.

It is for these reasons that mainstream schools, while welcoming the values underpinning movements towards inclusion, are still cautious about the impact that such developments will have on their pupils. This seems to be particularly the case for pupils with emotional and behavioural difficulties (Farrell and Polat, 2003; Farrell and Tsakalidou, 1999). Hence they see the potential negative consequences for acceptance, participation and achievement of placing children with SEN in their schools. However, as indicated by Norwich (2002), there are some LEAs which have drastically reduced the numbers of pupils that they send to special schools and hence they seem to have been successful in allaying the fears expressed by teachers in mainstream schools. Nevertheless the pattern is mixed and some LEAs still retain most of their special schools.

The findings from this systematic review should have the potential to inform future policy and practice in this complex area. For, if it is found that the presence of pupils with SEN in a school has no impact on the overall achievements of non-SEN pupils, then schools should have nothing to fear if they chose to admit more pupils with SEN. If, however, research suggests that the presence of such pupils

has a negative impact on attainment, then governments and LEAs should consider revising their policies.

1.4 Research background

The vast majority of studies and reviews on the impact of inclusion have focused on pupils with SEN rather than on their non-disabled peers. Some of these reviews, however, devote a section on the impact of inclusion on pupils without SEN (Harrington, 1997; Harrower, 1999). There is also a large volume of literature devoted to how students with SEN perform in relation to their non-disabled peers (Banerji and Dailey, 1995) although they may include sections on the impact for students without SEN.

There is, however, a limited number of small scale studies and reviews that have addressed the issue of the impact of inclusion on non-disabled students. Peltier (1997), for example, reviewed studies on the impact of inclusion on non-disabled children in terms of their academic growth, social growth, and teacher time and attention. He reported that, in terms of academic growth, the three studies that provided relevant evidence (one of them on preschool children) concluded that the inclusion had not had a negative impact on students' academic objectives. He also reviewed the results of two studies on the opinions of parents, teachers and students, neither of which reported any harm to the developmental progress of non-disabled pupils. In terms of social growth, Peltier (1997) found that all five studies he reviewed presented positive findings, mostly relating to reduced fear of human diversity; growth in social cognition; and the development of personal principles, and warm and caring friendships. However, the studies reviewed by Peltier (1997) mostly refer to experiences of non-disabled students rather than social outcomes. Finally, the only study he reviewed that has directly investigated the impact of inclusion on teacher's time and attention showed no significant time losses and interruptions due to the presence of pupils with SEN in the classroom.

Contrary to Peltier's (1997) review, Petch-Hogan and Haggard (1999) have found some negative attitudes among pupils without SEN towards inclusive practices and towards their disabled peers. However, overall, Petch-Hogan and Haggard (1999) found that the majority of seven studies they reviewed on academic and social outcomes of inclusion on students are either positive or neutral for all students concerned.

Harrower (1999) reviewed the literature concerning the inclusion of students with severe disabilities. Although the main focus is on the impact for pupils with SEN, he also reported six studies that measured the effects on students without disabilities. Most of them reported positive outcomes in terms of attitudes, acceptance, knowledge of disabilities and friendships; one study, however, concluded that placing a student with severe disabilities in regular education was not enough to reduce the negative stigma of SEN. The author also reviewed the research on parental views, and concluded that there were overall positive attitudes towards inclusion among parents of 'typical' children.

Moore, Gilbreath and Maiuri (1998) also reviewed the literature on how effective it is to educate students with disabilities in general education classrooms. The reviewers report findings on academic outcomes separately for students with mild disabilities, with significant disabilities and without disabilities, as well as overall findings on the impact on attitudes and relationships. As far as pupils without

disabilities are concerned, the authors say that they did not find any study reporting a negative impact on students without disabilities in terms of academic achievement. However, Moore *et al.* (1998) only reviewed three studies on academic achievement for students without disabilities. In terms of attitudes, experiences and relationships, they report overall positive experiences and improved attitudes towards inclusion by all students (with and without disabilities). The authors conclude in favour of inclusion for all concerned, claiming that any studies that report negative impacts could be accounted for by the lack of support and specialised education that the pupils receive.

Staub and Peck (1994/1995) concluded that inclusion not only 'does not harm non-disabled children' but that there are potential benefits. They describe the potential benefits in terms of the following categories: (a) reduced fear of human differences, (b) growth in social cognition, (c) improvements in self-concept, (d) development of personal principles, and (e) warm and caring friendships. Their review is not systematic but mainly focuses on evidence from a variety of sources, including ethnographic studies, experimental designs, and so on, and they provide a number of quotes from non-disabled peers. They do not report studies that have looked primarily in academic achievement, since, as they state, they interpret the findings mainly from the needs of all children. They view those needs mainly from a perspective of values and ethics, and on these grounds, they recommend inclusion.

In two more summary papers, Staub (1996, 1999) strongly supports inclusion as the best philosophy and practice for educating children with SEN. She argues that there is no research evidence showing that non-disabled students' learning would suffer because of inclusion, or that non-disabled peers will receive less attention and time from teachers while, on the other hand, there is research findings suggesting that friendships, self-esteem and personal principles will grow and be enriched in inclusive settings (Staub, 1996, 1999).

Salend and Duhaney (1999) reviewed four studies on the academic outcomes of inclusion on pupils without SEN. The findings indicated that pupils with SEN who attended an inclusive classroom did not interfere with the students' academic performance in terms of scores in various subjects and in terms of allocated and engaged instructional time. Furthermore, they found overall positive reported attitudes of students without disabilities towards inclusive classrooms, increased tolerance to individual differences, and greater awareness and sensitivity to human diversity and the needs of others. Salend and Duhaney also reviewed the impact on students with disabilities and on educators, where they have found a greater variety of reported outcomes for students and rather mixed reactions from educators.

Harrington (1997) reviewed the literature regarding the impact of full inclusion on pupils with SEN but he also partially reported on some of the studies' findings on the impact for pupils without SEN. In one study on the academic impact on non-disabled pupils, the effects were neutral while the few studies on social outcomes, mainly relating to acceptance and interactions, present a rather equivocal conclusion.

Manset and Semmel (1997) attempted a rather more systematic review of the literature on inclusive programmes for students with mild disabilities, primarily with learning difficulties. The review also includes effects for students who are at risk and some results for general education students. The review is systematic in terms of setting criteria for included studies: for example, published between 1984

and 1994, using objective measures of academic outcomes, involving programmes that involved school wide interventions and have as their main purpose the mainstreaming of pupils with disabilities. They found eight programmes reported in 11 studies, the results of which suggest overall that organisational and instructional changes incorporated in the programmes have overall positive outcomes on the achievement of non-disabled students. Some of these changes include redesigning general education provision, low student-staff ratio, opportunities for intensive one-to-one instruction, performance monitoring, etc. To their overall question that concerns the effectiveness of programmes for students with SEN, Manset and Semmel (1997) answer that 'the evidence is inconclusive' and that, although there is some evidence towards positive outcomes for these students, there is not a model of 'wholesale superior inclusive programming'. They suggest caution in interpreting such results due to the methodological concerns arising from some of the studies.

All studies and reviews, referred to above, have made a contribution to the literature on the impact of inclusion on pupils without SEN. However, they are all relatively small scale and in all but one case not systematic; taken together, they only make a modest contribution to knowledge in this complex area. The only review that used elements of systematic review methodology was carried out in 1997. Therefore it is appropriate to carry out this EPPI-Centre review both to update the existing knowledge base and to provide a more systematic and comprehensive review of the area.

1.5 Authors, funders and other users of the review

This review is the third EPPI-Centre review that has been carried out on behalf of the Inclusion Review Group. The two previous reviews were concerned with (a) the effectiveness of school-level actions for promoting participation by all students, and (b) the impact of paid adult support on the participation and learning of pupils in mainstream schools. As in the previous two reviews, members of this Group include academics and research staff who have recently completed the Economic and Social Research Council (ESRC) funded research Network project on understanding and developing inclusive schools. This EPPI-Centre review has been led by Dr Afroditi Kalambouka with support from Professors Peter Farrell and Alan Dyson. They have also been assisted by a research assistant, Mr Ian Kaplan. All these members of the Review Group work at the ESI (Educational Support and Inclusion), Faculty of Education, University of Manchester.

The Review Group has been supported by an Advisory Group that comprised members of the ESI, Department of Education, University of Manchester. We have also received some input from other potential users. Experience with previous reviews suggested that it is more productive to have genuine involvement from a small number of users rather than tokenistic involvement from a larger number. Accordingly, we invited Paul Rees (a principal educational psychologist) and Mike O'Connor (a former SEN adviser) to join the Advisory Group and hence to play a role in discussing the progress of the review and to help to shape the final report.

The range of potential users is large. It includes headteachers and teachers, local education authority officers and council elected members in LEAs which are (or are considering) pursuing inclusive policies, national policymakers and government officials, students and their families, disability support and lobby

groups and researchers. The findings of this review will be made available to its most immediate user community. A full report on the DfES project, including the literature review, has recently been published (Dyson *et al*, 2004). Two national dissemination conferences have also been planned and these provide ideal opportunities to 'showcase' the findings of the systematic review to a much wider audience of practitioners, policymakers and other stakeholders than would normally have access to them.

1.6 Review questions

This review seeks to answer the following question:

What evidence is there that the inclusiveness of schools has impacts on outcomes for the students without special needs in those schools?

The review question is quite narrow in focus. Although there is a large volume of literature on the impact of inclusion at both school and pupil level, the present review is only concerned with this impact when it refers to pupils without SEN. The vast majority of literature has focused on the impact of inclusion for the pupils included rather than those without SEN. In addition, although very important, the review will ignore evidence of the impact of inclusivity on the following groups: teachers, headteachers, managers and other school related staff, parents and carers of pupils with SEN and/or pupils without SEN. In addition, again, although inclusion at a very early age may be of crucial importance for all concerned, it is outside the scope of this study to look at the impact of inclusion when this concerns the very young ages or post-compulsory schooling.

2. METHODS USED IN THE REVIEW

This section describes the methods used to carry out this review. The methodology follows the EPPI-Centre guidelines for systematic reviews of research evidence. Firstly we provide an outline of our approach to involving users in this review. This is followed by a detailed description of how the studies were identified, using systematic methods of searching, screening and mapping. Finally, the section provides a summary of the procedures used for assessing the studies' quality and weight of evidence, how they were synthesised and the quality-assurance process.

2.1 User involvement

2.1.1 Approach and rationale

For this review to have any impact on policy and practice, it was essential to consider user involvement from the outset. For it is practitioners who need to be informed of its outcome and to shape their future policies and practices accordingly. LEA offices, headteachers and educational psychologists are key professionals who can shape developments for parents, pupils, schools and LEAs. For these reasons, we considered it essential to consider how these professionals could guide the review and comment on its findings. Furthermore, in view of the policy implications at government level, it was important to be guided by staff working at the DfES.

2.1.2 Methods used

In order to ensure the significant contribution of users in the process of review as well as in its final products, the following methods were followed.

First, the process of review and especially the ongoing results were presented and discussed with a group of approximately 20 educational psychologists, headteachers and LEA advisors. Their feedback was valuable in terms of clarifying some of the issues often raised in research studies, many of which centred on core questions, sharpening discussions on impact of inclusion, and exchanging opinions with the Review Group.

Secondly, it has to be stressed at this point that this particular review has a very specific and distinctive relationship to users. As referred to above, members of the Review Group were commissioned by DfES to undertake an empirical study of the relationship between 'inclusion and pupil achievement'. This involved analysing the National Pupil Database and conducting case studies in highly inclusive schools. The need for this study is reflected by the fact that the government is currently pursuing a policy of increased inclusion and wishes to do so on the basis of best evidence.

This study was managed by a DfES steering committee comprising officials, HMI, leading academics in the field of special and inclusive education, LEA advisors

and headteachers. This committee, which met four times during the course of the review, also served as an important, though informal, user group whose views and advice were helpful in shaping the review questions and in helping to interpret the findings.

2.2 Identifying and describing studies

This section describes how the relevant studies were defined by setting up a list of inclusion criteria. These were subsequently used to make exclusion criteria, from the most general to the more specific. It goes on to provide a summary of how the potential studies were identified by applying inclusion/exclusion criteria and how they were screened, illustrated with a few examples.

2.2.1 Defining relevant studies: inclusion and exclusion criteria

Inclusion criteria

We reviewed studies which met *all* the following criteria:

- They are in English (given limitations on available resources).
- They report on the results of empirical research rather than being purely theoretical or exhortatory.
- They are concerned with the phases of compulsory schooling.
- They report the outcomes of the 'intervention' (i.e. an increase in population inclusivity).
- They report these outcomes in relation to students without SEN or whole school populations (but not simply in relation to students with SEN).
- They report robust evidence of the impact of the intervention:
 - through a longitudinal study of one school or
 - by comparison with a similar but less inclusive school (with a lower level of population inclusivity) or
 - by comparison between different conditions within the same school (such as more and less inclusive classes) or
 - by some other equally robust means
- They are concerned with the impact of inclusion on pupils' personal, social and/or academic outcomes.

Exclusion criteria

The inclusion criteria above were used to set up exclusion criteria. Each study was screened by applying firstly the exclusion criterion 1, then 2, and so on. Studies were excluded if they met *any* of the following criteria:

- The study is not concerned with the impact of population inclusivity (exclusion criterion 1).

- The impact of population inclusivity is not on children without SEN (exclusion criterion 2).
- There is no reference to the impact on personal, social and/or academic outcomes (exclusion criterion 3).
- The study is not concerned with the compulsory phases of schooling (exclusion criterion 4).
- The study is not empirical research (exclusion criterion 5).
- The study is not written in English (exclusion criterion 6).
- The study does not provide evidence of robust measurements of pupil outcomes (e.g. reports pure attitudes of teachers, etc.) (exclusion criterion 7).

The number of studies excluded after applying these criteria is summarised in Figure 3.1. The justification for adopting these criteria is discussed below.

Topic

- There is a growing body of evidence on the so-called ‘school mix’ effect, that is, on the relationship between various demographic characteristics of a school’s population (social disadvantage, levels of prior attainment, levels of SEN) and outcomes for its students. The literature on this issue is interesting contextually for our review but was not central to its focus. We decided only to include such studies where they demonstrate that the composition of schools’ populations results in some part from the inclusion of students who elsewhere might be placed outside the mainstream, and investigate the impact on outcomes of such inclusion.
- As we have indicated, for the purposes of this review, differences in the population-inclusivity of a school were treated as an intervention in their own right. This means that we did not seek any studies which explored the other actions a school might take to maximise student outcomes (such as mentoring schemes, teaching assistant support or particular types of grouping practice). However, we report such actions where studies central to our concerns present evidence in relation to them since they have clear implications for practice.
- This review is restricted to the phases of compulsory schooling. Although the impact of population inclusivity may well be relevant to the wider educational context, the policy and practice frameworks in areas such as further and higher education, vocational training and lifelong learning are so different from those in compulsory schooling that an all-encompassing review would be an unmanageably large task.

Outcomes

- In terms of outcomes, our interest was in the full range of educational outcomes – academic, personal and social.
- However, we defined the term ‘student outcome’ quite narrowly as a change in the capabilities of students. These capabilities may be academic (e.g. increased/reduced knowledge and skills in a curriculum area) or social and

personal (e.g. increased/reduced self-esteem or ability to sustain friendships). We differentiated outcomes in this sense from 'impacts' which are effects which may or may not result in a change in capabilities. For instance, students may enjoy being in an inclusive classroom, and friendships between students with and without SEN may blossom. However, this does not *in itself* guarantee that any change in student capabilities (i.e. outcomes) will result.

- We have included studies which do not simply report the impact of some overall measure of inclusivity but which also explore the way in which the presence of different individuals and groups within the population impacts on outcomes.
- The current policy presumption in England is that inclusion is the preferred option for students with SEN. Whilst it is important to find and test the evidence for this presumption, that is not the focus for the current review and we therefore excluded studies which report outcomes *only* for such students. The key issue we addressed is whether 'including' some students has an effect on outcomes for other students. We were particularly interested, therefore, in searching for and including studies which report outcomes for students other than those who are 'included'.

Design

- We decided to search for studies which have robust means of identifying student outcomes and attributing these outcomes to schools' levels of inclusivity. It is not enough for studies simply to report outcomes in inclusive schools without some convincing means of attribution since there may be many other factors (teacher effectiveness, quality of leadership, levels of resourcing and so on) producing those outcomes. We anticipated that this would involve comparing outcomes in more- and less-inclusive schools or tracking changes in outcomes against variations in population-inclusivity. In effect, this means that we sought intervention studies, where some variation or difference in population-inclusivity is the intervention.
- Likewise, we were interested in studies where the report of outcomes has some prima facie trustworthiness. While we did not restrict ourselves to studies which use quantitative measures, we decided only to include studies which present direct rather than reported evidence of outcomes. We excluded studies which rely *exclusively* on school staff's perceptions of pupil outcomes. Not only is it inherently difficult for such staff to identify student outcomes, but they are likely to have vested interests in claiming outcomes of particular kinds (e.g. claiming that their own practices and principles generate positive outcomes, or that an imposed policy of inclusion is damaging their students). We, of course, included studies where staff perceptions formed part of a wider range of evidence and deal with the question of trustworthiness in these cases by applying our quality criteria. We also included studies which report pupils' own perceptions, corroborated or otherwise, since (assuming due methodological precautions have been taken) those perceptions constitute direct evidence of personal outcomes such as, for example, self-esteem and engagement with learning.
- Our first review (Dyson, 2002) indicated that there is a major weakness in the inclusive education research literature in dealing with student outcomes. In much of that literature, the question of outcomes is not addressed, or positive outcomes are asserted on the basis of minimal evidence, or evidence of

dubious trustworthiness (such as the uncorroborated perceptions of teachers). We, therefore, demanded a high quality of evidence in relation to outcomes, even if this meant excluding studies which might otherwise be relevant to our question. This was to avoid the danger of a systematic review simply replicating the existing biases within the field.

2.2.2 Identification of potential studies: search strategy

A search strategy was designed so as to allow optimal use of the time and money resources available. Initially, we identified the following as sources of research literature that were subsequently explored and searched. These sources were compiled using the expertise and resources available through Manchester University and its John Rylands Library. Work in this area has also been supported by the EPPI-Centre.

Firstly, citations in key articles (identified through research experience in the broader and also more specific areas of education, special education and inclusion) were used to formulate key search terms for use with electronic databases. Searches through electronic databases constituted the main strand of our search strategy. Searches in databases covered journal articles, books and book chapters, conference papers and proceedings, theses, dissertations and reports. A search strategy was developed for this part of the process. It involved the identification and combination of sets of search terms by which literature identified according to the protocol as relevant to the review has been classified within individual databases. Where databases have no such classificatory system such as 'subject headings', 'keywords' or 'descriptors', a set of 'free text' terms was devised and agreed between the Review Group members. Therefore, the final search strategy was formulated following consultation with members of the Advisory Group and members of staff at the EPPI-Centre. This set was subsequently tested out in a particular database and new terms were added or existing terms were altered or removed. When the Review Group was happy that, by using the devised set of terms (individual terms and their combination), in theory all the available items stored in the databases could be identified, the search was run in all relevant databases. Particular attention was paid to ensure that the devised set of terms would allow sufficient identification of relevant items without generating too many irrelevant hits.

- (a) *Inclusion and its synonyms and related terms*: Use of terms such as inclusion, integration, mainstreaming, and so on, allowed us to reflect the use of different terms to describe the same or similar practices across different countries and cultures. In addition, a variety of terms used in the place of inclusion allowed us to capture research on this area from different historical periods over the last few decades.
- (b) *Education and its synonyms and related terms*: Initial searches run without the use of terms relating to 'education', 'school', etc. were found to generate a large number of hits relating to integration/inclusion in the community, integration of elderly in different settings and so on. Therefore terms relating to education were inserted in the devised set of terms to clarify the search further by allowing us to discard hits relating to community integration as well as to restrict the search to the school age population.

- (c) *Outcomes and its synonyms and related terms*: The use of terms such as 'outcomes' and 'impact' covered academic achievement as well as social outcomes.
- (d) *Population and its synonyms and related terms*: Finally, terms relating to the population of people with SEN were added in order to refine the search further by keeping only those terms that relate to 'inclusion of people with SEN' and exclude results of hits relating to inclusion of pupils with English as an additional language, the inclusion of pupils from ethnic minorities, etc.

An example of the search strategy used for the ERIC database is found in Appendix 2.2.

Since substantial searching had already taken place at the initial phases of the review, we used a pearl-growing technique by locating known publications in databases and identifying the terms under which they have been indexed. We also used the bibliographies in known publications to identify others which were not found through the databases.

A few other sources were useful in the identification of potentially relevant studies. Those included personal contacts within the Review and Advisory Groups who in some cases were useful in identifying papers or providing directions for further searches. It was also particularly useful that some members of the Review Group had been involved in the DfES commissioned study on the relationship between inclusion and pupil achievement referred to earlier. These members had already undertaken a small literature review in the area and could provide the Review Group with references or other useful sources. Finally, journals which yielded a number of significant articles were handsearched and studies selected according to inclusion criteria as outlined in the protocol.

Searches were also carried out of websites, suggested by members of the review and Advisory Groups, of national and international organisations which commission and publish research in the field of inclusive education. Those did not directly identify studies that were finally included in the review. However, they were often useful in locating reviews, reports and other material that directed us to locating other useful studies. Also, there were several requests by email to known authors in the area to help us locate useful research, but when authors responded to our request, this strategy did not produce any results that had not been already located through the electronic database search.

2.2.3 Screening studies: applying inclusion and exclusion criteria

The majority of papers were identified through electronic searching. The citation details of all papers identified through electronic database searches were downloaded and saved in word files in order to be screened. Most papers were excluded using a screening process by which titles and/or abstracts were screened through the application of inclusion/exclusion criteria. This was achieved by the application of the inclusion criteria in a consecutive way: firstly, criterion 1 was applied; if this was met, the paper was checked against criterion 2; and so on. For many of the papers, it was clear from reading the title that it would not match the criteria (usually the criterion 1 was not met after reading the title). Many other papers required a careful reading of the abstract. If, after reading the abstract there was still considerable doubt on whether the paper was a potential

'include' or not, an effort was made to obtain the full hard copy (through the library, download or ordering). All potential 'includes' were saved into a different file and the full report was obtained (or it was attempted to be obtained) for the studies that appeared to meet all the criteria for inclusion. The inclusion and exclusion criteria were re-applied to the full reports and those that did not meet all of them were subsequently excluded. Details of the screening process and the results are presented in Figure 3.1.

The citation details, abstract and other notes were imported and saved in two EndNote files. One file contained all 'includes' that were data-extracted while the other file accommodated the details of all obtained papers that it was decided later to exclude from the data-extraction process and the in-depth analysis.

2.2.4 Characterising included studies

All studies meeting the inclusion criteria were keyworded using the EPPI-Centre (2002a) Core Keywording Strategy, version 0.9.7 (see Appendix 2.4). Additional keywords, which were specific to the context of the review, were added to those of the EPPI-Centre (see Appendix 2.4). These concerned the severity of special educational needs of included pupils, the exact type of SEN, and the area that the impact is reported on (academic or social outcomes) and for which group of pupils the outcomes are reported (pupils with SEN, pupils without SEN and/or both groups).

In coding the exact type of SEN, the Review Group decided that it would be best to adhere to the four broad categories used by the Code of Practice on Special Educational Needs DfES (2001): (a) cognition and learning, (b) communication and interaction, (c) behavioural, emotional and social development, and (d) sensory and/or physical needs. Since nearly all the studies were American, it was sometimes necessary to translate the US terminology to the above categories. Therefore, terms such as 'mental retardation', 'educable mental retardation', etc. were replaced by needs in 'cognition and learning' and so on, although the original terms were retained in 'details of special educational needs' in the EPPI-Centre data-extraction tool.

An effort was made to record the pupils' main special educational needs and not all additional needs that they may have had. For example, when a pupil had severe learning difficulties as the main registered difficulty but had to use hearing aids as well, this child was only coded under the 'cognition and learning' category rather than, in addition, under the 'sensory and/or physical needs' category. However, in many cases pupils with EBD had also moderate to severe learning difficulties and in these cases, both types of SEN were recorded. If pupils were thought to have EBD as their primary difficulty and only mild secondary learning difficulties, the 'behavioural, emotional and social development' category was used as the sole code.

2.2.5 Identifying and describing studies: quality-assurance process

In order to assure the quality of process, the following steps were taken:

- The application of the inclusion and exclusion criteria for the initially identified studies was conducted by pairs of the Review Group members, EPPI-Centre,

and/or the advisory team in samples of studies for each database (e.g. the first 50 studies identified in PsycINFO database or the first 10–20 in others in ERIC, etc).

- For the potentially included studies, the application of the inclusion and exclusion criteria and the keywording was conducted by pairs of the Review Group members, EPPI-Centre and/or the advisory team, working first independently and then comparing their decisions and coming to a consensus. Staff at the EPPI-Centre keyworded a sample of 14 reports to ensure consistency of the process. This process was particularly helpful in clarifying certain aspects of the review for all groups involved and for adopting a shared understanding of the studies' concepts and how these related to our specific review.

2.3 In-depth review

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

The inclusion/exclusion criteria were applied to all studies and, when there was uncertainty about the inclusion of a study, the criteria were refined and clarified further. This resulted in a set of very specific criteria and consequently the mapping exercise led to the identification of only a small number of studies. All studies in the descriptive map were also included in the in-depth review.

2.3.2 Detailed description of studies in the in-depth review

All studies included in the final report were cross-examined using a set of data-extraction questions. These are standard EPPI-Centre questions, and comprise 14 sections of questions ranging from administrative details, to methods, results and quality of study. These questions were complemented by the Review Group's specific questions.

The Review Group's specific questions concern areas of reported outcomes (academic or social); type of included students' SEN; the comparative element of the study; external standardised data provided in the study; the key impact of inclusion (negative or positive, etc.); the evidence of inclusivity; and, finally, the outcomes of inclusion (see Appendix 2.4).

To facilitate the process of coding the outcomes of inclusion, the following guidelines were adopted: the Review Group attempted to code the overall finding where the findings pointed in a single direction; where, however, the findings were not pointing in a single direction, it was decided that they would be described as mixed.

The study could be coded as either positive, negative and/or neutral outcomes rather than mixed if

- it reports two or more different types or areas of outcomes – for example, positive academic outcomes but negative social outcomes;

- different sub groups of pupils are reported to respond differently to the inclusion of pupils with SEN.

On the other hand, a study was categorised as mixed rather than positive, negative and/or neutral:

- if it reported different outcomes in the same general outcome area (e.g. some positive outcomes for specific mathematics skills but neutral outcomes for other mathematics skills).

The reason for this differentiation between mixed results and either positive or negative or neutral was that it was felt that results should be presented in this review as clearly and specifically as possible, whether they were positive, negative or neutral. On the other hand, to conclude that, for example, 15 out of 26 studies have mixed results would not be useful for either policy or practice implications.

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

An important part of the data-extraction procedure involved an appreciation for each study of the weight of evidence. This is a process by which reviewers are required to judge the study's evidence in relation to the study's question and, more importantly, in relation to the review's questions. This was a difficult exercise to carry out since there were only a few studies that directly focused on the impact of population inclusivity on pupils without SEN and therefore the Review Group decided to include studies (as long as they met the inclusion criteria) that either looked at this element indirectly (for example, by focusing on co-operative learning practices) or examined this element as part of other issues. Although the quality of these latter studies could be high for the purposes of the study itself, it would be weighed lower in relation to the purposes of this specific review.

The Review Group made use of the EPPI-Centre 'weight of evidence' tool. The two independent reviewers had to agree on the weight of evidence for each study in order to make the final judgment trustworthy. There are three key elements to this judgment: trustworthiness (or soundness of studies), appropriateness of design and analysis, and relevance of focus. More specifically:

- Soundness of the study (internal methodological coherence), based upon the study only (weight of evidence A)
- Appropriateness of research design and analysis used for answering the specific systematic review question (weight of evidence B)
- Relevance of the study topic focus (from the sample, measures, scenario, or other indicator of the focus of this specific systematic review) (weight of evidence C)
- An overall weight (D) taking into account the quality of evidence (A), the appropriateness of design (B) and the relevance of focus (C) (weight of evidence D)

The overall weight of evidence (D) was taken as an average of the soundness, appropriateness and relevance of the study.

2.3.4 Synthesis of evidence

The data were synthesised according to a framework that was developed during the data-extraction procedure. Initially, studies were examined and placed provisionally into a number of groups that it was thought would create meaningful clusters in which the results of the synthesis could be of some value. A repeated application of this exercise led to the development of the final conceptual framework in which the nature of the impact of population inclusivity is a key concept. The impact of population inclusivity was seen as important in two general key areas: (i) academic performance of students and (ii) social, emotional or other impact. Therefore, the main conceptual framework includes academic and social outcomes for students without SEN when pupils with SEN are included in their classrooms or schools. These outcomes can in general be positive, negative or neutral.

In addition, it was thought that outcomes for students might be different, depending upon the nature of SEN of the included students. This assumption was mainly based on some anecdotal evidence which suggested that the impact of inclusion of pupils with EBD might be more severe and tend towards the negative side for their non-disabled peers. Furthermore, such a distinction was strongly supported by members of the Review Group since it was felt that conclusions that were related to the nature of SEN group that was included could be considered in planning future policies on inclusion.

The Review Group does not claim that this is the only meaningful and useful way to synthesise the data and report the results. However, the framework adopted in this study views the impact of population inclusivity according to the following categories:

- area of key impact (academic or social outcomes)
- type of school (primary versus secondary schools)
- nature of SEN (adopting four main categories)
 - cognition and learning
 - behavioural, emotional and social development
 - sensory and/or physical needs
 - communication and interaction

Sometimes the decision to code a study under a specific category was problematic since, more often than not, studies were not carried out within a similar framework and, therefore, did not easily fall under one or another 'category'. It was therefore important for the Review Group to keep an open mind and remain flexible in applying this framework.

2.3.5 In-depth review: quality-assurance process

The quality-assurance process involved the following stages:

- Five studies were used as examples for moderating the tools. A member of the Review Group and a member of the EPPI-Centre link staff independently tried out both generic and review-specific data-extraction tools. The results were compared until a mutually agreed, moderated final data-extraction was obtained. Again, this process contributed further to the training among team

members of the Review Group. This process helped in checking the tools and shared interpretations.

Each of the remaining studies was data-extracted by pairs of reviewers in the Review Group. They followed the same approaches as those used in the reconciliation of the data-extraction products.

3. IDENTIFYING AND DESCRIBING STUDIES: RESULTS

This section describes the results for the first phase of the review. Eventually 26 studies were used to in the final review.

3.1 Studies included from searching and screening

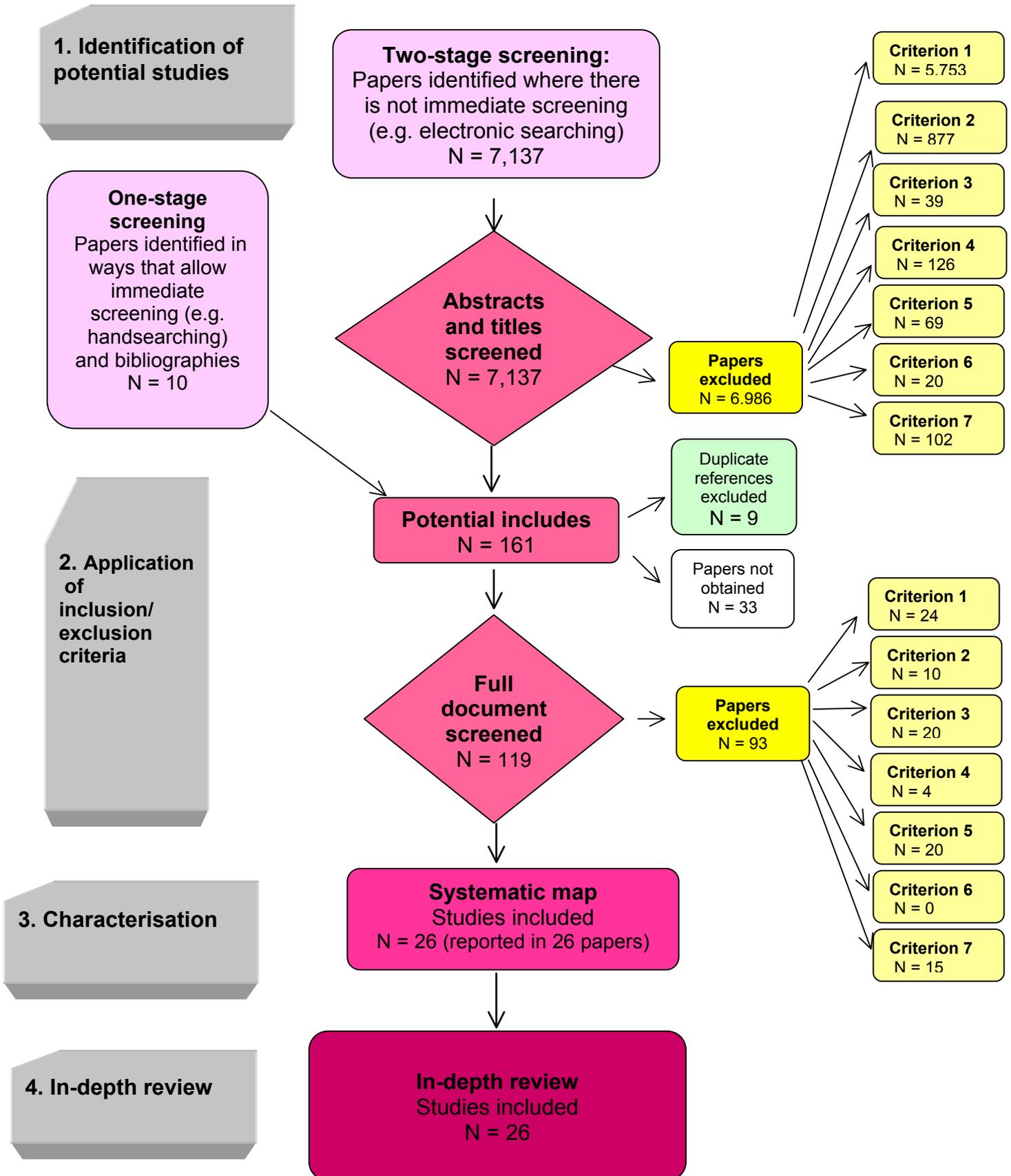
Figure 3.1 shows the process of filtering of publications from searching to mapping to synthesis. As can be seen, at the first stage of identification of potential studies 7,137 papers were initially identified through searches of electronic databases. Having screened all those papers' titles and/or abstracts, 161 papers were marked as 'potential includes'. Of the 6,986 excluded at this stage, the majority (82%) were excluded because they did not describe the impact of population inclusivity (exclusion criterion 1), 13% because they described the impact of inclusion on pupils with SEN and did not refer to the general population without SEN (exclusion criterion 2), 0.6% because they did not make any references to the impact on personal, social and/or academic outcomes (exclusion criterion 3), 2% because they referred to pre-school or post-compulsory provision (exclusion criterion 4), 1% because they referred to papers that described reviews rather than the results of empirical research (exclusion criterion 5), 0.3% because they were not written in English (exclusion criterion 6), and 1.5% because they did not provide a robust evidence of measurement (exclusion criterion 7).

Some reports could not be obtained. Of these, the majority were theses, mostly from the USA and Australia. Initially, a few of these theses were ordered through inter-library loan but either the British Library could not trace them or it was too expensive to obtain a copy. There were also a number of reports that could not be located, despite efforts to obtain them through inter-library loans.

Key to Figure 3.1

Criterion	
1	It does not describe the impact of population inclusivity.
2	It does not refer to the general population without SEN.
3	It does not report impact on personal, social or academic outcomes.
4	It focuses on pre-school or post-compulsory education.
5	It does not refer to empirical research.
6	It is not written in English.
7	It does not provide robust evidence of measurement.

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



3.2 Characteristics of the included studies

Of the studies that were initially marked as potentially included, the majority were excluded once the full text report had been read because either they did not concern the impact of population inclusivity, or they were not empirical pieces of research, or they did not include a robust evidence of measurement or the impact was not on personal, social and/or academic outcomes. Therefore, 26 studies were eventually included in the systematic map.

Of the 26 studies, 16 also looked at the impact of inclusion on pupils with SEN. For some of them, the impact of inclusion on pupils with SEN was the primary focus of the study and they only examined, in part, the impact on students without SEN. Some were case studies that evaluated a specific programme in a school or area, while others adopted experimental designs. The studies ranged over a period of 23 years, the earliest study being in 1982 and the most recent one in 2003. More specifically, there were five studies before 1990, 16 between 1990 and 1999, and five studies from 2000 onwards. These studies also vary widely in the way in which they are reported.

Study type

Table 3.1: Study type of included studies (N = 26)

Type of study*	Number of studies
Evaluation: naturally occurring	15
Evaluation: researcher-manipulated	11

* Codes for 26 studies; categories are mutually exclusive.

All studies were evaluations of practices/programmes (Table 3.1). Most of the evaluation studies investigated naturally occurring phenomena (N = 15) and, of those that were researcher-manipulated (N = 11), three were randomised controlled trials (RCT) and eight were non-randomised controlled trials.

Country

The majority of these studies were American (N = 22) and there were four studies from Australia, Canada and Ireland (Table 3.2). There were no UK studies that looked at the impact of population inclusivity. It has to be mentioned here that, while mapping the studies in the initial stages of identification, several studies were found where it was thought that they matched the majority of inclusion criteria, but were then excluded as they were not written in English (for example, a few German studies).

Table 3.2: Countries where studies were undertaken (N = 26*)

Country	Number of studies
USA	22
Canada	1
Australia	2
Ireland	1

* Codes for 26 studies; categories are mutually exclusive.

School setting

The studies mainly focused on primary school populations (table 3.3). Two studies did not specify the school level of the children or their ages, so the age was taken as 'all age'. One study took place in a special education school setting where the inclusion experiment took place within a local special school's facilities.

Table 3.3: Educational setting of studies (N = 26)

Educational setting*	Number of studies
Primary school	21
Secondary school	3
Special needs school	1
Other	2

* Codes for 26 studies; categories are not mutually exclusive.

Special educational needs of included students

This review used the four DfES categories of special educational needs: (a) cognition and learning, (b) sensory and/or physical needs, (c) communication and interaction and (d) behavioural, emotional and social development (see Table 3.4). Most of the studies reported inclusion of pupils with needs in cognition and learning, either as their primary or secondary need. However, in the majority of cases, pupils included in the mainstream had a variety of SEN, making it more difficult to measure the effect of the presence of children with a specific type of SEN on the pupils' outcomes.

Table 3.4: Special educational needs of the pupils included in the studies (N = 26)

What type of special educational needs to pupils have?*	Number of studies
Cognition and learning	24
Sensory and/or physical needs	11
Communication and interaction	9
Behavioural, emotional and social development	12
Not specified	0

*Codes are not mutually exclusive.

Focus of the outcomes

The majority of studies look at academic outcomes (N = 21), 13 studies look at social outcomes and four consider other outcomes, such as attendance.

Table 3.5: Focus of the outcomes (N = 26 studies)

In which of the following areas does the study report outcomes for pupils without SEN?*	
Academic outcomes	21
Social and personal (e.g. behavioural)	13
Other (e.g. attendance)	4

*Codes are not mutually exclusive.

3.3 Identifying and describing studies: quality-assurance results

During screening, a sample of 50 titles and abstracts was also considered by a member of the EPPI-Centre. Advisory Group members also helped to moderate the screening process, looking at samples of studies identified on each database. The studies included in the review were all keyworded by two reviewers; in 14 cases, this was a member of the EPPI-Centre. There were a few disagreements at these stages that were solved through moderation.

4. IN-DEPTH REVIEW: RESULTS

This chapter reviews the studies that were selected for the synthesis of evidence. It opens with a presentation of the studies' weight of evidence in terms of trustworthiness, appropriateness, relevance and overall weight of evidence. The chapter goes on to present all studies organised according to the type of outcomes they present (academic, social or others), the overall findings (neutral, positive, negative or mixed), and the overall weight of evidence. The chapter concludes with some overall comments on the key findings from the review.

All 26 studies included in the systematic map were selected to be included in the final data extraction and in the synthesis of evidence. The table in Appendix 3 gives details of these studies according to the review-specific questions.

The complete data-extraction records for each study can be found on the EPPI-Centre website on the review home page (<http://eppi.ioe.ac.uk/reel>). These show how each study was coded using the data-extraction tools, the main parts of which are common to all EPPI-Centre reviews of educational research. There are also comprehensive details of how the methodological processes behind each study can be explored, with more information than is feasible or desirable to include in this review itself. One can also find at the same website details of the specific review questions which are related to this review only.

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

The sixth inclusion criterion required that the studies would report robust evidence of the impact of the intervention by either providing longitudinal evidence of school(s)/classroom(s), or by comparison with similar but less inclusive school(s)/classroom(s) or by comparison between different conditions within the same school(s) or by some other equally robust means. The majority of studies provide robust evidence through comparisons, either between groups in the same or different schools (N = 11), or comparisons between classrooms within the same school (N = 7), or comparisons with similar but less inclusive schools (N = 2), or finally, comparisons between different conditions in the same groups (N = 4). Four studies report results based on data collected longitudinally while two studies provide some other robust evidence (Table 4.1).

Table 4.1: Comparative element (N = 26)

What is the comparative element in the study?*	Number of studies
Comparison with other similar but less inclusive school	2
Comparison with other classroom within the same school	7
Comparison between groups in the same or different schools	11
Comparison between different conditions in the same groups – in the same or different schools	4
Through a longitudinal study in one or more schools/classrooms	4
Other	2

* Codes are not mutually exclusive.

The data on which the results of the impact of inclusion were based varied widely between studies, and often studies present more than one type of standardised data. These are presented in detail in Table 4.2. Overall, it is clear from this table that researchers have no clear preference for using a particular type of standardised data. Rather, they use a variety of data, usually justifying their choice according to the study's specific aims and overall design. School/teacher reports, pupils' reports, systematic observations by teachers and data from national tests appear to be slightly more preferred compared with the other methods of data collection.

Table 4.2: Outcome data (N = 26)

What external or standardised data are provided as evidence of the outcomes to pupils without SEN?*	Number of studies
National tests	5
Teacher assessments	3
Group tests	2
Teacher rating scales	1
Personality tests	0
Other relevant attribute measures	3
Systematic observations (by researcher)	6
Systematic observations (by teachers/school staff)	1
School/teacher reports	6
Pupil interview data	3
Pupils' reports through questionnaires	6
Sociometric data	3
Ethnographic accounts	1
Other	11

* Codes are not mutually exclusive.

The studies also varied in terms of evidence they provide for population inclusivity. Sixteen studies report some kind of ratio: for example, the ratio of pupils with SEN to pupils without SEN in a class or in whole school was, say, 10%, or a class of, say, 25 students included three students with SEN. Fourteen studies specify the amount of time in the school day or the time in a specific curriculum area that was spent by the SEN pupils who were included. Thirteen studies specify that the needs of students were 'rather severe' while some of the remainder are somewhat vague in describing the severity of special needs of the included students.

Table 4.3: Evidence of inclusivity (N = 26)

What evidence of inclusivity does the study provide?*	Number of studies
In terms of number of SEN	16
per school/class (ratio e.g. >10%)	
In terms of severity of needs (e.g. severe needs of pupils)	13
In terms of hours of inclusion (total inclusion vs. partial inclusion, pull out programmes, resource rooms, etc.)	14
Other	5

* Codes are not mutually exclusive.

Most of the studies report overall neutral outcomes for pupils without SEN when their peers with SEN were included in their school, classroom or group. Nine studies report clear positive overall outcomes, while six report at least some negative outcomes. Finally, four studies report a mixture of outcomes.

Table 4.4: Key findings (N = 26)

What are the key findings on outcomes for pupils without SEN? *	Number of studies
Positive	9
Negative	6
Neutral (no change)	21
Mixed (some areas have improved while some have deteriorated)	4

* Codes are not mutually exclusive.

Weight of evidence

The 26 studies included in the synthesis were weighted according to the procedures outlined in section 2.3.3 which are included in the EPPI-Centre guidelines for all reviews. The results of this process are presented in Table 4.5.

Table 4.5: Results of assessment of weight of evidence for each study

Study	A Trustworthy	B Appropriate	C Relevant	D Overall weight
Affleck <i>et al.</i> (1988)	Low	Low	Low	Low
Bear <i>et al.</i> (1991)	High	High	Medium	High
Beuter (1984)	Medium	Medium	Medium	Medium
Block and Zeman (1996)	Medium	Medium	High	Medium
Brown (1982)	Medium	Medium	Low	Medium
Cawley <i>et al.</i> (2002)	Medium	Low	Medium	Medium
Daniel and King (1997)	High	Medium	Medium	Medium
Helmstetter <i>et al.</i> (1994)	Medium	Low	Low	Low
Hepler (1998)	Medium	Low	Low	Low
Hillen <i>et al.</i> (1992)	Medium	Low	Low	Low
Huber <i>et al.</i> (2001)	Medium	Medium	Medium	Medium
Hunt <i>et al.</i> (1994)	Medium	Low	Low	Low
Lundeen and Lundeen (1993)	Medium	Low	Low	Low
Mastropieri <i>et al.</i> (1998)	Medium	Low	Medium	Medium
McDonnell <i>et al.</i> (2003)	High	Medium	High	High
Obrusnikova <i>et al.</i> (2003)	Medium	Medium	Medium	Medium
Rankin <i>et al.</i> (1999)	Medium	Medium	Medium	Medium
Rarick and Beuter (1985)	Low	Low	Medium	Low

Study	A Trustworthy	B Appropriate	C Relevant	D Overall weight
Saint-Laurent <i>et al.</i> (1998)	Medium	Low	Low	Low
Sasso and Rude (1988)	Medium	Medium	High	Medium
Sharpe <i>et al.</i> (1994)	Medium	Medium	High	Medium
Shevlin and O'Moore (2000)	Low	Medium	Low	Low
Shinn <i>et al.</i> (1997)	High	Medium	Low	Medium
Stevens and Slavin (1995)	Medium	Medium	Medium	Medium
Tapasak and Walther-Thomas (1999)	Medium	Low	Medium	Medium
Willrodt and Claybrook (1995)	Low	Low	Medium	Low

As the table shows, most of the studies received a medium overall weight of evidence (N = 14 studies), 10 studies were weighted as low and only two studies were of high weight. The main justification for a low weight instead of medium was that the primary focus of studies was not the impact of inclusivity on pupils without SEN although this was partially examined, or that the studies had a confusing or vaguely reported methodology and process, or that the results were not convincing enough due to many methodological problems (see section 4.4 for more discussion). It has to be noted here that studies which directly focused on the impact of inclusivity on pupils without SEN and that could exhibit robust evidence on this were scarce. The main conclusions of this report are primarily based on studies that are of medium to high overall weight of evidence. However, studies of low weight are also included since they provide supporting evidence and contribute to the overall conclusion. Moreover, in the majority of cases, studies of low weight of evidence did not differ greatly in their conclusions from the medium to high weight studies.

4.2 Synthesis of evidence

Section 4.2.1 and 4.2.2 discuss the impact of inclusion of students with special educational needs in terms of academic and social outcomes. The majority of studies looked at academic outcomes (N = 21), 13 studies report social outcomes and 4 considered other outcomes (such as attendance, etc.). These have been described under the social outcomes. Tables 4.6 and 4.7 show the studies under each type of outcomes (neutral, positive, negative and mixed) and across the different weights of evidence (from high to low). In both sections, studies are presented in order of neutral, positive, negative and mixed findings.

It has to be made clear that the categorisation of studies into positive, negative, neutral or mixed was not always a straightforward process. Many of the studies reported the impact of inclusion in more than a single strand of outcomes. For example, some studies report academic outcomes in more than one curriculum area (often in language, reading and mathematics), or in more than one aspect of the same curriculum areas (e.g. long jump and passing in basketball in Physical

Education (PE)), or more than one measure for assessing a specific skill (e.g. test to assess knowledge in volleyball and observations in the actual practical skills).

Overall, an attempt was made to categorise studies under positive, negative, neutral or mixed outcomes through accepting the authors' own conclusions. However, in some cases it was felt that the authors' conclusions were not sufficiently justified by the presented findings and, as such, the Review Group members felt that studies like this should be categorised according to what interpretations the Review Group made from the findings. Fortunately, this was only the case in a couple of studies. For example, Cawley *et al.* (2002) conclude in favour of inclusion for pupils with SEN in the science classroom. However, the tables presented in the findings section (which are not accompanied by text) present results that the Review Group interpreted as negative.

Finally, and in line with the above discussion, an attempt was made to take the researchers' interpretation of their findings as reliable and accurate, and to categorise the studies according to the authors' overall reports on direction of outcomes. When the reported findings were based on statistical tests, the Review Group members took into consideration the authors' reports on statistical significance of differences or correlations. When there were problems with the reporting of statistical significance in tests (e.g. statistical significance reports were missing or there was some other problem with reporting), a comment is made in the description of the study's findings (e.g. Cawley *et al.*, 2002).

4.2.1 Academic outcomes

Table 4.6: Studies included in the in-depth review by cluster – academic outcomes

A: Neutral outcomes	B: Positive outcomes	C: Negative outcomes	D: Mixed outcomes
<i>High overall weight</i>	<i>High overall weight</i>	<i>High Overall weight</i>	<i>High overall weight</i>
McDonnell <i>et al.</i> (2003)			
<i>Medium overall weight</i>	<i>Medium overall weight</i>	<i>Medium overall weight</i>	<i>Medium overall weight</i>
Beuter (1984) Block and Zeman (1996) Huber <i>et al.</i> (2001) Obrusnikova <i>et al.</i> (2003) Rankin <i>et al.</i> (1999) Sharpe <i>et al.</i> (1994) Tapasak and Walther-Thomas (1999)	Mastropieri <i>et al.</i> (1998) Stevens and Slavin (1995) Shinn <i>et al.</i> (1997)	Brown (1982) Cawley <i>et al.</i> (2002)	Huber <i>et al.</i> (2001) Daniel and King (1997)
<i>Low overall weight</i>	<i>Low overall weight</i>	<i>Low overall weight</i>	<i>Low overall weight</i>
Affleck <i>et al.</i> (1988) Hillen <i>et al.</i> (1992) Hunt <i>et al.</i> (1994) Lundeen and Lundeen (1993) Rarick and Beuter (1985) Willrodt and Claybrook (1995)	Saint-Laurent <i>et al.</i> (1998)		

A: Academic outcomes – neutral

Most studies presenting neutral outcomes on the academic impact to students without difficulties when students with SEN were included focused on measures on reading and/or writing or language in general and mathematics (Affleck *et al.*, 1988; Huber *et al.*, 2001; Hunt *et al.*, 1994; McDonnell *et al.*, 2003; Rankin *et al.*, 1999; Sharpe *et al.*, 1994). Several other studies measured performance in elements relating to physical education (Block and Zeman, 1996; Obrusnikova *et al.*, 2003; Rarick and Beuter, 1985). Although, strictly speaking, PE is not considered an academic subject, the respective studies do however include measures taken from a curriculum subject, often including performance on concepts of PE and tests that go beyond motor skills.

High weight of evidence

The only one study that received overall ‘high’ weight of evidence in academic outcomes involved measurements in reading/language art and mathematics. **McDonnell *et al.* (2003)** examined the impact of inclusive educational programmes on the achievement of students with special needs and on their peers without SEN. In order to examine whether the presence of pupils with learning difficulties would have a negative impact on the educational achievement of pupils without disabilities, the researchers compared the academic performance of two groups of students from five elementary schools. Group 1 (experimental) comprised 324 students without disabilities who were enrolled in inclusive classes with students with SEN while Group 2 (control) comprised 221 students without disabilities who were enrolled in classes that did not include students with SEN. The experimental group classrooms included between one and seven pupils with difficulties in learning and cognition, ‘developmental delays’ and one pupil with autistic spectrum disorders. McDonnell *et al.* employed a quasi-experimental design and took measures on mandated state-level criterion referenced tests in reading/language arts and mathematics.

The results showed no significant differences in performance in reading/language arts and in mathematics subtests between students enrolled in inclusive and those enrolled in comparison classes (without students with SEN). Also, within-school comparisons of inclusion across the five schools revealed no statistically significant differences. Consistent with the findings of a previous study described below by Sharpe *et al.* (1994), McDonnell *et al.* (2003) conclude that the presence of students with SEN does not negatively affect the learning of students without disabilities.

Medium weight of evidence

Sharpe *et al.*’s (1994) study also involved (similarly to McDonnell *et al.*’s (2003) study) a comparison of achievement scores in reading, language arts and mathematics between two groups, an experimental group including students with SEN and a control group without students with SEN. However, their study had been undertaken almost a decade before McDonnell *et al.*’s (2003) research. Sharpe *et al.* conducted a study specifically to investigate the impact of inclusive school environments on the academic performance of general education students (GESs). Comparisons were made between two groups: (a) the inclusion group, consisting of 35 GESs educated in an inclusive environment, and (b) the comparison group, consisting of 108 GESs educated in a general education classroom without the presence of pupils with SEN. The researchers measured the improvement in academic performance of these two groups by comparing

group achievement test scores between pre- and post-intervention. The inclusive intervention consisted of mainstreaming five students with severe to profound learning difficulties, some of whom had significant challenges and one had severe emotional disorders. Achievement scores were taken in the areas of reading, language arts and mathematics. This was a post hoc study and therefore the measures included archival data collected from student files two years after the pilot inclusion programme.

Overall, the results of the study revealed no significant differences between the two groups on any of the academic measures. More specifically, the researchers examined the performance of students and found that the groups were generally homogeneous in terms that there were no initial differences (pre-intervention) between the two groups. Post-test, overall, there were no statistically significant evidence of performance differences between the two groups in the basic skill areas of reading, language arts and mathematics. Consistent with these results were the teachers' ratings on students' report cards, showing no overall differences between the two groups. An exception was found in the area of reading in teacher ratings, but this should be treated with caution due to the small cell numbers in expected frequencies.

Huber et al. (2001) also examined the impact of inclusion on classroom achievement in more depth than other studies in this review by investigating the differential impact of inclusion and inclusive practices on different ability groups of GESs: high, average and low achieving.

Achievement scores for 477 male and female GESs from first to fifth grade were sampled over three years. These pupils were educated in classrooms with different numbers of pupils with SEN, varying from 0 to 7, with the majority of classrooms having between 0 and 3 pupils with SEN. Therefore, during the two years of the implementation of inclusion and inclusive practices, students either received academic instruction with students with SEN or only with peers without SEN. Inclusive school practices comprised curricular changes and varying degrees of support, and it was hypothesised that they would have a differential effect on pupils at different achievement levels. Huber *et al.*'s argument was that 'by restructuring schools to support children with disabilities, the range of tolerance in the classroom will be moved downward, and would result in a classroom that is no longer meeting the needs of high achievers'. In addition, they were concerned that, if schools include a large numbers of pupils with SEN, it would place high demands on schools' time and resources, and therefore would make it too difficult for schools to meet the needs of individual students.

Measures included normal curve equivalent (NCE) scores on the total mathematics and reading sections of the Metropolitan Achievement Test (sixth edition) and the Stanford Achievement Test (eighth edition) (SAT) total mathematics NCE scores and total reading NCE scores.

The results revealed that student skill factors had a statistically significant effect on general education students' attainment (described below in mixed results). However, as far as the effect that the number of students with SEN had on scores for the class, it was found that there were no differences for reading, although there were some differences for mathematics. More specifically, when the reading incremental change score was the dependent variable, ANOVA results revealed that there were no significant differences among group means of students without SEN when classrooms contained different numbers of included students with disabilities. Huber *et al.* concluded that, in relation to the effect that the number of

pupils with SEN had on achievement, reading scores were not affected by the presence of their disabled peers.

Tapasak and Walther-Thomas (1999) set out to evaluate the first year of an inclusive education programme in a primary school. Their aim was to examine the social skills of students with SEN, their self-perceptions on academic and social competence (scholastic competence, social accountability, athletic competence, physical competence and behavioural contact), how frequently pupils with SEN are chosen as playmates and how positively they are viewed. Although the focus of their study is on pupils with SEN, the authors pose some open implicit questions about the ways students without disabilities are affected. They examined all students' perceived competence, report grades, and sociometric and peer nominations. The included students had a variety of difficulties, such as mild to moderate learning difficulties, emotional and behavioural difficulties, hearing impairments and other health impairments. There were between two and nine pupils with SEN in each inclusive classroom.

The study involved a total of 183 students in seven classrooms (with and without SEN), of whom 60 were primary-level students without SEN and 50 intermediate-level students without SEN in an inclusive-education programme. The researchers used the Play rating scale and the Pictorial Scale Perceived Competence to measure students' self-perception on cognitive, physical and academic competence as well as school records and report card grades.

The majority of the results concerned changes on the performance and social effects for pupils with SEN or comparisons between this group and the groups of pupils without SEN. The results for pupils without disabilities showed that, overall, pre- to post-intervention score increases for students without disabilities were not significant. On the Cognitive Competence subscale of the Pictorial Scale, the pre-post score increases for very young students without disabilities were not significant. On the Perceived Competence Scale for children in the third to fifth grades, there were no significant main effects for athletic competence, physical appearance, or behaviour subscales for the overall global scale scores. Teacher ratings also showed that there were no significant differences between autumn and spring (pre- and post-test measures) on academic competence for the groups (pupils with SEN and those without SEN).

This study did not arrive at any conclusions on the impact of inclusivity on pupils without SEN. Their more general conclusion refers to significant increases in self-perception of cognitive competence for primary students with and without disabilities. However, we could infer from their findings that including pupils with a variety of disabilities did not have an obvious negative impact on pupils' cognitive competence in primary schools. This conclusion should be treated with caution, taking the study's methodological limitations and its limited generalisability due to the small sample and the case study nature of the study. Also, as the authors acknowledge, there are several factors that limit the representativeness of these results: the population of included students; the fact that a number of students came from military backgrounds; and neither the teachers nor the students were randomly selected.

Rankin et al. (1999) investigated the effect on learning when including a student with disabilities in small group learning. The theoretical argument of the researchers was that, if the teacher spent consistent instructional time with a student with learning disability and the remaining students without SEN learned

the material, then this shows that the presence of the student with SEN does not adversely affect the classroom's learning.

The intervention took place in small groups of three or four students. Two conditions were designed: condition A involved teaching in a group of three GESs plus one student with difficulties in learning and cognition; and condition B involved teaching in the same group of three GESs plus another GES without special needs. Therefore, the independent variable was the group composition (with or without a pupil with SEN) and the dependent variable was the percent gain scores between a pre- and post test learning of five facts from the lessons taught by the teacher during the lesson. The five facts comprised discrete vocabulary facts or story comprehension questions given to 15 general education elementary school students (from kindergarten to second grade) who were taught in teacher directed small groups. The experiment involved an adapted alternating treatments design. A pre- and post-test on the facts was given to each student before and after the lesson. Improvement (or otherwise) was measured as the percent increase from pre- to post-test sessions.

The results showed that the students' gain scores were consistently higher on the post-test than the pre-test with no difference between the two conditions. More specifically, the mean gain scores were the same or higher in the groups with the student with disability during 92% of sessions, although the differences were not statistically significant. There was a small subset of sessions (22% of sessions) where the gain scores were higher in the condition with the students with SEN. Therefore, the general result was that the presence of the student with learning difficulties in teacher-directed learning in small groups did not have an adverse effect on the learning of vocabulary or story comprehension by this student's peers without SEN. **Rankin et al. (1999)** concluded that, when students with learning disabilities are included in either general education classrooms or in groups, they do not negatively affect the learning of GESs.

Two closely related studies have investigated the effects of the inclusion of pupils with SEN on their classroom peers without SEN in terms of academic achievement in PE. The first study by **Block and Zeman (1996)** measured the impact of including students with disabilities in regular physical education on their non-disabled peers. Two sixth-grade primary school classes were compared in an experimental design that involved (a) an experimental class with 28 GESs and three students with severe SEN, and (b) a control class 28 GESs. The measures comprised pre- and post-tests for basketball skills and attitudes (taken from a sub-sample of students) of regular education students towards their peers with disabilities. Academic performance was therefore measured as the improvement in specific skills in physical education, more precisely in basketball passing, shooting and dribbling. The three pupils with SEN were included in the specific PE class for three months prior to the study. The authors wanted to find out whether by including a student with severe disabilities they would necessarily have to compare the PE programme for students with and without disabilities, and how students really feel about having students with SEN in their PE classes. The results showed that there were no significant differences in mean gain scores between the two groups in passing or shooting. There were significant differences for dribbling, and the control group improved more than the experimental. However, the authors attribute this difference to the fact that the control group had initial lower dribbling skills and the gains made resulted in having about equal dribbling scores in post-test measures.

The authors conclude that, when students with severe disabilities are supported, they can be included in regular physical education environments without compromising the academic gains made by their peers without disabilities. They recommend caution in making generalisations from their results, mainly because the type of support provided in the school for students with SEN was somehow not realistic. More specifically, they had a fulltime adapted PE specialist for three students with SLD/PMLD plus two teaching assistants. They also recommend caution in interpreting these results due to the fact that this was a fairly controlled study in terms of subject and teacher selection, and they had limited measures to evaluate inclusion (only PE skill improvement and attitudes).

Following **Block and Zeman's (1996)** study, **Obrusnikova et al. (2003)** also looked at the effect of including a wheelchair-using pupil in a general physical education (GPE) programme on the remaining pupils in the setting. More specifically, they wanted to examine whether including a student who uses a wheelchair but is given no direct support in GPE compromises the motor skills and knowledge learning of elementary students without disabilities, what general education students think about having a disabled classmate, and how this impacts on their GPE programme. The main difference with the previous study of Block and Zeman was that, in the Obrusnikova *et al.* study, the student with physical disabilities was not provided with direct support. Based on the results of this previous study, Obrusnikova *et al.* hypothesised that there would be no significant differences between the experimental (fourth-grade inclusion) and the control (fifth-grade non-inclusion) class in motor skills (set, bump and serve) and knowledge acquisition of students, and between a pre-test and post-test in attitude scores. They therefore compared motor skills and knowledge acquisition between pupils in inclusive and non-inclusive classes respectively and attitude changes in pre-test and post-test measures.

The results showed that students in both classes improved in all measures (practical skills and knowledge) after the test period. However, results of ANCOVA showed that there was no statistically significant difference in the volleyball gains for any of the two classes. Also, there was no statistical significance in the gains between the inclusion and non-inclusion class. Obrusnikova *et al.* (2003) conclude that the inclusion of a physically disabled student who does not receive direct support from a specialised teacher/assistant is possible without compromising the learning of his/her classmates. Therefore, if proper teaching in GPE classes is provided, the inclusion of students with SEN does not negatively affect the learning of students without disabilities. The authors suggest that results showing a slightly higher improvement for the non-inclusive class (although not statistically significant) may be accounted for by the fact that students in the control classes were in fifth grade (compared with the fourth grade for inclusion class). Therefore their higher levels of performance and knowledge achieved may be attributable to maturation and experience as well as other human or environmental factors.

A methodologically different study, which also looked at, amongst other factors, motor performance, was the study by **Beuter (1984)**. Beuter carried out an etho-behavioural analysis of social behaviours of pupils with difficulties in learning and cognition, and of children without disabilities in an integrated educational setting. His aim was to evaluate objectively the social behaviours and interactions of students with SEN in integrated classrooms. However, as part of this, he used a motoric instructional programme and measured motor skills, which are here taken as a measure of academic performance.

The sample of pupils with SEN included 25 students with learning difficulties who were 'integrated' with third- and sixth-grade children in a motoric programme three times per week for over a six-month period. There were several groups, some including pupils with SEN and some not. For the experimental (integrated) groups the ratio of integration was 5–6 pupils with SEN to 20–24 pupils without disabilities. Measures in the motor tests included performance scores in three motor tests: the 20-yard sprint, standing long jump and softball throw. The results showed that the scores of performances in all three motor tests were in no way impaired by the integrated programme. Analysis of covariance of motor test scores of pupils without SEN in the third and sixth grade revealed no significant statistical differences between the experimental (inclusion) and control (non-inclusion) groups. Beuter concludes that integration in this experiment acted as a facilitator of motor performance.

Low weight of evidence

Three more studies, but with low weight of evidence, have focused on academic outcomes through measuring performance on the traditional subjects of language/reading and mathematics (Affleck *et al.*, 1988; Hunt *et al.*, 1994; Willrodt and Claybrook, 1995). **Affleck *et al.* (1988)** undertook a study to evaluate the efficacy of an integrated model called the integrated classroom model (ICM). The ICM is a programme designed to educate children with mild SEN in the same classrooms with regular education children for the entire school day. The paper reports the results of three of many studies that researchers undertook around ICM. For example, they compared student achievement data of ICM to achievement data in resource room programmes; cost-effectiveness between the two programmes; and pre- and post-test scores on the reading, mathematics and language subtests. As part of these studies, Affleck *et al.* report that the results of a study that compares the achievement of 39 pupils without SEN who were educated in ICM to the achievement of a randomly selected control group that did not include pupils with SEN. The integrated classrooms were composed of approximately one-third of pupils with SEN and two-thirds of average or above average regular education students. The target size of the classrooms was 24, eight of whom were SEN. Additional aide support was provided according to that proportion. The California Achievement Test Battery was used to collect achievement data. The results showed that there were no significant differences between the groups and Affleck *et al.* therefore conclude that regular education students' achievement is similar, regardless of whether they are educated in an ICM or in a regular classroom.

The Review Group had a number of concerns with this study and the way it reported. Overall, it appears to be incomplete since it reports on data presented in tables which are missing from the paper; the results are reported without any statistical evidence - no statistical significance differences are reported, the tests used are not named and there are no statistical significance levels. The original reports of these studies could not be located and it was difficult to judge the study in terms of methodology, procedures, results and general quality.

Willrodt and Claybrook (1995), in an unpublished dissertation, report the results of comparing mathematics and reading achievement among fifth-grade students between two suburban primary schools: one that utilised a traditional approach of pull-out special education classrooms and the other which had adopted an inclusion programme for pupils with SEN across all age ranges. Measures included the Texas Assessment of Academic Skills (TAAS) for both mathematics and reading. The TAAS scores were part of the mandatory assessments carried

out in the education district within which the schools were located. The results showed that there was no significant difference in the TAAS between the two schools, supporting a conclusion towards inclusion of pupils with SEN in schools without adversely affecting the performance of their peers without SEN. Willrodt and Claybrook's study has several shortcomings in relation to limited reporting of important aspects of the study, the sampling procedures and the lack of detail about the context of the 'pull-out school'.

Hunt *et al.* (1994) looked at the achievement of all students within the context of co-operative learning groups. Their main aim was to investigate the extent to which pupils with SEN would achieve in inclusion settings. As part of the study, the researchers also looked at the achievement of targeted academic skills by the students without disabilities. They used a pre-test/ post-test design to compare the achievement of targeted mathematics objectives of 10 pupils without SEN who participated in co-operative learning groups with the achievement of a control group of 10 pupils without SEN who were members of co-operative learning groups that did not include pupils with SEN. One pupil had severe intellectual and physical difficulties, one had multiple/severe disabilities, and one had autism and severe learning difficulties. They all had communication problems as well. The data mainly constituted measures of mathematical concepts and the results showed that there was a significant increase in the number of correct responses for both the group that included children with SEN and for the control group. Further, there were no statistically significant differences between the experimental and control group on the degree of change between pre-test and post-test scores.

Hunt *et al.* (1994) conclude that interactions in co-operative group learning by group members without disabilities did not negatively effect their level of achievement in mathematic concepts and that both groups significantly increased their mastery of the targeted mathematics objectives. Although this study directly assessed the impact of co-operative learning, indirectly it looked at what happens when a child with severe SEN is included in a classroom group.

Lundeen and Lundeen (1993), similarly to Hunt *et al.* (1994) also reported the results on the impact of inclusion on pupils without SEN as part of their study that aimed to evaluate the effectiveness of a collaborative teaching programme in a high school. Prior to the start of the programme, students with SEN were educated in resource/tutorial models. The collaborative teaching service delivery model involved enrolling students with special education needs in given subjects in which a regular educator and a special educator jointly taught the curriculum. Fifteen classes were included in the evaluation, involving a total of 318 students. Students with SEN mainly had learning difficulties, hearing impairments and behaviour disorders. Measures included reading comprehension scores, previous grades in traditional classes, grades in the collaborative teaching programme, comparison by student category, teaching team and content area interactions, mean grade point overall and by content area, and grade changes for individual students. Data were collected retrospectively from secondary sources (records) for the year before the collaborative teaching service delivery model and for the year after the programme.

Results comparing the performance of general education students between pre- and post-collaborative teaching model revealed that students' grades in English, social studies and science for the first semester of the collaborative teaching programme were substantially higher than for the previous school year (no collaborative teaching programme). More specifically, students taught in

collaborative classes presented an increase of almost half of a grade point. Also, almost half the students earned higher grades in classes in which the collaborative teaching programme was taught in the second semester compared with their performance in the same content area during the previous year. However, much of the grade improvements were not sustained through half of the school year. Therefore, at the end of the collaborative teaching programme, more than a third (39%) of students had better grades than the previous year, a third (34%) had poorer, and the remainder showed no change. Overall, the final grades of the students showed no significant change from the previous year. The authors conclude that all students' grades improved after their enrolment in the collaborative teaching programme. There are a number of concerns relating to the extent that the results of this study are trustworthy. Firstly, the research was not primarily aimed at studying the effects of student inclusivity, but at the effect of the collaborative teaching service delivery model. As such, the majority of the results focus on comparisons between pupils with SEN, and between pupils with and without SEN. In addition, any impact (neutral, and, in some cases, positive) could not directly be attributed to the presence, or numbers of pupils with SEN but also to the specific teaching model. Finally, Lundeen and Lundeen (1993) do not describe in detail the specific tests and analysis they performed, and there are also gaps in reporting of variables in sample details.

Hillen *et al.* (1992) carried out a study to examine the differential effects of the presence of children with special needs on teacher and class behaviour in mainstream settings. In order to do this, they undertook a case study in which they observed the teacher's behaviour on the academic and social life of pupils without SEN in two conditions: condition A, when pupils with SEN were integrated, and condition B when pupils with SEN were not present. Measures were taken across two domains: academic on-task behaviour and social behaviour. Students with SEN included two pupils with moderate learning difficulties and five with mild learning difficulties, who also had a variety of social-behavioural difficulties, reading deficits, language impairments, attention deficits and disruptive classroom behaviour. Both the teacher and the class were observed using the Observing Pupils and Teachers in Classrooms (OPTIC) schedule for all class and teacher.

The results showed that, when pupils with SEN were present, the level of on-task academic behaviour averaged at 75.41% compared with 74.57% when the SEN children were not present. Hillen *et al.* conclude that the presence of children with SEN (i) did not adversely affect the levels of on-task behaviour of pupils without SEN in classroom and (ii) their presence affected teacher behaviour in terms of higher levels of aversive responding.

Finally, similarly to Block and Zeman (1996) and Obrusnikova *et al.* (2003), **Rarick and Beuter (1985)** also report findings on the effects that educating pupils with SEN alongside their non-disabled peers can have on the physical education performance of those peers. Their primary aim was to determine if the integration of children with learning disabilities could successfully be accomplished. The project involved mainstreaming 25 students with learning difficulties in physical education classes with 85 students without disabilities. These were split into two groups: half were assigned to an experimental group (including students with SEN) and the remaining half to a control group. The measures, taken prior to the instructional programme and at the termination of the programme after five months, comprised videotapes at monthly intervals and tests in the standing long jump, softball throw for distance and 20-yard sprint.

Results in relation to the motor skills showed no significant differences between the integrated and the non-integrated groups of non-handicapped children on any of the three motor tests, revealing that integration of pupils with SEN in physical education classes does not adversely affect the performance of their non-handicapped peers. The authors conclude that students with learning difficulties can be successfully integrated in physical education classes. Rarick and Beuter's study is unusual in the sense that it attempted what they called "a reversed mainstreaming programme", in which, the general education students were integrated with pupils with SEN in the special education school settings. Since they had a control group in the same setting, the effect of the setting as a confounding variable can be minimised.

B: Academic outcomes – positive outcomes

As in the case of neutral outcomes, the majority of positive outcomes refer to measurements of reading/writing and mathematics (Saint-Laurent *et al.*, 1998; Shinn *et al.*, 1997; Stevens and Slavin, 1995). One study only includes primary measurements in science.

Medium weight of evidence

Mastropieri *et al.* (1998) set up a study to explore school factors associated with inclusive science instruction. Their main aim was to evaluate the achievement of students with disabilities with respect to their non-disabled peers, but as part of this, the authors also report on variables relating to inclusiveness of the classroom. Two groups were formed: an experimental group (one classroom) that incorporated activities-orientated science teaching of ecosystems, and a comparison group (two classrooms) that was based on textbook science learning of ecosystems. Four students with SEN (two with learning difficulties, one with physical difficulties and one with EBD) were included in the experimental group. Comparisons were made between the different groups in data (both qualitative and quantitative) that had been collected pre- and post-test.

The results suggested positive changes (measured by a multiple-choice test, a comprehension/performance test and an elaboration test in science) favouring the inclusive classroom in all cases. Mastropieri *et al.* conclude that their study provides further evidence concerning the effective inclusion of students with disabilities in science classrooms. However, the problem in drawing this conclusion is that there were two variables manipulated in the study: the inclusiveness of the classroom and the teaching approach. The authors do not differentiate between them and therefore the degree to which each variable accounts for the positive final effect is unclear.

Another study that measured the impact of inclusion for students without SEN, after using specific teaching strategies to promote inclusion, is the study by **Stevens and Slavin (1995)**. In a two-year study of the co-operative elementary school model (that uses co-operative learning as a philosophy for educational change) in two treatment and three comparison schools, Stevens and Slavin investigated the effect of co-operative elementary education on California Achievement Test scores in reading, language and mathematics. The sample included 1,012 students and measures comprised attitude measures, social relations measures and standardised test scores.

The results from post-tests one year after the set-up of the programme showed significant differences favouring the treatment group on reading vocabulary but no

significant differences on reading comprehension, language mechanics, language expression, mathematics computation and mathematics application. Post-tests after the end of the two-year programme showed significant effects favouring the treatment group in most of the measures (reading vocabulary, reading comprehension, language expression and mathematics computation). Also, after two years in the programme, students in the co-operative model had higher perceived ability in reading and language arts. The authors concluded mainly about the co-operative learning (as a useful basis for school restructuring and for producing important benefits for students). However, as in the study carried out by Mastropieri *et al.* (1998), the focus is on the effect of a teaching strategy that is expected to promote inclusion and therefore, not directly on the impact of inclusivity as such on the population of students without SEN.

Shinn *et al.* (1997) measured the effects of reintegrating 23 elementary-age students with, mainly, learning difficulties (20 with learning difficulties and two with sensory or communication difficulties) on these students and their low-reading peers. This study, similarly to that of **Huber *et al.* (2001)** that looked at how inclusion affects specific ability groups. The 23 students (who were previously educated in special education pull-out programmes) were placed into a general education classroom for reading instruction for 12 weeks. Their performance was then compared with the performance of a group of 66 students without SEN, using the curriculum-based measurement. The results are mostly presented as comparative achievements between students with SEN and their low-ability reading peers. Academic gains were measured through performance progress indicators that were collected at four different times: (i) initial (or baseline) performance, (ii) after four weeks of integration, (iii) after eight weeks of integration, and (iv) following 12 weeks of integration. The results for the sample of pupils without SEN showed that after four weeks of the reintegration, low-readers increased on some of the measures significantly but remained stable in others; while after eight weeks, students showed significant improvements in reading skills compared with reintegration measures. Since the study by Shinn *et al.* focuses mainly on the effects on pupils with SEN, the authors do not draw conclusion on the impact of inclusion on the whole population. However, compared with Huber *et al.*'s neutral overall results, Shinn *et al.* found some positive changes for the low ability peers. Huber *et al.* did report some positive changes in the scores of the low ability groups in mathematics but no significant change in reading.

One concern with the study by Shinn *et al.* is that about a quarter of the subset of low-reading peers (26%) received 'Chapter 1' reading services and therefore also had mild forms of SEN.

Low weight of evidence

Another study that partly focused on the effects of inclusion on the achievement of students without SEN is the study by **Saint-Laurent *et al.* (1998)**. This is another indirect report on the impact of inclusion on the school population using a specific educational model (PIER) and it is therefore difficult to evaluate what proportion of the results accounted for by the inclusion of pupils with SEN and what were accounted for by the impact of the specific model. The study actually evaluates the impact of an in-class service model (PIER) on the achievement of students at risk of school failure. PIER is a model in which pupils with SEN are educated alongside the general education students and it includes elements such as collaborative consultation, co-operative teaching, parent involvement, and strategic and adapted instructions. The model was implemented for one school

year in 13 different schools. The results were compared across two groups: the treatment group that utilised PIER and pupils with SEN were included in the general education classrooms, and the comparison group using traditional education models in which pupils with SEN were educated in resource classroom models.

Academic measures consisted of tests in reading, writing and mathematics, and were taken as pre-intervention in September and post-intervention in June. The results are described for general education students and for at risk groups that included pupils with SEN. Overall, it was found that the PIER programme benefits both at-risk and general education students in at least one academic area and that general education students were not held back by the presence of at-risk students who were present in the classroom. On the contrary, in reading and mathematics, the general education students benefited from the additional interventions that formed part of the model. Saint-Laurent *et al.* attribute this positive finding to the higher quality of teaching that the treatment group received and from the extra support from the classroom assistant.

C: Academic outcomes – negative outcomes

Medium weight of evidence

Brown (1982) studied the effect that classroom climate has upon students' learning and behaviour by investigating the difference between the social affective climate of typical primary classrooms and classrooms having a high number of students with SEN. Indirectly, therefore, the researcher examined the impact that the number of pupils with SEN had on the remainder of the class. The design involved analysis of four 'low-incident' classrooms, each with one learning or behavioural problem (L/BP) student and four 'high-incident' classrooms, each with four to eight L/BP students. Measures were collected of scores in the Barclay Classroom Climate Inventory for both academic (e.g. achievement/motivation, control/stability, etc.) and social (e.g. introversion, self-competency, etc.) outcomes.

The results showed that 'high-incident' classrooms had a negative impact on the class in terms of academic achievement. Students tended to be more unco-operative and withdrawn, and had more deficits in cognitive motivation. On the other hand, students in the 'low incident' classes tended to have more verbal and open behaviour. Brown concludes that the presence of students with learning and behaviour problems influences the classroom climate in a negative manner. He suggests, however, caution in taking this as an evidence towards selective placement decisions for pupils with this type of problems.

It is not very clear in this study whether the pupils described as having learning and problem behaviour are pupils with SEN that have been identified by an outside specialist (e.g. a psychologist) or whether they are pupils perceived by teachers as having these problems but who would not otherwise would have been characterised as having SEN. In addition, although the researchers used a collection of academic data, in their results section they do not differentiate in a clear manner between academic and social findings for students. Rather, the results they report are mainly social and the instruments they used reflect overall social measures (with some elements of academic measures as well).

In contrast to Mastropieri *et al.* (1998) (presented above), who conclude with a clear-cut positive impact of inclusion on students without SEN in the subject of

science, **Cawley et al. (2002)** have presented some negative outcomes when students with disabilities were included into the general education science classroom. The study describes a school-based project that concerns the design and implementation of a science project designed to include junior high school students with severe emotional/behavioural or learning difficulties into the general education classroom. In order to compare the performance between groups with students with SEN and groups without students with SEN, a variety of measures were taken. These included one-to-one interviews with teachers, observations during the science classes by teachers, questionnaires to teachers, scores from final exams and final grades scores, and school records such as attendance and discipline referrals.

The results relating to direct comparisons of academic grades between inclusion and non-inclusion classes are presented in a table without text to accompany the results and therefore we know nothing about whether the differences are significant or not. However, it seems that more GESs in non-inclusion classes pass final exams and grades in Grade 8, and that GESs in non-inclusion classes, Grade 7, had higher mean scores in final exams and final grades. Therefore, there seems to be a negative effect in terms of academic achievement. Despite these indications, **Cawley et al. (2002)** are in favour of inclusion of pupils with SEN in the general education science classroom (taking into consideration their results on social and other outcomes – these are presented below).

There is a problem with the presentation of **Cawley et al.**'s findings and with the conclusions since the authors only present means and standard deviations in academic attainment scores and their analysis or interpretations does not reflect some of the findings: for example, why does there seem to be a difference between final grades of pupils who are at general education inclusive and general education non-inclusive classes, and are these differences statistically significant?

D: Academic outcomes – mixed outcomes

Medium weight of evidence

Huber et al.'s (2001) study examined the differential impact of inclusion and inclusive practices on high, average and low achieving general education students. The study, which looked at the achievement of students over a three-year period, has been described in detail above (in the neutral results section).

The researchers found that overall there were no significant differences in the performance of students without SEN following the inclusion of pupils with SEN, but in some cases and in some groups they found mixed results. More specifically, the student skill factors had a statistically significant effect on incremental change in general education students' reading scores. As such, low achieving students (defined as students from the below-grade level skill group) and average achieving students (those who were from the within-grade level group) improved more than their high achieving peers. Similarly, in mathematics, student skill factors had a statistically significant effect on mathematics change scores, and again, students from the below-grade level skill group and students from the average level group, gained scores that were higher than students from the above-grade level skill group.

In addition they found that the number of pupils with SEN in a class was related to the students' mathematics scores but there was no clear pattern to this finding. More specifically, using the mathematics incremental change score as the

dependent variable, analysis of variance indicated significant differences among the group means of children educated with different numbers of students with SEN, while examination of the means revealed no clear direction to this finding.

Huber et al. (2001) conclude that, while low achieving general education students appeared to benefit academically, their higher achieving classmates appeared to lose ground when pupils with disabilities were included. As far as the effect of the number of pupils with SEN had on achievement, while the students' reading scores were not affected by the presence of their disabled peers, in mathematics, the effect was mixed.

The study carried out by **Daniel and King (1997)** aimed to determine the effects of students' placement in inclusive settings versus less inclusive or partial inclusive settings. The researchers compared three groups: (a) a non-inclusion control group with four classes (pupils with SEN were mainstreamed for a portion of the day with the majority being placed in pull-out programmes), (b) a random inclusion group with two classes (where students were randomly assigned across all classrooms in the school), and (c) a clustered inclusion group with six classes (where inclusion classrooms were formed that contained a higher percentage of pupils with SEN). The researchers collected a number of measures such as parent attitudes, student behaviours as reported by teachers, academic performance through SAT scores, and self-esteem reported by the students themselves. Overall the results indicated that the effects of inclusive programmes were mixed and difficult to decipher.

The results of measures of academic achievement were mixed and varied across grade and subject (mathematics, reading and spelling). For example, in the third grade, the students in inclusion groups did better in reading; in the fourth grade, the students in non-inclusion groups had higher gain scores in mathematics. The authors conclude in relation to academic achievement that there are few notable differences in academic achievement between inclusion and non-inclusion classrooms (defined as classrooms that have students for part of the day only).

4.2.2 Social outcomes

Table 4.7: Studies included in the in-depth review by cluster – social outcomes

A: Neutral outcomes	B: Positive outcomes	C: Negative outcomes	D: Mixed outcomes
<i>High overall weight</i>	<i>High overall weight</i>	<i>High overall weight</i>	<i>High overall weight</i>
Bear <i>et al.</i> (1991)			
<i>Medium overall weight</i>	<i>Medium overall weight</i>	<i>Medium overall weight</i>	<i>Medium overall weight</i>
Block and Zeman (1996) Obrusnikova <i>et al.</i> (2003) Sasso and Rude (1988) Sharpe <i>et al.</i> (1994) Tapasak and Walther-Thomas (1999)	Sasso and Rude (1988) Stevens and Slavin (1995)	Beuter (1984) Block and Zeman (1996) Brown (1982) Daniel and King (1997)	Cawley <i>et al.</i> (2002)
<i>Low overall weight</i>	<i>Low overall weight</i>	<i>Low overall weight</i>	<i>Low overall weight</i>
Helmstetter <i>et al.</i> (1994)	Hepler (1998) Hillen <i>et al.</i> (1992) Shevlin and O'Moore (2000)		Helmstetter <i>et al.</i> (1994)

A: Social outcomes – neutral

High weight of evidence

Bear *et al.* (1991) looked directly at the impact of integration on the self-perceptions of non-handicapped children and children with learning disabilities. They carried out two types of comparison: firstly, between groups of students without SEN when they are in inclusive and non-inclusive settings, and secondly, between students with and without SEN when they are in inclusion settings. Data were collected through self-completion questionnaires that included measures on global self-worth as well as domain-specific self-perceptions. More specifically, they used the self-perception measures (SPP-C) with sub-scales for scholastic competence, behavioural conduct and global self-worth. The examiners read the statements and pupils had to fill in the questionnaire.

The researchers hypothesised that self-perceptions of scholastic competence, behavioural conduct and global self-worth would be higher among the students without SEN in integrated settings than in non-integrated ones. Contrary to their predictions, scores of the 339 elementary students overall provided limited evidence on this. More specifically, differences between the two groups of students without SEN were not statistically significant except in self-perceptions of global self-worth among boys. Boys in integrated classes were found to have higher self-perceptions of global self-worth than did boys in non-integrated classes. The authors explain this finding through the social comparison theory and do not arrive at any general conclusions in favour or against inclusion. They also recommend caution in interpreting these results due to the fact that their sample was restricted to third-grade students.

Medium weight of evidence

Sharpe *et al.* (1994), who investigated the impact of inclusive school environments in academic achievement (see above), also examined the impact of

inclusion on general education students' behaviour. Again, the data comprised archival data collected two years after the pilot inclusion programme but this time they concerned behavioural areas of conduct and effort, taken from records of students' effort on report cards. Comparisons between two groups, an inclusive and a non-inclusive group, revealed that there were no significant differences between the two groups on behavioural measures and, more specifically, that there was no evidence of a decline in behavioural performance.

Block and Zeman (1996) also looked at the social outcomes of inclusion in terms of attitudes of pupils without SEN. Measures in attitudes comprised the Children's Attitudes Towards Integrated Physical Education Revised (CAIPE-R) questionnaire. Overall, they found that there was no difference in gain scores in sport-specific attitude between the inclusion and the non-inclusion group. The authors conclude that inclusion of pupils with severe disabilities in basketball session settings do not appear to negatively affect attitudes toward inclusion in PE and that, when properly supported, students with SEN can be included in regular physical education settings.

Similarly, **Obrusnikova et al. (2003)**, examining differences between pre-test and post-test measures in attitude scores, found that the attitude responses towards including a student with a disability in general physical education in both classes tended to be positive and stable during a two-week volleyball unit. She went further to examine relevant factors in attitudes and found that a positive attitude was significantly associated with students' family experience of disability, but there was no association with student general education experience or student general physical education experience. Students in the inclusive class and those with a family experience of disability were slightly more accepting of the student with disability.

Tapasak and Walther-Thomas (1999), in their study examining primarily effects for students with SEN (described under 'Studies on academic outcomes' above), also report some social outcomes for students without disabilities. They found that there was a decrease in the number of students without disabilities who were nominated by their peers as being shy/sensitive. It is inferred that this finding is viewed by the authors as a positive outcome in terms of social competence of the students. Overall, however, the authors report that there were no significant differences between pre- and post-test on teacher ratings of classroom behaviour problems, showing that the behaviour of pupils without SEN did not deteriorate as a result of the inclusion of pupils with SEN. A significant group main effect was observed on the sociability leadership subscale. Students without disabilities received more peer nominations for the ability to get along with and guide others (compared with students with disabilities – although this might have been the expected result). There were no effects on the aggressive disruptive subscale.

Sasso and Rude (1988) also investigated the social status of non-handicapped children as a result of integration efforts (see below for a more comprehensive review of the study). The study overall concludes that there were positive outcomes for pupils when their peers with SEN were included in their educational programmes. However, Sasso and Rude differentiated their results between students with high- and low social status, and their results for high-status students were mainly neutral.

Low weight of evidence

Helmstetter et al. (1994) considered the outcomes of interactions between students with moderate or severe disabilities, and their non-disabled peers in a national study in American high schools. The study mainly aimed at investigating whether non-disabled students would (i) experience changes in their self-concept, and (ii) modify their perceptions of other people and alter significantly their beliefs when students with disabilities were included in their school; it also aimed to investigate how these modifications might be evaluated. They administered a survey questionnaire to 166 high school students who had regular interaction with at least one pupil with SEN. Their students responded to a Likert scale survey of 29 benefit items and nine difficulty items. The items were clustered in 10 areas, including views on understanding the feelings and beliefs underlying the behaviour of others, fear of human differences, tolerance of others, development of self-concept, development of personal principles and personal development among others.

Comparisons were made between students' views of their benefits between two periods: before they had any contact with students with SEN and after they had some contact. Also, the data were correlated throughout the period of time they were in contact with students with disabilities both inside and outside school. The results overall were neutral. More specifically, when the results were examined across the amount of contact with pupils with SEN (both in school and outside school), there was no significant main or interaction effects. That is, the amount of time students spent with their peers did not overall affect their perceived benefits from this interaction. Overall, students reported positive outcomes of inclusion experiences with students with SEN, unaffected by the amount of contact with peers with SEN. **Helmstetter et al. (1994)** concluded that, overall, students without disabilities may benefit from inclusion experiences with students who have significant disabilities.

B: Social outcomes – positive*Medium weight of evidence*

Sasso and Rude (1988) investigated the social effects of integration on non-handicapped children. More specifically, the study measured the degree of social status change with high- and low-status primary school students without SEN who participated in a peer initiation intervention with children with SEN. The 12 integrated children had severe needs, primarily in learning and cognition, although a few students had severe multiple handicaps or autism. Although pupils did not have EBD as a primary type of SEN, they were reported to have additional behavioural problems, including non-compliance, social withdrawal and aggressive disruptive behaviours.

Two groups of non-handicapped students were formed, an experimental and a control. The experimental group received one-hour training to prepare them to be initiators and were instructed to engage in play activities during the morning recess session for seven weeks. The control group was present in the playground but did not receive any training and was not encouraged to interact with disabled students. Measures included a peer nomination sociometric tool, and data were collected as pre- and post-scores.

The results showed that students who were exposed to the peer initiation programme made significant sociometric gains over the control group, although

both treatment and control group sociometric scores were similar for the pre-test. The most impressive findings according to Sasso and Rude were the gains evidenced by low status non-handicapped students. The authors conclude that integration efforts can have a positive social effect on non-handicapped participants, especially to low-social status students. In attempts to explain the social status scores gained by these students, Sasso and Rude provide different explanations, such as the increased attention received by these students, despite their reservations that all non-handicapped students in both groups were given virtually identical levels of attention by the teachers and researchers.

This study is interesting in that, although it involves groups that could both be described as being physically integrated groups, only one of them is actually regarded as an 'integrated' group.

Stevens and Slavin's (1995) study, investigating the effects of co-operative elementary education on students' achievement (see above), also included social relations measures in which students were asked to list the names of their friends in the class. The students included in the schools in the co-operative models primarily had learning difficulties. Comparisons between pre- and post-intervention measures showed that, on the post-intervention measure, students in the co-operative model listed significantly more friends than did students in the comparison schools. The authors perceive this as a positive effect of the co-operative model on the students' social relationship, and therefore, as a social outcome.

Low weight of evidence

Hillen *et al.*'s (1992) study, presented above under 'Academic outcomes', found neutral outcomes for the students without SEN in terms of pupils' academic on-task behaviour when pupils with learning and emotional/behavioural difficulties were present in the classroom. However, the authors also report some positive outcomes in terms of the teacher's approval and lower levels of disapproval in the class with the integrated pupils with SEN compared with the control group. Although these results directly refer to the observed teacher's behaviour rather than the pupils', higher positive rewarding from teachers may imply lower levels of negative behaviour by students and more positive student outcomes. In view of the quasi-experimental case study nature of this research, Hillen *et al.* call for caution for generalisation of the results.

Shevlin and O'Moore (2000) undertook an evaluation study for an Irish programme linking mainstream pupils and their counterparts with severe and profound learning difficulties and assessed the social benefits for all involved. More specifically, the researchers studied the students' emotional reactions when pupils with SEN are included (a) before the inclusion programme had taken place, and (b) after the inclusion programme. They measured aspects such as being comfortable, being confident, their 'humanity', understanding others better, personal benefits and so on – as reported by the students without SEN.

Overall, the results showed the inclusion link programme had a positive impact on the emotional reactions of young people towards their peers with learning difficulties. Such results referred to increased insights into the lives of their peers with learning difficulties, recognition of a common humanity shared with their peers, overcoming of micro-conceptions about learning difficulties and adoption of a positive approach to their peers.

Shevlin and O'Moore (2000) conclude that mainstream pupils can and do benefit from interactions in inclusive practices and suggest such practices are an 'urgent educational priority'. However, this report includes several methodological gaps, such as that there is no information about the age of students (primary or secondary), or about the data-collection measures, procedures and methods of data analysis.

Hepler (1998) looked at the impact of social integration of children with emotional/behavioural difficulties on their non-disabled peers. The aim of the study was to determine if children with and without SEN in a school would both enjoy and learn behavioural, cognitive and affective skills in a social skills programme undertaken with a local mainstream school. The intervention therefore involved a social skills programme with children at the fifth grade. Two groups were formed according to the number of children with EBD, and according to the level of their social status: group I was the average/high status group including one child with SEN and three without SEN, and group II was the low status group with two children with SEN and two without SEN. The children with EBD came from a day treatment centre in the area. Several measures – including sociometric ratings, role-play, conversation, play, negative comments, solitary play – were taken before and after the programme. These data were collected through sociometric rating scales, role-play tests, observations and students' evaluation of the programme.

Results from the sociometric ratings showed that students without SEN increased their sociometric rating (although it is not specified whether these differences were statistically significant). Also, observational data showed that the children improved in three-quarters of the skills from the pre- to post-test period and they decreased the number of negative responses. Students also made other gains, such as in new knowledge related to the use of specific behavioural skills, entering an ongoing activity and increased awareness of new approaches for accommodating negative comments from peers. Overall, the results suggested that children without SEN participating in the study benefited from their interactions with the children who had SEN.

C: Social outcomes – negative outcomes

Medium weight of evidence

Beuter (1984) measured social behaviours in an etho-behavioural analysis of behaviours of pupils with SEN and pupils without disabilities in an integrated educational setting (described above). Social behaviours were measured by coding data collected through videotaping (playbacks of videotapes which had been taken four times during the experiment) and entering them onto a computer. Then the researchers, assisted by trained observers, quantified social behaviours and social interactions in terms of frequency, duration, sequence and directionality. The results showed that including children with learning difficulties did not facilitate social integration between them and their non-disabled peers. Pupils without SEN showed a tendency to withdraw socially towards the end of the programme. It is inferred here that this result, together with results on pupils with SEN not increasing their social interactions, were taken as negative and Beuter concludes that 'integration' acted as an inhibitor of social interaction.

Another study reporting some partial negative social outcomes is the study by **Block and Zeman (1996)**. The main focus of their study was on the impact of including students with disabilities in regular physical education classes; however,

they also measured the attitudes of pupils to inclusion. Although, overall, they found no difference in attitudes, the inclusion class had a significant advantage in pre-test attitudes and showed a slight drop-off in general attitude towards inclusion (i.e. the control class had more subjects with gains in general attitude compared with inclusion). This fact could have been taken as a small indication that inclusion somehow could have negative effects on pupils' attitudes. The authors, however, say that the fact that pupils with SEN had been included for three months prior to the set-up of the study may have affected the initial higher attitudes of the inclusion class.

Brown's (1982) study (described above) was among the few studies that has solely reported clear negative effects of the inclusion of pupils with SEN on students without disabilities. An analysis of the impact of the inclusion of different numbers of pupils with SEN revealed that, when a higher number of pupils with behavioural and learning difficulties were present in classroom, students tended towards significantly greater impulsive, acting-out behaviour with mood swings and unco-operative attitudes; on the other hand, students in the normal classes tended towards significantly greater controlled, tolerant, stable, and predicted behaviour. Furthermore, students in the learning and behaviour problem classrooms were more reticent, shy, tense and unhappy, while students in the normal classes had more open and outgoing behaviour. Brown concludes that having students with learning and behaviour problems in a classroom negatively influences its climate.

Similarly, the study by **Daniel and King (1997)** reports very strong negative social outcomes for students placed in an inclusive classroom, especially when the classroom contained a higher proportion of students with SEN. The three groups studied (as described above) were the non-inclusion group with pupils with SEN mainstreamed for only part of the day, the random inclusion group, and the clustered inclusion group with a higher percentage of pupils with SEN.

The study reports negative social outcomes for students in the inclusive setting, varying according to the degree of inclusivity. As far as behavioural problems were concerned, reported by teachers and parents on the internalising and externalising subscales of the Child Behaviour Checklist, it was found that students in the clustered inclusion classroom (that contained a higher proportion of included students with SEN compared with random groups) experienced more behaviour problems. The authors attribute this to possible frustration or boredom that students may feel in classrooms with a high variety of ability levels. In addition, students in all grades in the inclusion classroom uniformly reported lower self-esteem compared with their peers in non-inclusion classrooms (again, defined as classrooms with students with SEN included for part of the day only, while for the majority of the day being educated in pull-out programmes).

D: Social outcomes – mixed outcomes

Medium weight of evidence

Cawley et al.'s (2002) study on the outcomes of a school-based project in science to include school students with severe emotional and learning difficulties reports some positive and some negative results. As seen above, in terms of academic achievement, although the authors do not clearly report negative outcomes they present tables of results where some mean scores in science testing have been dropped. Despite this, the authors in general conclude in favour of inclusion models. They also examined some social outcomes, measured as the

total numbers and comparative rates of discipline referrals between classes with included students with SEN and classes without students with SEN, as well as between classes of science teaching and classes in other subjects. The authors report mixed outcomes in terms of behavioural measures when students with learning and emotional/behavioural problem were included in science classes. The behaviour was measured in terms of discipline referrals during the science class when students with SEN were included and was compared with classes that had no students with SEN. The results from examining discipline referrals (taken from the school's records) showed that this varied across different classes (such as science or non-science class) and grades. For example, the authors report that, in one case of an inclusion class, there were 50% fewer discipline referrals than a class not having students with SEN. However, in other cases, the inclusion classes seemed to have more discipline referrals than the non-inclusion class. This varied across grades (grades 7 and 8 were studied) and across subjects (either in science or other than science subjects).

Low weight of evidence

The study by **Helmstetter et al. (1994)** has been described above under 'Neutral outcomes' in terms of students not changing their views when these were examined generally as a result of time spent with students with SEN. However, this study also revealed some mixed findings in other variables. For example, there are some positive findings in terms of students being more responsive to the needs of others and showing an increased appreciation of diversity when they had more than one hour's weekly contact with peers with disabilities. In addition, when the type of interaction was examined as an independent variable, there were some statistically significant effects reported, such as the students who shared a class with a person with a disability had a significantly lower benefit score in valuing relationships with persons with disabilities but significantly higher benefit score in tolerance of others. However, Helmstetter et al. (1994) overall conclude that students without SEN may benefit from inclusion experiences.

4.2.3 The impact of inclusion across type of SEN and age

Examining the impact of the students' type of SEN on the overall outcomes of pupils without disabilities was somewhat complicated since the majority of studies involve samples of students who either had a variety of SEN or were individual students with more than one single type of SEN. In the latter case, the students' main type of SEN was taken into account. Tables 4.8 and 4.9 present the key outcome for each study across type of SEN for primary and secondary schools.

Table 4.8: Impact of inclusion across type of SEN: primary schools

	Cognition and learning		EBD		Sensory/Physical		Communication	
	Academic	Social	Academic	Social	Academic	Social	Academic	Social
Primary	Affleck <i>et al.</i> (1988)	Bear <i>et al.</i> (1991)	Affleck <i>et al.</i> (1988)	Sharpe <i>et al.</i> (1994)	Huber <i>et al.</i> (2001)	Tapasak and Walther-Thomas (1999)	McDonnell <i>et al.</i> (2003)	Tapasak and Walther-Thomas (1999)
	Beuter (1984)	Sharpe <i>et al.</i> (1994)	Huber <i>et al.</i> (2001)	Tapasak and Walther-Thomas (1999)	Obrusnikova <i>et al.</i> (2003)	Obrusnikova <i>et al.</i> (2003)	Hunt <i>et al.</i> (1994)	Sasso and Rude (1988)
	Hillen <i>et al.</i> (1992)	Tapasak and Walther-Thomas (1999)	Lundeen and Lundeen (1993)	Hepler (1998)	McDonnell <i>et al.</i> (2003)		Tapasak and Walther-Thomas (1999)	
	Huber <i>et al.</i> (2001)	Hillen <i>et al.</i> (1992)	Sharpe <i>et al.</i> (1994)	Sasso and Rude (1988)	Hunt <i>et al.</i> (1994)		Rankin <i>et al.</i> (1999)	
	Hunt <i>et al.</i> (1994)	Sasso and Rude (1988)	Tapasak and Walther-Thomas (1999)	Brown (1982)	Tapasak and Walther-Thomas (1999)		Willrodt and Claybrook (1995)	
	Lundeen and Lundeen (1993)	Shevlin and O'Moore (2000)	Willrodt and Claybrook (1995)		Rankin <i>et al.</i> (1999)		Saint-Laurent <i>et al.</i> (1998)	
	McDonnell <i>et al.</i> (2003)	Stevens and Slavin (1995)	Saint-Laurent <i>et al.</i> (1998)		Willrodt and Claybrook (1995)		Shinn <i>et al.</i> (1997)	
	Rankin <i>et al.</i> (1999)	Beuter (1984)	Mastropieri <i>et al.</i> (1998)		Lundeen and Lundeen (1993)			
	(1985)	Daniel and King (1997)	Brown (1982)		Saint-Laurent <i>et al.</i> (1998)			
	Sharpe <i>et al.</i> (1994)		Huber <i>et al.</i> (2001)		Mastropieri <i>et al.</i> (1998)			
	Tapasak and Walther-Thomas (1999)				Shinn <i>et al.</i> (1997)			
	Willrodt and Claybrook (1995)							
	Mastropieri <i>et al.</i> (1998)							
	Saint-Laurent <i>et al.</i> (1998)							
	Shinn <i>et al.</i> (1997)							
	Stevens and Slavin (1995)							
	Daniel and King (1997)							
	Huber <i>et al.</i> (2001)							

Code:  neutral  positive  negative  mixed

Table 4.9: Impact of inclusion across type of SEN: secondary schools

	Cognition and learning		EBD		Sensory/Physical		Communication	
	Academic	Social	Academic	Social	Academic	Social	Academic	Social
Secondary	Block and Zeman (1996)	Block and Zeman (1996)	Lundeen and Lundeen (1993)	Cawley <i>et al.</i> (2002)	Lundeen and Lundeen (1993)			
	Lundeen and Lundeen (1993)	Helmstetter <i>et al.</i> (1994)	Cawley <i>et al.</i> (2002)					
	Cawley <i>et al.</i> (2002)	Shevlin and O'Moore (2000)						
		Block and Zeman (1996)						
		Cawley <i>et al.</i> (2002)						
		Helmstetter <i>et al.</i> (1994)						

Code:  neutral  positive  negative  mixed

The discussion below is mainly based on studies that report high or medium weight of evidence; when the low weight of evidence is taken into account, this is made clear.

Cognition and Learning

Almost all studies involved samples of students who had difficulties in cognition and learning either as the main or as a secondary type of SEN. Overall, the inclusion of pupils with cognitive and learning difficulties had a neutral effect on their peers without difficulties. However, this varies across the type of outcomes (academic and social) and the school type (primary and secondary).

In terms of academic outcomes in primary schools, there were six studies that present results that were considered to be high or medium weight of evidence (and six that were of low weight of evidence) which report neutral outcomes when students with cognitive and learning difficulties were included in the primary schools; three studies presented positive outcomes (and one that was low weight of evidence) and one study presented mixed outcomes (and one that was low weight of evidence). There was no study that reported negative academic outcomes for a primary school population when students with learning difficulties, as the main type of SEN, were included. The six studies considered medium or high weight of evidence that presented neutral outcomes reported on the inclusion of pupils who either had learning and cognitive difficulties as the main type of SEN or a mixture of difficulties. In Beuter's (1984) study, all 25 included students had cognitive and learning difficulties. Also, Beuter (1984) reports the ratio of inclusion as approximately one pupil with SEN to four pupils without SEN. In the remaining five studies (Huber *et al.*, 2001; McDonnell *et al.*, 2003; Sharpe *et al.*, 1994; Tapasak and Walther-Thomas, 1999; Rankin *et al.*, 1999), the majority of pupils had learning difficulties as the main type of SEN and, in four of these studies, some students had severe to profound disabilities (Beuter, 1984; McDonnell *et al.*, 2003; Rankin *et al.*, 1999; Sharpe *et al.*, 1994).

On the other hand, of the three medium weight of evidence studies reporting positive academic outcomes for the primary school level students when students with cognitive and learning difficulties were included, no student had severe or profound difficulties (Mastropieri *et al.*, 1998; Shinn *et al.*, 1997; Stevens and Slavin, 1995). Stevens and Slavin (1995) only refer to their sample in general as being 'learning disabled' without mentioning any other details, while Mastropieri *et al.* (1998) and Shinn *et al.* (1997) imply that their samples' learning difficulties were mild to moderate.

The picture for social outcomes when pupils with cognitive and learning difficulties are included in primary schools is slightly more positive (compared with mainly neutral academic outcomes) since there are four studies reporting positive outcomes (two of which are of medium evidence) and three (all medium or high evidence) reporting neutral overall findings. Of the two 'medium weight' studies with positive findings, both refer to the inclusion of pupils with severe to profound learning difficulties (Sasso and Rude, 1988; Sharpe *et al.*, 1994).

There are few studies which investigated the effects of inclusion at the secondary school level. Of the medium weight of evidence studies, Block and Zeman (1996) reported neutral academic outcomes for the rest of the class when three students with moderate to severe learning and cognitive difficulties were included, while the same study also reports some negative outcomes for some other academic skills

as did Cawley *et al.* (2002). This later study, however, involved samples of students who collectively have a mixture of learning difficulties and EBD.

Emotional and behavioural difficulties

Of the studies that report either positive or neutral academic and social outcomes when primary pupils with EBD were included, those of medium weight of evidence had pupils with EBD as a sub-sample of the total sample with SEN or students with other difficulties (mainly learning) who had additional emotional/behavioural problems.

There are only two studies in which the included students had EBD as the main difficulty. The first, a study by Hepler (1998), reports that primary school students benefited socially from their interactions with included students with emotional/behavioural difficulties. However, this study has been rated as low in the overall weight of evidence, since it mainly looks at the effects of a programme rather than inclusion as such.

The second study in the medium weight of evidence category, Brown (1982), reported negative outcomes, both academic and social, for primary school pupils when their peers with EBD, as the primary type of SEN, were included. He found that the greater the number of students with behavioural (and learning) problems in the classroom, the more it influences the classroom climate and learning environment in a negative way in terms of all students' cognitive abilities and social behaviour.

At the secondary school level, there were only two studies that report the outcomes of the inclusion of pupils with SEN on the school population. One of them has been rated as low in the overall weight of evidence, while the other, which involves samples with learning or behavioural problems, reports negative academic and mixed social outcomes.

Therefore, the evidence in the studies reviewed here suggests that at primary school level, the impact of inclusion of students with EBD on outcomes for other children can sometimes be negative (Brown, 1982).

Sensory and/or physical needs, and communication and interaction

There are no studies that considered the outcomes of the inclusion of pupils with sensory and/or physical needs or with communication/interaction difficulties on secondary school population other than the study by Lundeen and Lundeen (1993). This study has, however, been rated as low in terms of overall weight of evidence as it studied the direct effects of a collaborative teaching service delivery model rather than inclusivity.

Some of the remaining studies that report positive or neutral academic outcomes for primary school pupils involved sub-samples of pupils with hearing impairments, health impairments, speech and language disorders and multiple disabilities as the main type of SEN (Mastropieri *et al.*, 1998; Rankin *et al.*, 1999; Shinn *et al.*, 1997; Tapasak and Walther-Thomas, 1999).

Finally, it has to be mentioned here that there is not a single study that has reported negative or mixed outcomes of the inclusion of students with sensory and/or physical needs, and/or communication/interaction difficulties. An exception is the study by Huber *et al.* (2001) whose sample included some pupils described

as having health impairments, who, however, were also quite likely to have additional learning or behavioural/emotional needs.

4.3 In-depth review: quality-assurance results

Generally all staff involved in data-extraction were in agreement about the main points to draw out of the review, and about the overall quality and impact of the studies. However, the reviewers consistently raised two issues that needed careful scrutiny before deciding whether studies should have been included and/or the extent to which it was possible to draw conclusions from the findings that related to our research question

The first of these related to the focus of the studies that were being reviewed. There were several studies that were not directly related to the impact of inclusivity on non-SEN school populations but which, instead, focused on the impact of some intervention programme on pupils with SEN in a mainstream setting. The impact on non-SEN pupils was almost incidental to the main study. For example, the study by Saint-Laurent *et al.* (1998) actually focused on the impact of a mainstream class service delivery model (PIER) on pupils with SEN. Other similar studies also looked at the impact of a specific programme on pupils with SEN. For example, Hunt *et al.*, (1994), Lundeen and Lundeen (1993), and Stevens and Slavin (1995) studied inclusion within the context of co-operative learning, and Hepler (1998) studied the impact of inclusion on non-SEN students within a programme that mainly involved social skills learning for those with SEN. All these studies also included outcome measures for pupils without SEN. The problem with studies such as this is that it is difficult to account for how much of the impact on students without SEN is attributed to the specific programme or simply to the inclusion of pupils with SEN. After some discussion among the members of the Review Group, it was agreed that the above studies would be included in the data-extraction and in the final in-depth review but would be given a low overall weight of evidence.

The second issue concerned differences in judgments on specific questions in the data-extraction instrument and in the consequent process of reaching a judgment in relation to the quality of studies. Overall, there were only a few disagreements and this was considered to be inevitable since the reviewers often came from different epistemological backgrounds and therefore tended to form different judgments in relation to the quality of a specific piece of research. In general, these disagreements were easily resolved between the two reviewers.

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

As referred to earlier, there were various activities in which users and other staff were involved in supporting the review. This process was prevalent at all stages of the review in a number of ways.

First, at a meeting with 30 educational psychologists in June 2004 we were able to share the preliminary findings and to gauge their reactions. Second, at regular steering committee meetings in relation to the Inclusion and Pupil Achievement

project, different aspects of the review in relation to pupils without SEN were shared with DfES staff, LEA advisors and headteachers. Third, at informal seminars at the School of Education the Review Group has shared the findings with staff and students.

5. FINDINGS AND IMPLICATIONS

In this chapter, we consider the key outcome of this review, including the availability and appropriateness of the studies that met the inclusion criteria, the strengths and limitations of the review, the key findings, and implications for policy, practice and research.

5.1 Summary of principal findings

5.1.1 Identification of studies

Initially 7,137 papers were identified through searches of electronic databases. Having screened all their titles and/or abstracts, 161 papers were marked as possible ones to be included. Of the 6,986 that were excluded at this stage, the vast majority (82%) were excluded because they did not describe the impact of population inclusivity. In addition, 13% that describe the impact of inclusion on pupils with SEN did not refer to the general population without SEN. A small number (0.6%) did not make any references to the impact on personal, social and/or academic outcomes. Although 2% were concerned with the impact of population inclusivity on the general population, their focus was on pre-school or post-16 provision and a further 1% were reviews of the literature rather than the results of empirical research. A small number, 0.3%, were not written in English and, finally, 1.5% did not provide a robust evidence of measurement.

Of the 161 potentially included papers, nine turned out to be duplicates. It was not possible to obtain paper copies of 33 of them mainly because they were PhD or MPhil theses, mostly from the USA and Australia. Initially, a few of these theses were ordered through inter-library loan, but either the British Library could not trace them or it was too expensive to obtain copies. Full text documents of the remaining 119 papers were obtained. This led to a further reduction in the number of papers to the eventual number of 26 studies reported in 26 papers that were subjected to keywording and the data extraction process.

5.1.2 Nature of studies selected for in-depth review

All 26 studies that were selected for inclusion in the data extraction and in the synthesis of evidence were evaluation studies, 15 of which were naturally occurring and 11 'researcher manipulated' involving some form of experimental design. Of the 26 studies, 16 also looked at the impact of inclusion on pupils with SEN and, for some of these, this was the primary focus of the study, although they also included the impact on students without SEN. The studies ranged over a period of 22 years, the earliest study being in 1982 and the most recent one being published in 2003. More specifically, there were five studies before 1990, 15 between 1990 and 1999, and five studies from 2000 onwards.

The majority of these studies were American (N = 22); there were also two studies from Australia, one from Canada and one from Ireland. There was a slight preponderance of studies in which the included pupils experienced difficulties in

the area of cognition and learning, although pupils with other types of difficulties were also mentioned. The majority of the studies (21) focused on academic outcomes and these were measured in a wide variety of ways, including class tests, national examinations and teacher ratings. Typically, outcome measures of one group, where the pupils had worked alongside pupils with SEN, were compared with similar groups within in the same school or in a similar school, who had not worked alongside such pupils.

Virtually all the studies (21) focused on the outcomes of inclusion for primary aged pupils. The nature of the inclusion experienced by the pupils with SEN is described in different ways. In some studies (16), this is described as the proportion of pupils with SEN in a mainstream class, whereas in others (11) it is described as the number of hours per week (or day) that a child with SEN spent in a mainstream class. Some studies describe inclusion in both ways.

5.1.3 Synthesis of findings from studies in in-depth review

In order to make overall sense of the findings from the systematic review, it is important to remember that many of the studies did not focus on one type of outcome only, academic or social. Although 12 focused solely on academic outcomes and 5 addressed social outcomes, there were 9 that considered both academic and social outcomes. Moreover, some studies (e.g. Helmsteter *et al.*, 1994; Huber *et al.*, 2001) considered different aspects of either academic or social outcomes. All of this means that the 26 studies yielded more than 26 findings. Indeed, as Tables 4.6 and 4.7 show, the 26 studies yielded 40 different findings and Table 4.4 indicates that, of these, 9 were positive, 6 negative, 21 neutral and 4 mixed.

When the studies are subdivided further to include outcomes for primary and secondary aged pupils and where pupils with different types of SEN were included (see Tables 4.8 and 4.9), there are no less than 78 different outcomes. Of these, 19 are positive, 7 negative, 45 neutral and 7 mixed. Table 5.1 compares key findings for academic and social outcomes with outcomes related to the age of the non-SEN pupils and the SEN type of the included findings.

Table 5.1: Summary of findings and outcomes from the review

	Outcomes (adapted from Tables 4.8 and 4.9)		Outcomes (adapted from Tables 4.8 and 4.9)	
Positive	9	23%	19	23%
Negative	6	15%	7	9%
Neutral	21	53%	45	58%
Mixed	4	10%	7	9%
Total	40		78	

Essentially the overall conclusions from Table 5.1 are similar, whether the results are summarised as findings or outcomes. Around 25 percent are positive, between 9 and 15 percent are negative, between 53 and 58 percent are neutral, and around 10 percent are mixed.

Overall, these findings indicate that placing children with SEN in mainstream schools is unlikely to have a negative impact on academic and social outcomes for pupils without SEN. These findings confirm the conclusions from the literature reviewed in section 1.4. It can be seen from Table 5.1 that, if neutral and positive

findings are grouped together, 76% of findings and 81% of outcomes indicate that there will be no adverse effects on pupils without SEN of including pupils with special needs in mainstream schools.

A closer look at the findings reveals a number of further points:

- There is slightly more emphasis in the studies on academic, rather than social, outcomes (22 out of 40 findings) and the main outcomes on the academic side are neutral, for over half of all studies. However, of the six studies with negative findings, four focus on social aspects suggesting that negative findings might be more likely to occur on this dimension. Indeed neutral, positive, negative and mixed findings are more evenly spread across the social dimension, 7, 5, 4 and 2 respectively (see Tables 4.6 and 4.7).
- Some of the findings (e.g. Saint-Laurent *et al.*, 1998) suggest that the inclusion of pupils with SEN in primary schools can have a positive impact on the achievement of their mainstream peers, particularly if the support offered to the pupil with SEN is well managed. This confirms some of the key findings from the EPPI-Centre review on the impact of paid adult support (Howes *et al.*, 2003).
- Tables 4.8 and 4.9 indicate that most of the outcomes relate to primary aged pupils (65 out of 78) and the main focus of the majority is on the impact of placing pupils with difficulties in the area of cognition and learning. In the primary phase, there are slightly more studies that focus on the impact of placing pupils with EBD in mainstream schools than those with sensory and physical impairments. In general, there are few studies on the impact of placing pupils with difficulties in communication and interaction, particularly at the secondary level.
- A higher proportion of outcomes in the secondary phase refer to the negative impact of placing pupils with SEN in mainstream schools: 3 out of 13 outcomes as opposed to 4 out of 65 at the primary phase. This suggests that there may be more problems in managing inclusion successfully in secondary schools.
- Table 4.8 indicates that negative outcomes are not related to one SEN type. However, there is one study (Brown, 1982) which focused exclusively on the impact on mainstream pupils of placing children with EBD in a school. His findings stressed the negative outcomes for pupils of such placements and confirms the views expressed in other studies that are not part of this review (e.g. Dyson *et al.*, 2004) that pupils with EBD cause most problems for other children when placed in mainstream schools compared with other pupils with SEN.
- From the studies reviewed, it is not possible to draw any conclusions as to the impact of placing pupils with SEN in mainstream schools on achievement across different curriculum areas. The majority of studies either use class-based assessments or standard curriculum measures (mathematics, literacy, etc.) or they rely on an achievement test (e.g. the Californian Assessment Battery). There are, however, a few that refer to pupils' learning of motor skills in physical education lessons. None of the findings indicate that the 'inclusion effect' is more or less serious for any one particular curriculum area.

5.2 Strengths and limitations of this systematic review

To our knowledge, this is the first systematic review of the literature that has focused on the relationship between the inclusion of pupils with SEN and the achievement of their peers without SEN. Other reviews, mostly from the USA, have not been strictly systematic, nor have they been comprehensive and covered all types of SEN, academic and social outcomes and the full age range of pupils in mainstream schools. We believe that our search of the databases and other grey literature has been comprehensive and that it is unlikely that we have missed other studies that have addressed this area.

There are, however, a number of limitations that should be considered when judging the overall weight of the review's findings and in considering the implications for policy and practice.

1. There were several studies in which the main focus of the research was on the impact of inclusion for pupils with SEN and the element that addressed the impact on non-SEN peers was secondary. Although this is not strictly speaking a limitation, it might reflect the fact that the impact of inclusion on non-disabled pupils has always been of secondary importance.
2. Almost all the research has been carried out in the USA and, given the different contexts – assessment arrangements; special needs support services, range of provision, etc. – it is important to be cautious in generalising the findings of the review to the UK.
3. It was unfortunate that we were unable to obtain a small number of studies because the British Library was unable to locate them or because it was too expensive to get physical access to them. In addition, the review might have missed important studies carried out in other European countries because of language restrictions.
4. Of more fundamental concern, perhaps, is the slightly loose or uncertain way in which the term 'inclusion' was defined by the authors of the studies that were reviewed. It was not always clear whether the inclusion arrangements involved fulltime placements in mainstream class; whether and to what extent such placements were supported; and whether pupils were withdrawn to other special classes for certain lessons and for how long. In addition, in one study, Rarick and Beuter (1985) focused on what they called 'reversed mainstreaming' where the general education students were integrated with pupils with SEN in the special education school settings. All this means that it is not possible to judge from the review whether or not certain types of inclusion arrangements were associated with particular academic or social outcomes.
5. In organising the studies, we used the categories from the Code of Practice as a way of grouping SEN type. These categories are specific to the UK and not used in other countries, in particular the USA. Therefore we had to make a judgment as to which Code of Practice category the pupils in a study seemed to fit. This was not always easy and many studies seemed to involve pupils who spanned more than one group.

6. The majority of studies focused on the impact of inclusion when included students had more than a single SEN type (as a group or in individual students). However, since the majority of studies included in the review provided a single overall result for the impact of having students with a wide range of SEN types in the class(es), a danger of masking effect is entailed. More specifically, it is possible that there might have been positive impacts and negative impacts associated with particular types of SEN but these could have been lost in the overall analysis.
7. There is a scarcity of studies that have considered the impact of placing pupils with SEN on the academic and social outcomes for secondary aged pupils.
8. When grouping the studies, we decided to include those that focused on outcomes related to physical education as 'academic' outcomes. Although there were only three such studies and the findings were in line with those that focused on more traditional academic outcomes, it is important to bear in mind the broader interpretation of the term 'academic' outcomes referred to in this review.
9. Finally, it is worth noting that all but one of the studies was quantitative involving 'measures' of one kind or another. There appears to be a dearth of qualitative studies that have addressed our research question and which meet the criteria for inclusion in this review.

Taken as a whole, the above limitations indicate that it is important to be cautious when coming to an overall conclusion about the findings of the review. In particular, the dearth of UK research in this area, the difficulty in determining the extent and range of special needs that the pupils experienced, the limited number of studies in secondary schools and the fact that the type of inclusive arrangement were not always clearly defined are all relevant caveats that should be borne in mind when planning future policy and practice in this area.

Nevertheless, the review has coincided with the publication of the report on the DfES study on the relationship between inclusion and pupil achievement in English schools (Dyson *et al.*, 2004). As the results of this study were not available when this review was being carried out, the findings are not included here. However, the DfES study found that, for all practical purposes, there was no relationship between the placement of pupils with SEN in mainstream schools and the achievements of pupils without special needs. This is an important finding and is in line with the overall conclusions to this review. This suggests that the limitations discussed above might not have affected the overall findings of the review and hence we can be more certain in coming to the overall conclusion that, by and large, placing children with SEN in mainstream schools is unlikely to have a significant impact on overall levels of achievement among pupils without special needs.

5.3 Implications

5.3.1 Policy

Since 1997, the English education system, like many across the world, has been committed to moving in a more inclusive direction. Amongst the many things that

this commitment can mean, it certainly implies an intention to maintain as many children as possible in mainstream settings and to reduce the reliance on segregated special schools. There is nothing in this review or in the DfES study in which we have been involved (Dyson *et al.*, 2004) to suggest that this commitment is likely to have a significant impact on overall levels of attainment in mainstream schools. Even taking into account the limitations referred to above, the findings suggest that, in relation to the impact on the attainments of non-SEN pupils, the government, LEAs and schools should have no concerns about pursuing the inclusion agenda.

This is not to say, of course, that policymakers should ignore problems faced by schools in trying to become more inclusive and raise standards. There is evidence from other literature (Fox *et al.*, 2004), that some schools are concerned about inclusion, particularly for pupils with emotional and behavioural difficulties. Therefore, policymakers should pursue inclusion policies in an informed way, consulting with all relevant stakeholders at all times. In addition, the lack of studies in the secondary sector suggests that schools and LEAs should pursue the inclusion agenda with some caution and, where possible, commission research that can explore this complex area in more depth.

5.3.2 Practice

In relation to the implications of this review for practice, it is important at the outset to remind ourselves about the nature of the studies that were included in the review. By and large, these studies were of 'minimum effort' inclusion in that, with the exception of the studies by Hepler (1998), and Shevlin and O'Moore (2000), no specific effort had been made to help the non-SEN children to adjust to the presence of their non-SEN peers. Hence, for the majority of these studies, there was little or no investment in the success of inclusion, in relation to non-SEN pupils. This puts the generally neutral or positive findings of the review into some form of context. For, if we had reviewed studies where special efforts had been made to make the inclusion effective for non-disabled children, then we might have seen even more positive results. The implications for practice are that, even in 'minimum effort' inclusion, the impact on academic and social outcomes is likely to be positive or neutral.

On the whole, therefore, this review suggests that schools, parents and LEA professionals should have no concerns about the impact of inclusion on achievement of pupils without SEN, especially in primary schools. This applies across all four categories of SEN. However, these studies and other research reviews (Harrower, 1999; Farrell, 2000) indicate that successful inclusion does not occur in a vacuum. Parents, teachers and pupils need to be fully committed to the idea, programmes of work have to be carefully planned and reviewed regularly; and support staff need to work flexibly as a team, and receive appropriate support and training. In these general conditions, schools should feel able to include pupils with SEN without fearing that it will damage the attainments of the remaining pupils. However, this is not to say that mainstream schools should include huge numbers of pupils with SEN so that the balance of the school as a whole is affected. Although we found no studies where large groups of pupils were included, there is anecdotal evidence from some secondary schools which have high proportions of pupils with learning difficulties (e.g. over 30%) that the impact on the ethos and possibly the achievements of a school might be adversely affected if large numbers of pupils with SEN were included. This was also supported by the cases studies in the DfES study (Dyson *et al.*, 2004).

In addition, it is important to remember that that around a quarter of findings were negative in relation to both social and academic outcomes. Although small in number, this does raise questions that practitioners need to address when considering whether to include more children with SEN: How many should be included? How should the support be managed? Which other agencies should be conducted to provide support? Should the inclusion arrangements involve fulltime placements in mainstream classes or should some form of withdrawal be used?

5.3.3 Research

A number of implications for further research arise from this review. First, we found few studies that focused on secondary schools. Given concerns that parents and others have expressed about the viability of inclusive arrangements in secondary schools (e.g. Fox *et al.*, 2004), further research needs to address the question about the impact of inclusion on the achievements of non- SEN groups in a secondary context. The DfES study referred to earlier suggests that an analysis of data from the National Pupil Dataset (NPD) reveals a very small, and for all practical purposes insignificant, negative relationship between inclusion and pupil attainment in secondary schools. There are huge variations between schools, suggesting that it is detailed studies at the individual school level that are likely to yield most promising findings.

Second, the majority of studies in this review centred on the impact of including relatively small numbers of pupils on a small number of their peers. Studies that included larger groups are needed to test the robustness of our findings more thoroughly. Again the DfES study has made a contribution here, although we were unable to study the impact of including pupils with different types of SEN on the attainments of pupils without SEN.

Third, there is a need for more longitudinal research that could trace the relationship between inclusion and the achievements of non-SEN pupils over time.

Fourth, research needs to be carried out that links the attitudes of teachers, other support staff and parents in a school towards inclusion to see if this is an important intervening variable that could have an impact on academic and social outcomes for all the pupils.

Fifth, it would be interesting to study the effects on social and academic outcomes of preparing children and young people for the inclusion of pupils with disabilities, particular those with more severe problems. One would predict that preparing young people to live and work alongside their peers with disabilities should have a positive impact on a range of academic and social outcomes.

Finally, more studies should focus on the views of pupils without SEN about inclusion. Given current interest in involving users in planning, carrying out and evaluating research, it is surprising that so few studies actually focus on the pupils' views.

6. REFERENCES

6.1 Studies included in map and synthesis

Affleck JQ, Madge S, Adams A, Lowenbraun S (1988) Integrated classroom versus resource model: academic viability and effectiveness. *Exceptional Children* **54**: 339–348.

Bear GG, Clever A, Proctor WA (1991) Self-perceptions of nonhandicapped children and children with learning disabilities in integrated classes. *Journal of Special Education* **24**: 409–426.

Beuter A (1984) Ethobehavioural analysis of the social behaviours of trainable mentally retarded and intellectually normal children in an integrated educational setting. *American Corrective Therapy Journal* **38**: 11–18.

Block ME, Zeman R (1996) Including students with disabilities in regular physical education: effects on nondisabled children. *Adapted Physical Activity Quarterly* **13**: 38–49.

Brown W (1982) Classroom climate: possible effects of special needs on the mainstream. *Journal for Special Educators* **19**: 20–27.

Cawley J, Hayden S, Cade E, Baker-Kroczyński S (2002) Including students with disabilities into the general education science classroom. *Exceptional Children* **68**: 423–435.

Daniel LG, King DA (1997) Impact of inclusion education on academic achievement, student behaviour and self-esteem, and parental attitudes. *Journal of Educational Research* **91**: 67–80.

Helmstetter E, Peck CA, Giangreco MF (1994) Outcomes of interactions with peers with moderate or severe disabilities: a statewide survey of high school students. *Journal of the Association for Persons with Severe Handicaps* **19**: 263–276.

Hepler JB (1998) Social integration of children with emotional disabilities and nonhandicapped peers in a school setting. *Early Child Development and Care* **147**: 99–115.

Hillen R, Houghton S, Bain A (1992) The effects of integrating children with special needs into mainstream classes: a case study. *Educational and Child Psychology* **9**: 52–56.

Huber KD, Rosenfeld JG, Fiorello CA (2001) The differential impact of inclusion and inclusive practices on high, average and low achieving general education students. *Psychology in the Schools* **38**: 497–504.

Hunt P, Staub D, Alwell M, Goetz L (1994) Achievement by all students within the context of cooperative learning groups. *Journal of the Association for Persons with Severe Handicaps* **19**: 290–301.

- Lundeen C, Lundeen DJ (1993) Effectiveness of mainstreaming with collaborative teaching. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association. Anaheim, CA, USA: November 19–22.
- Mastropieri MA, Scruggs TE, Mantzicopoulos P, Sturgeon A, Goodwin L, Chung S (1998) A place where living things affect and depend on each other: qualitative and quantitative outcomes associated with inclusive science teaching. *Science Education* **82**: 163–179.
- McDonnell J, Thorson N, Disher S, Mathot-Buckner C, Mendel J, Ray L (2003) The achievement of students with developmental disabilities and their peers without disabilities in inclusive settings: an exploratory study. *Education and Treatment of Children* **26**: 224–236.
- Obrusnikova I, Valkova H, Block ME (2003) Impact of inclusion in general physical education on students without disabilities. *Adapted Physical Activity Quarterly* **20**: 230–245.
- Rankin DH, Logan KR, Adcock J, Angelucci J, Pittman C, Sexstone A, Straughn S (1999) Small group learning: effects of including a student with intellectual disabilities. *Journal of Developmental and Physical Disabilities* **11**: 159–177.
- Rarick GL, Beuter AC (1985) The effect of mainstreaming on the motor performance of mentally retarded and nonhandicapped students. *Adapted Physical Activity Quarterly* **2**: 277–282.
- Saint-Laurent L, Dionne J, Giasson J, Royer E, Simard C, Pierard B (1998) Academic achievement effects of an in-class service model on students with and without disabilities. *Exceptional Children* **64**: 239–253.
- Sasso G, Rude HA (1988) The social effects of integration on nonhandicapped children. *Education and Training in Mental Retardation* **23**: 18–23.
- Sharpe MN, York JL, Knight J (1994) Effects of inclusion on the academic performance of classmates without disabilities: a preliminary study. *Remedial and Special Education* **15**: 281–287.
- Shevlin M, O'Moore AM (2000) Creating opportunities for contact between mainstream pupils and their counterparts with learning difficulties. *British Journal of Special Education* **27**: 29–34.
- Shinn MR, Powell-Smith KA, Good III RH, Baker S (1997) The effects of reintegration into general education reading instruction for students with mild disabilities. *Exceptional Children* **64**: 59–79.
- Stevens RJ, Slavin RE (1995) The cooperative elementary school: effects on students' achievement, attitudes and social relations. *American Educational Research Journal* **32**: 321–351.
- Tapasak RC, Walther-Thomas CS (1999) Evaluation of a first-year inclusion programme: student perceptions and classroom performance. *Remedial and Special Education* **20**: 216–225.
- Willrodt K, Claybrook S (1995) Effects of inclusion on academic outcomes. Unpublished research report. Texas, USA: Sam Houston State University.

6.2 Other references used in the text of the report

Ainscow M (1991) Effective schools for all: an alternative approach to special needs in education. In: Ainscow M (ed.) *Effective Schools for All*. London: David Fulton, pages 1-19.

Ainscow M (1999) *Understanding the Development of Inclusive Schools*. London: Falmer.

Audit Commission (2002) *Special Educational Needs: A Mainstream Issue*. London: Audit Commission.

Ballard K (Ed.) (1999) *Inclusive Education: International Voices on Disability and Justice*. London: Falmer Press.

Banerji M, Dailey RA (1995) A study of the effects of an inclusion model on students with specific learning disabilities. *Journal of Learning Disabilities* **28**: 511–522.

Brantlinger E (1997) Using ideology: cases of non recognition of the politics of research and practice in special education. *Review of Educational Research* **67**: 425–459.

Department for Education and Skills (2001) *Special Educational Needs Code of Practice*. London: HMSO.

Dyson A (1999) Inclusion and inclusions: theories and discourses in inclusive education. In: Daniels H, Garner P (eds) *World Yearbook of Education 1999: Inclusive Education*. London: Kogan Page, pages 36-53.

Dyson A, Farrell P, Polat F, Hatcheson G, Gallanaugh F (2004) *Inclusion and Pupil Achievement*. London: DfES.

Dyson A, Howes A, Roberts B (2002) A systematic review of the effectiveness of school-level actions for promoting participation by all students. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Egelund N (2000) Special education in Denmark. *European Journal of Special Needs Education* **15**: 88–99.

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

EPPI-Centre (2002a) *Core Keywording Strategy: Data Collection for a Register of Educational Research (Version 0.95)*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

EPPI-Centre (2002b) *Guidelines for Extracting Data and Quality Assessing Primary Studies in Educational Research (Version 0.94)*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Evans J, Lunt I (2002) Inclusive education: are there limits? *European Journal of Special Needs Education* **17**: 1–14.

- Farrell P (2000) The impact of research on developments in inclusive education. *International Journal of Inclusive Education* **4**: 153–162.
- Farrell P, Ainscow M, Howes A, Franham J, Fox S, Davis P (2004) Inclusive education for all: dream or reality? *Journal of International Special Needs Education* **7**: 7–11.
- Farrell P, Ainscow M (eds) (2002) *Making Special Education Inclusive: From Research to Practice*. London: Fulton.
- Farrell P, Polat F (2003) The long term impact of residential provision for pupils with emotional and behavioural difficulties. *European Journal of Special Needs Education* **18**: 277–293.
- Farrell P, Tskalidou K (1999) Recent trends in the re-integration of pupils with emotional and behavioural difficulties. *School Psychology International* **20**: 323–338.
- Fox S, Farrell P, Davis P (2004) Factors affecting the successful inclusion of pupils with Down's syndrome in primary schools. *British Journal of Special Education* **31**: 184–191
- Harrington SA (1997) Full inclusion for students with learning disabilities: a review of the evidence. *The School Community Journal* **7**: 63–71.
- Harrower JK (1999) Educational inclusion of children with severe disabilities. *Journal of Positive Behaviour Interventions* **1**: 215–230.
- Hegarty S (1991) Towards an agenda for research in special education. *European Journal of Special Needs Education* **6**: 87–99.
- Howes A, Farrell P, Kaplan I, Moss S (2003) The impact of paid adult support on the participation and learning of pupils in mainstream schools. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.
- Jupp K (1992) *Everyone Belongs*. London: Souvenir Press.
- Lipsky DK, Gartner A (1997) *Inclusion and School Reform: Transforming America's Classrooms*. Baltimore: Paul H Brookes.
- Lunt I, Norwich B (1999) *Can Effective Schools Be Inclusive Schools?* London: Institute of Education, University of London.
- Manset G, Semmel MI (1997) Are inclusive programmes for students with mild disabilities effective? A comparative review of model programmes. *The Journal of Special Education* **31**: 155–180.
- Meijer C (1998) *Integration in Europe: Provision for Pupils with Special Educational Needs. Trends in 14 European Countries*. Middlefart: EADSNE.
- Mittler P (2000) *Working Towards Inclusion: Social Contexts*. London: David Fulton.
- Moore C, Gilbreath D, Maiuri F (1998) Educating students with disabilities in general education classrooms: a summary of the research. Unpublished research

- report. Eugene, Oregon, USA: Western Regional Resource Centre [ERIC: ED419329].
- Norwich B (2002) *LEA Inclusion Trends in England 1997-2001, Statistics on Special School Placements and Pupils with Statements in Special Schools*. Bristol: CSIE.
- Ofsted (2000) *Educational Inclusion: Guidance for Inspectors and Schools*. London: Ofsted.
- Peltier GL (1997) The effect of inclusion on non-disabled children: a review of the research. *Contemporary Education* **68**: 234–238.
- Petch-Hogan B, Haggard D (1999) The inclusion debate continues. *Kappa Delta Pi Record* **35**: 128–131.
- Salend SJ, Duhaney LMG (1999) The impact of inclusion on students with and without disabilities and their educators. *Remedial and Special Education* **20**: 114–126.
- Skrtic TM (1991) The special education paradox: equity as the way to excellence. *Harvard Educational Review* **61**: 148–206.
- Staub D (1996) Inclusion and the other kids. *Learning* **25**: 76–78.
- Staub D (1999) On inclusion and the other kids: Here's what research shows so far about inclusion's effect on nondisabled students. Unpublished report. Newton, MA, USA: Education Development Centre, National Institute for Urban School Improvement.
- Staub D, Peck CA (1994/1995) What are the outcomes for nondisabled students? *Educational Leadership* **52**: 36–40.
- Thomas G, Vaughn M (2004) *Inclusive Education: Readings and Reflections*. Maidenhead: Open University Press.
- UNESCO (1994) *The Salamanca Statement and Framework for Action on Special Needs Education*. Paris: UNESCO.

Appendix 1.1: Advisory Group membership

Members of the Educational Support and Inclusion Group, School of Education

Professor Mel Ainscow
Ms Alkmini Lachana
Dr Andrew Howes
Dr Clare Lennie
Dr Filiz Polat
Dr Neil Humphrey
Mr Peter Hick
Mr Peter Jenkins
Ms Sarah Rooney

Local education authority (LEA)

Paul Rees, Assistant Principal Educational Psychologist, South Glamorgan
Mike O'Connor, Former SEN adviser, Cheshire

Editor of professional journal

Dr Sue Ralph, Editor, *Journal of Research in Special Educational Needs*

Appendix 2.1: Inclusion and exclusion criteria

Inclusion criteria

In order to be included, a study should meet all seven criteria described below:

- The study is written in English.
- The study reports on the results of empirical research rather than being purely theoretical or exhortatory.
- The study is concerned with the phases of compulsory schooling.
- The study reports the outcomes of the 'intervention' (i.e. an increase in population inclusivity).
- The study reports these outcomes in relation to students without SEN or whole school populations (but not simply in relation to students with SEN).
- The study reports robust evidence of the impact of the intervention:
 - through a longitudinal study of one school, or
 - by comparison with a similar but less inclusive school (with a lower level of population inclusivity), or
 - by comparison between different conditions within the same school (such as more and less inclusive classes), or
 - by some other equally robust means
- The study is concerned with the impact of inclusion on pupils' personal, social and/or academic outcomes.

Exclusion criteria

Each study is screened by applying the following exclusion criteria in the following order from 1 to 7. A study is excluded when one (or more) of the following exclusion criteria applies:

- The study is not concerned with the impact of population inclusivity (exclusion criterion 1).
- The impact of population inclusivity is not on children without SEN (exclusion criterion 2).
- There is no reference to the impact on personal, social and/or academic outcomes (exclusion criterion 3).
- The study is not concerned with the compulsory phases of schooling (exclusion criterion 4).
- The study is not empirical research (exclusion criterion 5).
- The study is not written in English (exclusion criterion 6).

- The study does not provide evidence of robust measurements (e.g. reports pure attitudes of teachers, etc.) (exclusion criterion 7).

Cut-off date

The Review Group decided not to set a specific cut-off date, in the knowledge that earlier studies would be rather limited. In addition, given that policies vary between countries and the review focuses on studies that include literature from the USA and Australia, it was thought likely that studies should be included irrespective of when they were carried out.

Appendix 2.2: Search strategy for electronic databases

Databases searched

Australian Education Index 1976–December 2003
ASSIA: Applied Social Sciences Index and Abstracts
British Education Index 1976–December 2003
ERIC: CIJE & RIE 1966–1983
ERIC: CIJE & RIE 1984–1989
ERIC: CIJE & RIE 1990–September 2003
PsycINFO 1967 – 2004
Social Sciences Citation Index (ISI Web of Science)
ZETOC: Electronic Table of Contents

ERIC search strategy

Searched via Dialog@Site from 1966 to September 2003

NB1: All the terms refer to free text term searched for in title or abstract.

NB2: The wildcard character, a question mark (?) replacing the ending of a term, is used to find variations of a root word. For example, entering INTEGRAT? would retrieve records for INTEGRATION, INTEGRATED and INTEGRATING, etc. Also, the proximity search technique was used to delineate the distance between search terms. For example, PERSONAL..DEVELOPMENT, where the term DEVELOPMENT follows within two words of the term PERSONAL. This search would retrieve PERSONAL AND SOCIAL DEVELOPMENT and so on.

#A Inclusion

INTEGRAT? OR (MAINSTREAM?) OR (INCLUD? GENERAL CLASS?) OR (INCLUD? GENERAL EDUCATION) OR (INCLUS?)

#B Education

PUPIL? OR (STUDENT?) OR (CLASS?) OR (EDUCAT?) OR (SCHOOL?)

#C Achievement

SOCIAL .. DEVELOPMENT OR (ACADEMIC .. DEVELOPMENT) OR (PERSONAL .. DEVELOPMENT) OR (ATTAINMENT?) OR (ACADEMIC RESULT?) OR (LEARN? RESULT?) OR (SOCIAL RESULT?) OR (CLASS? RESULT?) OR (STUDENT? RESULT?) OR (PUPIL? RESULT?) OR (TEST? RESULT?) OR (EDUCATION? RESULT?) OR (SCHOOL? RESULT?) OR (ACHIEV?) OR (OUTCOME?) OR (EFFECTS) OR (EFFECT) OR (IMPACT) OR (TEST? PERFORMANCE) OR (ACADEMIC PERFORMANCE) OR (LEARN? PERFORMANCE) OR (CLASS? PERFORMANCE) OR (STUDENT? PERFORMANCE) OR (PUPIL? PERFORMANCE) OR (SCHOOL? PERFORMANCE) OR (EDUCATION? PERFORMANCE)

#D Population

SEN OR (BEHAVIOURAL PROBLEMS) OR (BEHAVIOURAL PROBLEMS) OR (SYNDROME) OR (ASPERG?) OR (AUTIS?) OR (CEREBRAL .. PALS?) OR (BLIND) OR (DEAF) OR (SPECIAL NEEDS) OR (RETARDAT?) OR (IMPAIRM?) OR (HANDICAP?) OR (DISORD?) OR (DISAB?) OR (DIFFICULT?)

Final results: #A AND #B AND #C AND #D (=3396 studies for ERIC between 1990 and September 2003)

Table 2.2.1: Example of terms used in relation to inclusion, SEN and impact: ERIC – CIJE & RIE

Inclusion	No.	Education	No.	Outcomes	No.	Population	No.
Inclus?	7839	School?	170982	Impact	21903	Difficult?	17538
Includ? general education	81	Educat?	352101	Effect	14795	Disab?	27878
Includ? general class?	8	Class?	88794	Effects	25790	Disord?	8949
Mainstream?	5339	Student?	206817	Outcome?	26399	Handicap?	4557
Integrat?	30871	Pupil?	4388	Achiev?	44731	Impairm?	6224
				Education? performance	316	Retardat?	4229
				School? performance	1928	Special needs	6052
	40054 ↑		377306 ↑	Pupil? performance	86	Deaf	2544
				Student? performance	6148	Blind	1444
				Class? performance	536	Cerebral .. pals?	268
				Learn? performance	850	Autis?	1723
				Academic? performance	2251	Asperg?	153
				Test? performance	1045	Syndrome	3703
				Social result?	160	Behavioural problems	494
				School? Result?	993	Behavioural problems	7
				Education? result?	916	SEN	49
				Test? result?	3326		
				Pupil? result?	16		61569 ↑
				Student? result?	1219		
				Class? result?	364		
				Learn? result?	598		
				Academic result?	174		
				Attainment	4710		
				Personal .. development	840		
				Social .. development	3224		
				Academic .. development	269		
					118053 ↑		

ISI Web of Sciences

NB1: The wildcard character, asterisk (*) replacing the ending of a term, is used to find variations of a root word. For example, entering DISABILIT* would retrieve records for DISABILITY and DISABILITIES, etc.

TS=((SEN OR behavioural problems OR syndrome OR asperger* OR autism OR cerebral palsy OR blind OR deaf OR special needs OR retardation OR imparment OR handicap OR disorder OR disabilit* OR difficult*)AND (inclusion OR mainstream OR integration) AND (pupil OR student OR class OR education OR school) AND (social development OR academic development OR personal development OR attainment OR academic results OR test results OR educational results OR school results OR achievement OR outcome OR effects OR impact OR test performance OR academic performance OR learning performance OR class performance OR student performance OR pupil performance OR school performance OR educational performance))

Appendix 2.3: Journals handsearched

The majority of the journals available for handsearching were either available online or their contents were included in the databases that were searched by the Review Group. Therefore, an initial search of journals, such as Remedial and Special Education, the European Journal of Special Needs Education and the International Journal of Inclusive Education, did not produce any results of studies that had not already been retrieved using the electronic databases.

In addition, the contents of some of the journals that are considered to be key journals on the area and/or they often contained studies marked as 'potential includes' were also handsearched or searched online. More specifically, the contents of the following journals were handsearched:

European Journal of Special Needs Education (online: 1997–2004)

Exceptional Children (online: 1990–2004)

Focus on Autism and Other Developmental Disabilities (online: 2000–2004)

International Journal of Inclusive Education (hard copies)

Remedial and Special Education (online: 2004)

The Journal of Learning Disabilities

Finally, a researcher spent approximately half a day in the offices of two of the Professors in Inclusive Education where she went through and handsearched all material related to inclusion.

Appendix 2.4: EPPI-Centre keyword sheet, including review-specific keywords

V0.9.7 Bibliographic details and/or unique identifier

<p>A1. Identification of report Citation Contact Handsearch Unknown Electronic database (Please specify.)</p> <p>A2. Status Published In press Unpublished</p> <p>A3. Linked reports <i>Is this report linked to one or more other reports in such a way that they also report the same study?</i></p> <p>Not linked Linked (Please provide bibliographical details and/or unique identifier.) </p> <p>A4. Language (Please specify.) </p> <p>A5. In which country/countries was the study carried out? (Please specify.) </p>	<p>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.)</p> <p>A7. Curriculum Art Business studies Citizenship Cross-curricular Design and technology Environment General Geography Hidden History ICT Literacy – first language Literacy further languages Literature Mathematics Music PSE Physical education Religious education Science Vocational Other (Please specify.)</p>	<p>A8. Programme name (Please specify.) </p> <p>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.)</p> <p>A10. Age of learners (years) 0–4 5–10 11–16 17–20 21 and over</p> <p>A11. Sex of learners Female only Male only Mixed sex</p>	<p>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.)</p> <p>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</p>
--	---	---	--

Review-specific keywords

A.1 Are any learners described as having severe, complex or profound needs?

A.1.1 Yes

A.1.2 No

A.2 Type of special educational needs (*Please use as many keywords as apply.*)

A.2.1 Cognition and learning

A.2.2 Communication and interaction

A.2.3 Behavioural, emotional and social development

A.2.4 Sensory and/or physical needs

A.2.5 Not specified

A.3 The study reports impact on: (*Please use as many keywords as apply.*)

A.3.1 Academic outcomes

A.3.2 Other (e.g. personal, social, etc.)

A.4 Does this study report outcomes on: (*Please use as many keywords as apply.*)

A.4.1 Learners with SEN

A.4.2 General population (learners without SEN)

Appendix 3: Details of studies included in the in-depth review

Item	Study outcomes	What type of SEN do pupils have?	Are any learners described as having severe, complex or profound needs?	What is the comparative element of the study?	What external or standardised data are provided as evidence of the outcomes?	Does this study report outcomes on students with and/or without SEN?	What are the key findings of the study?	What evidence of inclusivity is provided?
Affleck <i>et al.</i> (1988)	Academic	Cognition and learning EBD	No	Between groups	National tests	With and without SEN	Neutral	Number of SEN Severity of SEN Hours of inclusion
Bear <i>et al.</i> (1991)	Social and personal	Cognition and learning	No	Between groups	Attributive measures	With and without SEN	Neutral	Number of SEN Severity of SEN Hours of inclusion
Beuter (1984)	Academic Other	Cognition and learning	No	Between groups	Observational (research)	With and without SEN	Neutral Negative	Number of SEN Hours of inclusion
Block and Zeman (1996)	Academic Social and personal	Cognition and learning		Other class	Questionnaire Other	Without SEN	Neutral Negative	Severity of SEN
Brown (1982)	Academic Social and personal	Cognition and learning EBD	No	Between groups	Group tests Interviews Questionnaire Ethnographic	With and without SEN	Negative	Number of SEN
Cawley <i>et al.</i> (2002)	Academic Social and personal Other	Cognition and learning EBD	Yes	Other class	National tests Observational (teach) Reports Other	With and without SEN	Mixed Negative	Hours of inclusion

Item	Study outcomes	What type of SEN do pupils have?	Are any learners described as having severe, complex or profound needs?	What is the comparative element of the study?	What external or standardised data are provided as evidence of the outcomes?	Does this study report outcomes on students with and/or without SEN?	What are the key findings of the study?	What evidence of inclusivity is provided?
Daniel and King (1997)	Academic Social and personal	Cognition and learning Sensory and/or physical Communication and interaction		Between groups	National tests Questionnaire Reports	With and without SEN	Mixed Negative	Hours of inclusion
Helmstetter <i>et al.</i> (1994)	Social and personal	Cognition and learning	Yes	Other	Questionnaire	Without SEN	Mixed Neutral	Severity of SEN Other
Hepler (1998)	Social and personal	EBD	No	Between groups Different conditions	Observational (research) Questionnaire Sociometric	With and without SEN	Positive	Number of SEN Severity of SEN Hours of inclusion
Hillen <i>et al.</i> (1992)	Academic Social and personal	Cognition and learning	No	Different conditions	Observational (research)	Without SEN	Positive Neutral	Severity of SEN Hours of inclusion
Huber <i>et al.</i> (2001)	Academic	Cognition and learning Sensory and/or physical EBD	No	Between groups	Other	Without SEN	Neutral Mixed	Number of SEN
Hunt <i>et al.</i> (1994)	Academic	Cognition and learning Sensory and/or physical Communication and interaction	Yes	Between groups	Teacher assessment	With and without SEN	Neutral	Number of SEN Severity of SEN

Item	Study outcomes	What type of SEN do pupils have?	Are any learners described as having severe, complex or profound needs?	What is the comparative element of the study?	What external or standardised data are provided as evidence of the outcomes?	Does this study report outcomes on students with and/or without SEN?	What are the key findings of the study?	What evidence of inclusivity is provided?
Lundeen and Lundeen (1993)	Academic	Cognition and learning Sensory and/or physical EBD	No	Different conditions	Reports	With and without SEN	Positive Neutral	Number of SEN Hours of inclusion
Mastropieri <i>et al.</i> (1998)	Academic Other	Cognition and learning Sensory and/or physical EBD	No	Other class	Other	Without SEN	Positive	Number of SEN Severity of SEN
McDonnell <i>et al.</i> (2003)	Academic	Cognition and learning Sensory and/or physical Communication and interaction	Yes	Other class Between groups	National tests	With and without SEN	Neutral	Hours of inclusion
Obrusnikova <i>et al.</i> (2003)	Academic	Sensory and/or physical	No	Other class	Attributive measures Observational (research) Other	Without SEN	Neutral	Number of SEN
Rankin <i>et al.</i> (1999).	Academic	Cognition and learning Sensory and/or physical Communication and interaction	Yes	Different conditions Other	Observational (research) Other	With and without SEN	Positive Neutral	Number of SEN Severity of SEN
Rarick and Beuter (1985)	Academic	Cognition and learning	No	Between groups	Observational (research) Other	With and without SEN	Neutral	Other

Item	Study outcomes	What type of SEN do pupils have?	Are any learners described as having severe, complex or profound needs?	What is the comparative element of the study?	What external or standardised data are provided as evidence of the outcomes?	Does this study report outcomes on students with and/or without SEN?	What are the key findings of the study?	What evidence of inclusivity is provided?
Saint-Laurent <i>et al.</i> (1998)	Academic	Cognition and learning Sensory and/or physical Communication and interaction EBD	No	Between groups	Other	Without SEN	Positive	Other
Sasso and Rude (1988)	Social and personal	Cognition and learning Communication and interaction EBD	Yes	Other class	Sociometric	Without SEN	Positive Neutral	Severity of SEN Hours of inclusion Other
Sharpe <i>et al.</i> (1994)	Academic Social and personal	Cognition and learning EBD	Yes	Other class Longitudinal	Group tests Reports Other	Without SEN	Neutral	Severity of SEN Hours of inclusion
Shevlin and O'Moore (2000)	Social and personal	Cognition and learning	Yes	Longitudinal	Interviews	Without SEN	Positive	Number of SEN Severity of SEN Hours of inclusion
Shinn <i>et al.</i> (1997)	Academic	Cognition and learning Sensory and/or physical Communication and interaction	No	Longitudinal	Teacher assessment Teacher rating scales Reports Interviews Other	With and without SEN	Positive	Number of SEN
Stevens and Slavin (1995)	Academic Social and personal	Cognition and learning	No	Other school	Questionnaire Other	With and without SEN	Positive	Number of SEN Hours of inclusion

Item	Study outcomes	What type of SEN do pupils have?	Are any learners described as having severe, complex or profound needs?	What is the comparative element of the study?	What external or standardised data are provided as evidence of the outcomes?	Does this study report outcomes on students with and/or without SEN?	What are the key findings of the study?	What evidence of inclusivity is provided?
Tapasak and Walther-Thomas (1999)	Academic Social and personal Other	Cognition and learning Sensory and/or physical Communication and interaction EBD	No	Longitudinal	Teacher assessment Attributive measures Reports Sociometric	With and without SEN	Neutral	Number of SEN Severity of SEN Hours of inclusion
Willrodt. and Claybrook (1995)	Academic	Cognition and learning Sensory and/or physical Communication and interaction EBD	No	Other school	National tests	With and without SEN	Neutral	Other