EPPI CENTRE

IOE LONDON
Leading education and social research
Institute of Education
University of London

What is the effect of block scheduling on academic achievement?

A systematic review

Technical report written by Kelly Dickson, Karen Bird, Mark Newman and Naira Kalra

EPPI-Centre
Social Science Research Unit
Institute of Education
University of London

EPPI-Centre report no. 1802T · February 2010

SSRU Social Science Research Unit

TECHNICAL REPORT

What is the effect of block scheduling on academic achievement?

A systematic review

TECHNICAL REPORT

Review conducted working with support staff

Report by Kelly Dickson (EPPI-Centre)
Karen Bird (EPPI-Centre)
Mark Newman (EPPI-Centre)
Naira Kalra (EPPI-Centre)

The results of this systematic review are available in four formats. See over page for details.

The results of this systematic review are available in four formats:

SUMMARY	Explains the purpose of the review and the main messages from the research evidence
REPORT	Describes the background and the findings of the review(s) but without full technical details of the methods used
TECHNICAL REPORT	**Includes the background, main findings, and full technical details of the review**
DATABASES	Access to codings describing each research study included in the review

These can be downloaded or accessed at **http://eppi.ioe.ac.uk/reel/**

The EPPI-Centre reference number for this report is 1802T.

This report should be cited as: Dickson K, Bird K, Newman M, Kalra N (2010) What is the effect of block scheduling on academic achievement? A systematic review. Technical Report. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

ISBN: 978-1-907345-02-9

CONTENTS

List of abbreviations

DCSF	Department for Children, Schools and Families
EPPI-Centre	Evidence for Policy and Practice Information and Co-ordinating Centre
GPA	Grade Point Average
QAT	Quality Assessment Tool
SAT	Standard Attainment Test
TAAS Test	Texas Assessment of Academic Skills Test
WoE	Weight of evidence

Abstract

What do we want to know?

Block scheduling is one approach to school scheduling. It typically means that students have fewer classes (4-5) per day, for a longer period of time (70-90 minutes). There are three main types of block schedule investigated in this review, comprising the following:

4 x 4: four blocks of 80–90 minute classes in one day, with students taking four subjects in one term

A/B: classes of 70-90 minutes each for 3/4 different subjects on every alternating day

hybrid: five classes per day, between 55 and 90 minutes in length

The in-depth review asks the following:

Does block scheduling result in higher levels of student attainment than traditional scheduling?

Studies used different measures of academic achievement across different academic subjects. These included test results in Mathematics, English, Science, exam scores or average grade scores across different subjects.

Sub-questions were also asked in the in-depth review and these investigated whether the effect of block scheduling varied by type of block schedule and type of subject(s) taught.

Who wants to know and why?

Interested parties include policy-makers and schools interested in whether teaching subjects in extended 'blocks' of time will improve achievement at Key Stages 3 and 4 in the National Curriculum.

What did we find?

Only 12 of the 14 studies included in the in-depth review provided the data necessary for statistical meta-analysis to assess the effectiveness of different types of block scheduling on academic achievement. The 12 studies were considered to be of medium weight of evidence and two were considered to be of low weight of evidence, overall, for this review.

Where we were able to combine data to produce summary effect sizes, we found that 4 x 4 block scheduling resulted in higher cross-subject achievement than traditional schedules. However, the outcome average cross-subject achievement could conceal worsening performance in some subjects and better performance in others.

For single subject outcomes:

In Science, A/B block scheduling resulted in higher results than traditional schedules.

In Mathematics and English, the evidence was unclear, with studies showing both better and worse results for block scheduling compared with traditional scheduling.

What are the implications?

There is not conclusive evidence in this review to support the introduction of policy guidance on the use of block scheduling in secondary schools. Findings do not indicate that participating in block schedules would produce negative outcomes for pupils across subjects, but the findings on positive effects are not strong enough to recommend their implementation.

How did we get these results?

We searched six key educational bibliographic databases and seven key websites. We applied

inclusion and exclusion criteria to build up a ‘map’ of relevant studies. Additional criteria were applied to the studies in the map, which produced the 12 studies that were synthesised to answer the in-depth review questions.

Where to find further information

http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2476

CHAPTER ONE

Background

1.1 Aims and rationale for current review

The organisation of school and how time should be spent during the school day has been under discussion since the education system came into existence. Questions about the optimal length of the school year or school day, and how much time is afforded to which subjects continue to be asked by educational policy-makers and educational professionals. This review has been commissioned by the Department for Children, Schools and Families (DCSF) with the initial aim of identifying research evidence on the structure of the school day, and specifically how time is used in schools. As the review developed in consultation with the DCSF, the specific area of block scheduling was chosen to study in depth.

1.2 Definitional and conceptual issues

Given the broad nature of the review topic, 'structure of the school day', a systematic map question was developed through discussion with the DCSF. Key definitions were developed to help define the scope of the review. These definitions were used to devise the search strategy and inform the inclusion exclusion criteria.

1.2.1 Key definitions

Structure of the school day was conceptualised as how time is used in schools to organise pupil learning and activities, and how time can influence learning. It included the following potential areas that we could search for and systematic 'map' the research literature on:

Length of the school year: For this review, the organisation of the school year refers to how many days are included in the school year, when the school year starts and how terms are organised within the school year.

Length of the school day: The length of the school day is concerned with the number of hours pupils are expected to attend school and the start times of schools. It is not, however, about extended school activities which fall outside the scope of this review.

Length of lessons: This is primarily about the time allocated to different subjects either during the school day or across the school term/year.

Break times: This refers to how non-lesson time is organised in the school day.

Multiple-shift schooling: This refers to when pupils attend school – for example, one cohort of students attends in the morning and a different cohort in the afternoon.

Organisation of examinations: The review is interested in studies which look at when to set exams in the school day (i.e. morning or evening) and across the school year (i.e. in the autumn, spring or summer term).

Time of day: This is concerned with when subjects are taught (e.g. placing certain subjects in specific timeslots) and student preferences for the time at which teaching should take place (e.g. whether take place am or pm).

1.2.2 Key outcomes

The review was interested in the following pupil outcomes:

Attainment: Specific academic achievement, context and country specific (e.g., GCSEs, 'A' Levels, High School diploma, Standard Attainment Tests (SATS))

Academic skills: Ability and acquisition of skills (e.g. reading, writing, Mathematics, Science, sport etc.

Learning: The experience of learning, including attention, cognition and memory

Personal and social skills: Acquisition of personal and social skills

Pupils' views: For example, a pupil's satisfaction with, or perception of, an intervention

Schools: School climate, staffing, overall school performance

1.3 Policy and practice background

There are three key policy background issues relevant to this review:

- The requirement for progression in pupil achievement across all key stages
- Key Stage 3 flexible curriculum and personalised learning agenda
- Standardised school year

1.3.1 Progression in pupil achievement

In England and Wales, there has been a drive to raise standards of achievement at all Key Stages since the introduction of the National Curriculum in 1988. The Children's Plan (DCSF, 2007) included an ambitious set of goals for 2020, with particular focus on improving educational achievement. At Key Stage 2, 60% of pupils are expected to achieve level 4 in both English and Mathematics (i.e. three out of five pupils). The aim is that from Key Stage 2 onwards, local authorities set educational targets that indicate how they will improve the proportion of pupils making progress through each Key Stage level. For example, all pupils achieving level 4 in English or Mathematics at the end of Key Stage 2 are expected to be capable of progressing to level 5. In addition, all pupils averaging level 6 or above in English and Mathematics, and 30% of those averaging level 5 at the end of Key Stage 3, are expected to be capable of achieving five A*-C grades at GCSE and equivalent, including English and Mathematics. Any schools which currently set their targets below 30% at any key stage will be required to set more ambitious targets for 2010 and 2011. Targets for children in care, Black and Ethnic Minority groups, and pupils with special educational needs and learning difficulties and disabilities will also need to be set and show improvement and progress in order to raise standards for these groups of children (DCSF, 2008a, 2008b).

The emphasis on ensuring pupils progress through each Key Stage to increase their educational attainment has implications for all aspects of teaching and learning, the organisation of the curriculum in schools, and the personalised learning agenda.

1.3.2 Flexible curriculum and personalised learning

Individual schools in England and Wales have responsibility for designing, organising and timetabling the curriculum, and have a choice about which year pupils take their National Curriculum end of Key-Stage tests. Schools can make decisions regarding which subjects they prioritise, the amount of time allocated to each subject, the number of lessons in the school day, the number of terms in a school year (with consultation with the LEA), and how the curriculum is taught across the key stages (DfES, 2002). The DES Circular 7/90 simply recommends that a teaching week includes at least 24 hours at Key Stage 3 and 23.5 hours at Key Stage 2. There are no constraints on the way that the National Curriculum subjects can be distributed or timetabled across these Key Stages. The only requirement is that the programme of study for each subject is completed by the end of the key stage. Thus, designing a curriculum framework in secondary schools allows individual schools to be creative about how the school day is structured (DfES, 2004). This has implications for deciding the length of the school day and how much time is spent on different subject areas of the curriculum. Designing a flexible curriculum links with the personalised learning agenda because it encourages schools to think resourcefully about how each school is best organised to meet the academic and welfare needs of pupils.

The secondary national strategy for school improvement produced guidance on delivering a condensed Key Stage 3 curriculum (DfES, 2004). This allows schools to adjust the proportion of teaching timing for different subjects when required to support students at different levels of attainment and rates to progress through Key Stage 3 and into Key Stage 4. The 2006 update of this policy reports on the ongoing evaluation of a two-year Key Stage 3 condensed curriculum project (DfES, 2006). They found that, when some schools plan a programme of study, they look to remove repetition and duplication of subject matter in the curriculum in order to make the best use of time available in the school day, across school terms and across the school years, rather than extending the total amount of school time pupils spend in learning a topic or subject.

1.3.3 Standardised school year

Although schools have a certain amount of autonomy when organising the school day, it is the responsibly of each local educational authority to standardise terms and holidays nationally (Eurybase, 2008). There has been discussion about the length of the school year and how many terms it should include to maintain teaching momentum and maximise pupil learning. In 2001, the Local Government Association independent commission into the organisation of the

school year proposed that the current three-term, six half-term structure be modified into six terms of six or seven weeks each, with exam results being published at the end of term six and the summer holidays remaining at five to six weeks long but brought forward to include the whole of July and first half of August (Independent Commission on the Organisation of the School Year, 2002). By 2004, there was an agreement to have a standardised arrangement of school terms; however, not all English schools and local authorities have signed up to the proposals. Although the majority of schools continue to start the school year as near to the first of September as possible and have six weeks holidays in July and August, they do not all agree to equalising learning blocks, or to establishing a spring break in April regardless of Easter (LGA, 2004, 2007).

1.4 Research background

1.4.1 Length of lessons

The traditional school day mostly involves six to eight periods/classes in a day where each class/period is for approximately around 50 minutes (Scroggins and Karr-Kidwell, 1995). Research in the US on the length of lesson has mostly looked at block scheduling, which is an alteration to the traditional structure and duration of lessons in a school day (Trenta and Newman, 2001). One of the cited aims of block scheduling is to allow greater time to be focused on student-oriented activities in order to promote in-depth discussion and increased interaction, and hence deeper thinking by providing the time for discussion. Teachers are said to be more likely to engage in learning-oriented activities and to use a variety of teaching strategies.

1.4.2 Length of the school day

Much of the research on the duration and length of the school day has questioned whether the half day or full day at school is more beneficial for children attending kindergarten/nursery (aged 5 and 6). A full day at school is often preferred by parents as this reduces the number of changes that a child is exposed to in a day. The argument in favour of half-day school often relies on the reduced cost associated with it (Rothenberg, 1995). Some research concludes that, if all things are kept equal, there are no differences between students attending full-day or half-day kindergarten/nursery in the outcomes such as developmental gains, attendance and quality of curriculum (Nunnelley, 1996). Other longitudinal research comparing the two finds that full-day is more beneficial and that students in full-day kindergarten programmes show more positive behaviour, greater classroom involvement and better academic performance at later stages in life (Rothenberg, 1995).

1.4.3 Length/structure of the school year

Year Round Education/Schooling (YRE)

Internationally, Year Round Education can be described as an alteration to the traditional school calendar in a manner that permits 'continuous education' and short but frequent breaks (Worthen and Zsiray, 1994). The main rationale behind this design is that shorter breaks would improve retention and provide a more efficient teaching system. There are various types of YRE schedules. The traditional schedule involves a 45 school day period followed by a 15-day vacation period. These can be single track (where all students attend school on the same day) or multi-track programmes (staggered attendance). Besides this, there are also various other variations to the YRE programme and hence it may not be possible to generalise the findings from one design to the other.

Overall, YRE has had positive results in terms of increasing student attendance rates; it has a positive impact on student and teacher attitudes and behaviour; and students maintain or improve their academic performance (Shields and LaRocque, 1996). At the same time, YRE can cause inconvenience to friendships when friends are in different tracks and can make co-ordinating holidays with the rest of the family more difficult. Single track YRE programmes can cost more than traditional programmes, while multi-track programmes can save costs but only at the district level (Shields and LaRocque, 1996; Worthen and Zsiray, 1994).

Extended Year Schedule (EYS)

This design is often labelled as Year Round Education. However, the Extended Year Schedule is different from YRE in its aim. In the US, EYS aims to increase the number of school days in a year from the traditional 180 to 220/240 as in some other parts of the world (Worthen and Zsiray, 1994). However, despite the additional number of school attendance days, there is little support from studies, mainly conducted in the US, that doing so would improve academic achievement. Even if performance does increase, the additional cost may not be worth the benefits (Worthen and Zsiray, 1994).

1.4.4 Time of day

Research suggests that students have preferences for the time of day in which they learn (Callan, 1998). Studies indicate that such preferences vary according to age (Klein, 2001; Wheeler, 1995), academic ability (Milgram et al. 1993), and ethnicity (Dunn and Griggs, 1990; Lam-Phoon, 1986). These studies suggest that matching the time of day to student preferences for learning can improve pupil achievement, behaviour and attendance (Ammons et al., 1995; Harp and Orsak,

1990; Lynch, 1981; Virostko, 1983). As part of this body of research, some studies have considered whether core subjects, such as Mathematics and Literacy, should be taught at particular times of day (Klein, 2001; Sjosten-Bell, 2005), or in a specific order (Engin, 2006).

1.5 Authors, funders and other users of the review

The Review Group comprises members of staff from the EPPI-Centre. The team members have experience of undertaking systematic reviews in education and social policy, and teaching the theory and methods of systematic reviews in education.

The Social Science Research Unit's Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), based at the University of London's Institute of Education, is acknowledged as a centre of excellence for conducting secondary research of direct relevance to policy-makers in the United Kingdom and beyond. It has pioneered the development of systematic review methods for 'social interventions' since 1993 and is also a formal partner of the Campbell Collaboration.

1.6 Review questions and approach

This review seeks to address these gaps by assessing the UK and international evidence on the impact of the use of time in schools. Using systematic review methodology to determine what the evidence says with regard to the use of time in schools, the review will therefore focus on addressing the following key question:

What research has been undertaken on the use and influence of time in schools?

CHAPTER TWO

Methods used in the review

2.1 Type of review

A two-stage review model was adopted. The first stage consisted of identifying all studies that met the review inclusion criteria. Descriptive information about these studies was collected and presented in the form of a 'map' of research in the field on 'the use and influence of time in schools'. The map provides a basis for informed discussion and decision making between the Review Group and review users about the focus of the second stage in-depth review. The in-depth review provides a detailed investigation of a more focused subset of the wider literature. The review can be focused in a number of different ways that can correspond to particular policy or practice priorities. As the in-depth review was focused and narrowed down to look at block scheduling, a second set of inclusion criteria was developed from the in-depth review question and applied to the studies initially identified in the map. Detailed data extraction has been undertaken at this stage to facilitate synthesis of the findings of the selected studies in order to provide answers to the in-depth review question(s).

2.2 User involvement

2.2.1 Approach and methods used

The review has been informed by the commissioners and relevant policy-makers at the DCSF and two substantive topic specialists acting as project consultants. Both have played a key role in informing the progress of the review at two points in the review process:

April 2008: Scope of the review, including (a) specifications of the review questions and (b) draft specifications of the inclusion exclusion criteria

August 2008: Interim report: moving from the systematic map to the in-depth review.

An initial meeting was set up with the DCSF in April to discuss the scope of the review. Subsequent to this meeting, the majority of communication with the DCSF and the project consultants was conducted via email.

2.3 Identifying and describing studies

2.3.1 Defining relevant studies: inclusion and exclusion criteria

The exclusion and inclusion criteria were applied to the papers identified using the search strategy. This focused the search process and ensured that only relevant papers were described (mapping) and reviewed. Any papers that were not excluded, after applying the criteria, were requested for review.

Studies were excluded if they met the following criteria:

- Not on the structure of time within the school day or year
- Not reporting on children and young people aged 5 to 16
- Not on mainstream, maintained and independent school
- Of the following study type:
  - Descriptive
  - Methodology
  - editorial, commentary, book review
  - policy document
  - resource, textbook
  - bibliography

- theoretical paper
- position paper
- reviews (systematic and non-systematic)

• Not reporting data on pupils or the school

• Not published or reported in English

• Not published or reported between 1988 and the present

The inclusion and exclusion criteria can be found in Appendix 2.1.

2.3.2 Identification of potential studies: search strategy

Key search terms were determined by the review question and the inclusion and inclusion criteria, and these were supplemented by 'pearl growing' further key search terms from papers identified through handsearching.

Journal articles: Searches were undertaken using a wide range of electronic bibliographic databases.

Handsearching of print sources (e.g. relevant journals and textbooks)

Searching of 'grey' literature through databases, conference proceedings and research-funders

The search strategy can be found in Appendix 2.2.

The EPPI-Centre's specialist web-based systematic review software, EPPI-Reviewer, was used to keep track of and code studies found during the review.

2.3.3 Screening studies: applying inclusion and exclusion criteria

Inclusion and exclusion criteria have been applied successively to (i) titles and abstracts, and (ii) full reports. Full reports were obtained for those studies that appeared to meet the criteria or where we had insufficient information to be sure. These reports were entered into a second database. The inclusion and exclusion criteria were re-applied to the full reports and those that did not meet these initial criteria were excluded.

2.3.4 Characterising included studies

Included studies were coded for contextual (i.e. characteristics of the intervention and population) and methodological information using section A to E of the EPPI-Centre Data Extraction and Coding Tool for Education Studies V2.0. There are two levels of coding for data extraction. The first level for all studies included in the map provides data for the purposes of describing or mapping the overall field of research on the topic area. The second level of coding provides detailed information about studies included in the in-depth review necessary for the purpose of description, quality assessment and synthesis. Additional coding, which was specific to the context of the review, was added to those of the EPPI-Centre. All the coded studies have been added to the larger EPPI-Centre database, REEL, for others to access via the website.

2.3.5 Identifying and describing studies: quality-assurance process

All team members involved in screening and coding took part in moderation exercises where results were discussed to ensure consistency in interpretations of the review inclusion criteria and the coding tool. For the initial title and abstract screening, double screening was carried out on 500 papers and, for the second round of screening, on full reports, the criteria were independently applied by a second screener to 20 percent of the reports. A rate of agreement of 90 percent was required before proceeding to independent screening. The remaining sample of potential includes were then screened independently by single reviewers on EPPI-Reviewer. The team leader also carried out independent audits of each team member's screening decisions and coding on a random sample of papers. Where a reviewer was unable to reach a decision, consensus was reached through discussion with the team leader, and occasionally a third team member.

2.4 In-depth review

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

The mapping exercise identified many studies relevant to the review question, What research has been undertaken on the use of time within the school on pupils or schools?. The studies identified covered a diverse range of topics, such as the influence of the time of day on pupil learning, the length of the school year, the length of the school day, and the length of lessons within a school day.

The breakdown of studies into topic focus, and pupil and schools outcomes, as categorised in the map, provided the starting point for selecting studies to include in the in-depth review (see Chapter 3). The EPPI-Centre members in conjunction with the DSCF, discussed what topic focus to look at, and the decision was made to synthesise studies that reported on the length of lesson time, specifically studies on block scheduling and their relationship with academic achievement.

The in-depth review question was as follows:

Does block scheduling result in higher levels of student attainment than traditional scheduling?

Sub-questions included the following:

Does the effect of block scheduling vary with type of block scheduling?

Does the effect of block scheduling vary with subject(s) taught?

In-depth review exclusion criteria:

- Not an evaluation of the effect of block scheduling
- Not reporting academic achievement using standardised test(s)
- Not pupils with a mean age between 11 and 16
- No control group in the report.

2.4.2 Detailed description of studies in the in-depth review

Studies identified as meeting the inclusion criteria were analysed in depth, using the EPPI-Centre's detailed data-extraction guidelines (EPPI-Centre, 2002), together with its online software, EPPI-Reviewer (Thomas and Brunton, 2006).

2.4.3 Assessing quality of studies and weight of evidence (WoE) for the review question

The findings and conclusions of the included studies were judged and weighted according to three dimensions:

The quality of the execution (internal methodological coherence) of the studies, based upon the study only, and how judged/criteria (WoE A). The Home Office Quality Assessment Tool (QAT) was used to assess WoE A. The QAT assesses the methodological quality of each study in four key areas: sample selection, bias, data collection and data analysis. A scoring system is used to measure the overall quality, with possible scores ranging from 4 to 20. Studies with the lowest scores are considered to be most methodologically robust.

The appropriateness of the research design and analysis used for answering the review question, and how judged/criteria (WoE B). Studies were judged on WoE B according to the methods used for selecting the sample and how studies dealt with baseline differences between the control and intervention groups. Studies which used random allocation were considered to be the most appropriate method for answering the review question and would warrant a High judgment. Studies that used a matched sample design or statistically controlled for differences in the analysis were considered Medium. Those studies that did not control for differences between the control and intervention groups were considered Low on WoE B.

The relevance of the study topic focus (WoE C). WoE C was judged by the representativeness of the study population. Other measures of relevance for the in-depth review question were comparable across included studies: all studies were conducted in the US, within middle or high schools (geographical or school context) and used similar relevant measures of academic achievement. Thus, studies were judged on WoE C according to the representativeness of the sample: small number of schools (1-3) and low number of students (below 1,000) = Low; small number of schools (1-3), but high number of students (over 1,000) = Medium; large number of schools but low number of students (under 1,000) = Medium; and large number of schools and high number of students= High.

An overall weight, taking into account WoE A, B and C (WoE D).

For this review, the overall WoE D was an average of A, B and C (see Appendix 2.5 for further details).

2.4.4 Synthesis of evidence

The synthesis was organised by the review question which is concerned the effectiveness of different types of block scheduling. Prior to the synthesis, it was hypothesised that outcomes of studies might vary, depending on the following:

- the type of block scheduling
- the type of academic achievement measured
- the quality of the study

The synthesis was undertaken with the aim of exploring patterns of effect sizes using these study characteristics. The categorisation of the type of intervention was determined by the authors' description of the length of lessons and on which days lessons were taught. This allowed the reviewers to group together similar types of school schedules according to how time was blocked in any given subject area.

The Review Group converted the different outcome measures used in the individual studies (although all were measures of academic achievement) to a standard metric that facilitated combination of the individual study results into a weighted average effect size. Effect size calculation allowed exploration across different types of block schedules and allowed comparison between the results of the individual studies. These were calculated from either the raw outcome data and/or the statistical 'result' given in a study. Where data was not available to calculate an effect size, the Review Group contacted the study authors to obtain it.

Studies reported other outcome measures in addition to academic achievement, but these were not included in the synthesis. All academic achievement outcomes for which effect sizes could be calculated

were included; where effect sizes could not be calculated, the findings from the studies were reported separately. Where there was more than one effect size calculated for a study, the outcome that was most frequently used in other studies was used first in any meta-analysis.

The synthesis explored the following patterns:

- Does the effect favour the experimental intervention or control intervention?
- Is the direction of effect the same or different across similar interventions?
- How small or big is the effect size?
- Does the effect size estimate exclude the possibility of the opposite effect?

The statistical technique of meta-analysis was used to produce a weighted average of the individual study effect sizes in each group. This was only undertaken where there was at least one high or medium WoE rated study.

Meta-analysis was completed using techniques (random effect models) that give different weightings to each individual study, and the statistical measures of similarity are provided in each case where a weighted average effect size is given.

2.4.5 Deriving conclusions and implications

An interpretation framework was developed in order to help to summarise and interpret the strength and outcome of the evidence provided. The framework is based on the number and quality of the studies that have evaluated different types of block scheduling, and where the quality is judged to be sufficient the direction (i.e. positive or negative) and size of the weighted average effect size. Further details are provided in Appendix 2.4.

CHAPTER THREE

Identifying and describing studies: results

3.1 Studies included from searching and screening

Figure 3.1 illustrates the process of filtering from searching to mapping and finally to synthesis. Table 3.1 gives the origin of all reports found and those subsequently included in the systematic map.

A total of 8,054 citations were identified through systematic searches of six electronic databases. The number of citations identified in each database is documented in Table 3.1. Of the 8,054 citations identified, 1,860 were duplicates and were excluded when citations were uploaded onto EPPI-Reviewer (Thomas and Brunton, 2006).

The largest yield of the 8,054 citations identified came from ERIC (4,192) and PsycInfo (3,250).

After excluding duplicates, titles and abstracts were screened using the exclusion criteria, described in section 2.2.1. The majority of papers excluded at this stage (5,144) did not meet our first inclusion criterion: that is, they were not on the structure of time within the school day or year. The second most common exclusion criterion was study: if a paper did not report a research study, then it was not included.

The initial screening yielded 674 papers potentially relevant to our review. A further 8 papers were identified through handsearching. Allowing for papers that we were unable to obtain in time (85), 597 papers went through to full screening.

At this second, more detailed stage of screening, a further 453 papers were excluded, again most commonly on the grounds that they did not meet our first criterion for inclusion or that they were not a paper reporting a research study. This resulted in a final total of 130 studies that met our criteria for inclusion in the systematic map. The database closed on Friday 16 August 2008. Papers received after that date will be included in future updates to this review.

Table 3.1: Identification of studies

Attribute	Found*
SSCI	148
BEI	92
IBSS	28
ERIC	4,192
PsycInfo	3,248
AEI	346
Total	8,054

*Mutually exclusive

3.2 Characteristics of the included studies (systematic map)

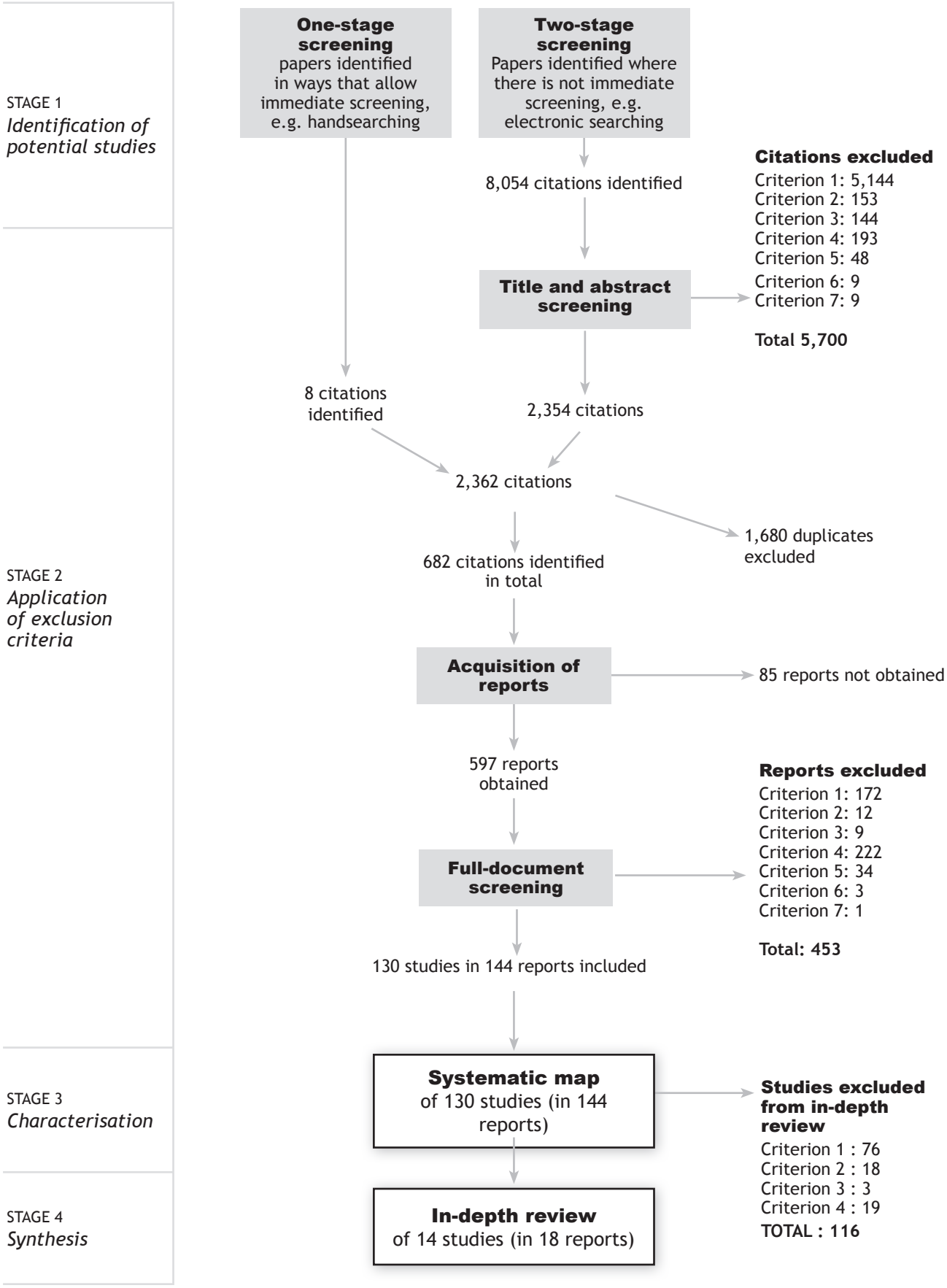
The 130 studies included in the map have been analysed using Section A to E of the EPPI-Centre Data Extraction and Coding Tool for Education Studies V2.0 (EPPI-Centre, 2006) and a set of review-specific keywords (both in Appendix 2.4). The description which follows is based on the data extracted with those tools and provides just a snapshot of selected aspects of the studies included.

Table 3.2: Country (N=130)

Attribute	Found*
Australia	1
Canada	3
Germany	1
Israel	2
Sweden	4
Turkey	1
UK	1
USA	117
Total	130

*Mutually exclusive

Figure 3.1 Filtering of papers from searching to map to synthesis



Although the studies have been conducted across a range of countries, Table 3.2 shows a Western and specifically USA bias, as the overwhelming majority of studies have been conducted in the USA (N=117). A small minority of studies have been conducted in Sweden (N=4), Canada (N=3) and Israel (N=2) and only one study is included from the UK. This represents a large gap between studies from these countries and those conducted in the USA.

Table 3.3: Status of report (N=130)

Attribute	Number*
Published as a journal article	62
Published as a report	45
Published as a conference paper	18
Unpublished	5
Total	130

*Mutually exclusive

The map includes studies that have been published as journal articles (N=62, Table 3.3) and/or reports (N=45). A small proportion of the studies included in the review are papers delivered at conferences (N=18) and unpublished but electronically available Master and PhD theses (N=5).

Table 3.4: Purpose of study (N=130)

Attribute	Number*
Exploration of relationships	13
What works?	117
Total	130

*Mutually exclusive

Table 3.5: Evaluation study focus (N=117)

Attribute	Number**
Outcomes	100
Process	30
Total	130

**Not mutually exclusive

Table 3.6: Study design (N=130)

Attribute	Number**
Case study	6
Case-control study	46
Cohort study	10
Cross-sectional study	5
Document study	1
Ethnography	1
Experiment with non-random allocation to groups	22
One-group post-test only	3
One-group pre-test only	10
Random experiment with random allocation to groups	1
Secondary-data analysis	7
Views study	29
Total	141

**Not mutually exclusive

Of the 130 studies included in the map, 117 studies are evaluative and ask a 'What works' question; only 13 studies look at the relationship between different phenomenon (Table 3.4). Of the 117 studies asking a 'What works?' question, 100 report outcomes of an intervention and 30 look at the process of implementing an intervention (Table 3.5). Thirteen studies report both an outcome and process evaluation. The most common approach to evaluating an intervention was to employ a case-control study design method (N=46, Table 3.6). Many of these types of studies collected pupil data retrospectively and included pupil self-reported measures. The second largest study type were studies reporting views (N=29). In many cases, they concerned the perceived impact of an intervention.

Table 3.7: Type of intervention: 'use of time' (N=130)

Attribute	Number**
Length of the school day	22
Length of lessons	68
Break times	4
Multiple-shift schooling	2
Length of school year	31
Time of day	11
Timing of exams	1
Length of terms/semesters	15
Pupils' use of time	2
Pupils' organisation of learning time	3
Total	159

**Not mutually exclusive

Table 3.8: Type of 'use of time' and educational setting (N=130)

Attribute**	Kindergarten	Primary school	Secondary school
Length of the school day	17	4	6
Length of lessons	1	7	65
Break times	0	4	2
Multiple-shift schooling	0	1	1
Length of school year	4	26	16
Time of day	0	8	7
Timing of exams	0	0	1
Length of terms / semesters	1	9	8
Pupils' use of time	0	1	1
Pupils' organisation of learning time	0	0	2
Total	23	60	109

**Not mutually exclusive

Table 3.9: Type of 'use of time' and pupil achievement (N=130)

Attribute*	School age**	Literacy	Numeracy	Student grades (across subjects)	Exam results (across subjects)
Length of school day	Primary**	16	13	1	1
	Secondary	4	4	1	1
Length of lessons	Primary	4	2	0	0
	Secondary	19	18	34	11
Break times	Primary	1	1	0	0
	Secondary	0	0	0	0
Multiple-shift schooling	Primary	0	0	1	0
	Secondary	1	0	0	0
Length of school year	Primary	8	7	5	4
	Secondary	11	10	5	6
Length of terms	Primary	5	5	3	2
	Secondary	5	5	3	3
Time of day	Primary	1	1	3	0
	Secondary	1	1	3	1
Organisation of teaching subjects	Primary	1	1	2	0
	Secondary	0	0	0	0
Pupils' organisation of learning time	Primary	0	0	0	0
	Secondary	0	0	1	0
Total		77	68	62	29

*Not mutually exclusive
**Primary includes pupils aged between 4-10 enrolled in kindergarten, nursery and/or primary schools and secondary includes pupils enrolled in mainstream schools aged 11-16.

The studies included in the map look at different aspects of the use of time in schools. Table 3.7 shows that a large subset of studies included in the map look at the length of lessons within the school day (N=68); the majority of these studies focus exclusively on the introduction of block scheduling in secondary schools. Thirty-one studies are on the length of the school year, and evaluate the impact and/or delivery of 'year round schooling'. Twenty-two studies look at the length of the school day, many of which look at the impact of attending either half-day of full-day kindergarten. A smaller proportion of the studies look at influence that the time of day has on pupil learning (N=11). In addition, there were some studies which looked the concept of timetable-free schools: how pupils organise their own learning (N=3) and how pupils use their time in schools (N=2).

The categories are not mutually exclusive, because a number of studies report on more than one change in the way time is used in schools.

The categories in Table 3.8 are not mutually exclusive because a number of studies report changes in the use of time in more than one educational setting. The length of the school day was primarily investigated in kindergartens with pupils aged between 5 and 6 (N=17). Studies which looked at the length of the school day in secondary schools (N=6) also looked at primary schools (N=4). Studies addressing the length of lessons were mostly conduced with pupils aged 11-16 in secondary schools (N=65). All the studies looking at break times included primary schools in their sample. The length of the school year focused mostly on primary schools (N=26), but also included secondary schools (N=16). Studies which looked at the impact of the time of day were also concerned with both primary and secondary school pupils. However, the two studies looking at how pupils might organise their day without the use of the timetable were conducted in secondary schools.

Pupil achievement was measured in different ways across and within the studies. All the studies measured pupil achievement either by testing pupils in specific areas (such as Literacy and Numeracy) or by measuring average academic ability across a range of subjects using each pupil's grade or exam results. Table 3.1 provides a breakdown of the number of studies measuring pupil achievement (within or across subjects) by the type of intervention and the school age of pupils.

Studies investigating the length of the school year focused on primary school aged pupil (4-10 year-olds) and had a high number of studies which reported outcomes for Literacy (N=16) and Numeracy (N=13) compared with student grades (N=1) and exam results (N=1). Studies on the length of lessons were mostly conducted with 11-16 year olds in secondary schools. A larger proportion of those studies collect data on student grades (N=34) and exam results (N=11), there are still a number of studies which report Literacy (N=19) and Numeracy (N=18) outcomes in this age group.

Studies investigating the length of the school year used all four types of outcome for both primary school aged pupils (Literacy N=8, Numeracy N=7, student grades N=5, exam results N=4) and secondary school aged pupils (Literacy N=11, Numeracy N=10, student grades N= 5, exam results N=6). The same pattern applied to studies investigating the length of terms/semesters. Studies investigating the impact of the time of day, the organisation of teaching subjects and/or how pupils organise their own time in general were less well reported, focused less on measuring pupil achievement, and, in some cases, did not report any academic outcomes at all.

Table 3.10: Type of 'use of time' and student attendant and drop-out rates (N=130)

Attribute**	Pupil attendance rates	Pupil drop-out rates
Length of the school day	3	2
Length of lessons	17	8
Break times	1	1
Multiple-shift schooling	0	0
Length of school year	7	2
Time of day	1	2
Timing of exams	0	0
Length of terms / semesters	3	1
Pupils' use of time	0	0
Pupils' organisation of learning time	0	0
Total	32	16

**Not mutually exclusive

In addition to reporting student academic achievement, many studies also reported student attendance rates and drop-out rates. This was mostly reported by studies looking at the length of lessons (attendance rates N=17, drop-out rates N=8) and length of the school year (attendance rates N=7, drop-out rates N=2).

Some of the studies in the review also reported on pupils' non-academic outcomes. These included self-esteem, relationship outcomes, psycho-social well-being, pupil motivation and pupil attention rates. Studies which looked at the length of lessons, length of the school day and the impact of the time of day on learning were interested in, and reported, on all five outcomes (see Table 3.11).

Table 3.12 indicates that only a small minority of studies reported on school outcomes (climate and teaching) and costs compared with pupil outcomes. The review did not set out specifically to identify studies which included economic costs and did not search economic databases. However, five studies reporting costs on the three main ways of changing the use of school time have been identified.

Many of the studies included in the map use pupil self-reported measures. Table 3.13 shows that 10 of the studies looking at the impact of time of day and changing when lessons were taught reported student's satisfaction with those changes. Of the 68 studies evaluating different lesson lengths, 27 report students' perceptions of the intervention, including their perception of the effectiveness of the intervention.

Table 3.11: Type of 'use of time' and non-academic pupil outcomes (N=130)

Attribute**	Pupil's self esteem	Relationships outcomes	Psycho-social well being	Pupils' motivation	Attention rates
Length of the school day	1	2	2	0	1
Length of lessons	2	4	2	2	1
Break times	1	1	1	0	0
Multiple-shift schooling	0	0	0	0	0
Length of school year	2	2	4	2	2
Time of day	2	2	1	1	3
Timing of exams	0	0	0	0	0
Length of terms / semesters	1	0	0	0	0
Pupils' use of time	0	0	0	0	0
Pupils' organisation of learning time	1	1	0	2	0
Total	10	12	12	8	7

**Not mutually exclusive

Table 3.12: Type of 'use of time' and school outcomes (N=130)

Attribute**	School climate	Teaching	Costs
Length of the school day	2	2	1
Length of lessons	3	3	1
Break times	1	1	0
Multiple-shift schooling	0	0	0
Length of school year	2	2	3
Time of day	2	2	0
Timing of exams	0	0	0
Length of terms / semesters	1	0	0
Pupils' use of time	1	1	0
Pupils' organisation of learning time	1	1	0
Total	13	12	5

**Not mutually exclusive

Table 3.13: Type of 'use of time' and pupil's self reported outcomes (N=130)

Attribute**	Satisfaction with intervention	Perception of intervention	Perception of impact / effectiveness	Perception of implementation
Length of the school day	4	0	3	1
Length of lessons	0	27	11	21
Break times	1	1	2	1
Multiple-shift schooling	1	0	0	0
Length of school year	0	0	0	0
Time of day	10	2	8	1
Timing of exams	0	2	0	2
Length of terms / semesters	0	2	0	0
Pupils' use of time	5	1	4	1
Pupils' organisation of learning time	0	1	1	2
Total	21	36	29	30

**Not mutually exclusive

3.3 Summary of results of map

Overall, we identified a total of 130 research studies looking at the use of time in schools. The majority of which were journal articles or reports publishing evaluations of interventions conducted in the US.

The most common approach to evaluating an intervention was to employ case-control study design methods (N=46), many of which included self-reported measures (see Table 3.13).

The majority of studies identified in the map focused on the length of lessons (N=68), length of school year (N=31) and length of the school day (N=22).

Studies reported outcomes on academic achievement (see Table 3.9) and student attendance and drop-out rates (see Table 3.10).

Some of the studies in the review also reported on pupils' non-academic outcomes (see Table 3.11) and school climate and teaching (see Table 3.12).

CHAPTER FOUR

In-depth review: results

4.1 Selecting studies for the in-depth review

Chapter 3 described the findings of the first stage of the review process and presented the results of a systematic and broad descriptive mapping of the 130 relevant studies identified. This chapter describes the second stage of the review process, the findings of an in-depth review, and the synthesis of the quality and findings of a subset of studies relevant to answering the narrower in-depth review question agreed in consultation with the DCSF.

The in-depth review question was as follows:

Does block scheduling result in higher levels of student attainment than traditional scheduling?

Sub-questions for the in-depth review examined the effect by subject, outcome measure and type of block scheduling (intervention). These included the following:

What is the effect of block scheduling (all types, only by 4 x 4 or only A/B) on (i) academic achievement, (ii) Mathematics achievement, (iii) Science achievement, (iv) English achievement, and student exam scores/grades?

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

Fourteen studies met our inclusion criteria for the in-depth review. All the studies were published between 1995 and 2004 and evaluated different types of block schedules. The studies were all conducted in North America (USA N=13, Canada N=1) and investigated the UK equivalent of secondary school pupils aged 11-16. Pupil level data was collected and examined to evaluate the effectiveness of block scheduling. Most studies evaluated either the 4 x 4 block schedule (N=8) or the A/B block schedule (N=5).

Some studies also examined hybrid schedules (N=3) and the impact of extending a single lesson (N=2). Of the 14 evaluations included in the in-depth review, only one used a quasi-randomised design. The remaining 13 studies used a retrospective study design whereby the outcomes of students who were already enrolled in schools delivering block scheduling (the intervention group) were compared with the outcomes of students already enrolled in school using a traditional schedule (the control group). In some cases, this would be the same school at two different points in time: that is, comparing the outcomes of a cohort of students before block scheduling had been implemented with a later cohort of students who followed the block schedule.

4.2.1 School schedule types

Traditional schedule: Students enrolled in traditional schedules participate in six to eight classes per day for 40-60 minutes per class period. Each class takes one year to complete. All the studies in the in-depth review used the traditional schedule as their control group.

4 x 4 block schedule: This design consists of four blocks of extended duration classes (80–90 minutes each) per day and allows students to take up to four different subjects in one term, and up to eight courses over two terms in a school year.

A/B block schedule: The 'alternating block schedule' or the A/B block schedule organises the school day into classes of 70-90 minutes each for three to four different subjects on every alternating day. The A/B block schedule can mean that six to eight subjects are studied throughout the year but on alternate days; the classes are again clubbed into 'blocks' and are of a longer duration than traditional classes.

Hybrid block: Three studies look at the impact of hybrid models of block scheduling. Hybrid models usually operate with five classes a day instead of

Table 4.1: Study, type of block scheduling, outcome measured

Study	Type of block scheduling	Outcome measured
Cobb et al. (1999)	4 x 4 block	GPA, Mathematics, Reading, Writing
DiRocco (1997)	A/B block	GPA, Mathematics, Reading, Science
Hughes (2004)	4 x 4 block	GPA
Lapkin et al. (1997)	Half-day; 80-minute lessons	French
Lewis et al. (2003)	4 x 4 block, A/B block	Language Arts, Science
Marchette (2003)	4 x 4 block, A/B block	Science (Biology)
McCreary and Hausman (2001)	A/B block, hybrid block	Mathematics, Science
Rice et al. (2002)	70-minute lessons	Mathematics
Schreiber et al. (2001)	4 x 4 block; hybrid block	Mathematics, Language, Reading
Schroth and Dixon (1995)	A/B block	Mathematics
Texas Education Agency (1999)	A/B block, modified alternate block schedule	Mathematics, Reading, Writing
Veal (1999)	4 x 4 block, hybrid block	GPA
Walker (2000)	4 x 4 block	Mathematics
Zhang (2001)	4 x 4 block scheduling	Algebra, Biology, English, ELP, History

four in the block schedule, and six in a traditional schedule. Lessons are longer than 55 minutes, but shorter than the 90- minute lessons you would find with the 4 x 4 and A/B models.

Single block lesson: Two studies looked at the impact of a block lesson on student performance. In one study researchers looked at teaching French in half-day and 80-minute slots, and another looked at teaching Mathematics in lessons of 70 minutes or longer. In both cases, the amount of time spent on each subject was not increased over the school year but concentrated into a block similar to the 4 x 4, A/B and hybrid schedules.

4.2.2 Outcome measures: academic achievement

As outlined in 4.2 the most frequent measure of achievement was in (i) Mathematics, (ii) Science, (iii) English and (iv) GPA/school grades (see Table 4.1); this latter measure was often given as an averaged standard metric that described students' scores in exams or grades across different academic subjects. To be included in the in-depth review, studies were required to measure outcomes using standardised tests. In most cases, researchers opted for valid measures of achievement commonly used in the local educational state area. For example, studies conducted in Texas (N=2) used the TAAS test (Texas Assessment of Academic Skills) another study used the North Carolina End of Course tests. Some studies used SATs or tests specific to the course subject. Four studies also looked at the improvement of GPAs across academic subjects, rather than directly measuring academic achievement in one core subject.

4.2.3 Case studies

Case study 1: Cobb et al. (1999) 4 x 4 block scheduling

Students in one Colorado high school attended four, 90-minute courses every day of the school week for a minimum of one and a half years. Students were mixed sex, aged 13-17, and overwhelmingly white, reflecting both the community and neighbourhoods in which the schools were located. When compared with similar students in schools following traditional schedules, these students performed significantly less well on standardised Mathematics tests, but had consistently higher GPAs.

Case study 2: DiRocco (1997) Alternative day (A/B) block scheduling

A middle school in rural, central Pennsylvania, originally implemented alternate day block scheduling as an experiment led by one teaching team with their seventh-grade students (aged 12-13). This schedule meant that classes met every other day, in 82-minute periods, as opposed to every day for 40 minutes. Students had three classes per day instead of six.

The teaching team were impressed with this format and so sought permission to continue the schedule. The school principal agreed to run the alternate day schedule for one cohort of students across multiple subjects (English, Mathematics, Social Studies, Reading and Foreign Languages), during their seventh and eight grades. The study compared the final course averages, GPA and standardised achievement tests of these students with the previous cohort of students who had experienced the traditional schedule. Compared with students attending the traditional format, students of the alternate block schedule had

higher course marks across some subjects in their seventh grade and all subjects in their eighth. The author claims that this demonstrates that students required some time to adjust to the new schedule.

Case study 3: Lapkin et al. (1997) 80-minute and half-day French lessons

The Carlton Board of Education (Ontario, Canada) implemented and evaluated a second language course that had proven successful in teaching English in Quebec. The programme involved teaching French to seventh-grade, English-speaking students (aged 12-13) in a middle school. Teaching followed one of three alternate schedules: (1) a half day of instruction of French over a 10-week period, (2) 80 minutes of French instruction per day for five months, and (3) 40 minutes of French per day for 10 months (traditional schedule). Students were assigned to one of these classes, with minor exceptions, on a random basis. The actual amount of time devoted to French instruction remained the same across all models. The Carlton Board of Education arranged for one teacher to lead all three groups, using the same curriculum, in the 1993-94 school year. The half-day schedule was run from October through to December, the 80-minute model from January to June, and the traditional model throughout the year. The evaluation found that there were no significant differences in academic performance between the three groups in French listening comprehension and French speaking. Compared with the traditional model, students in the block schedule models presented academic gains in French reading and writing.

Case study 4: Walker (2000) Block scheduling (all types)

Set against a context of growing national interest in block scheduling and school reform, this study examined the prevalence and impact of block scheduling in the state of Kansas, USA. The author conducted a short telephone survey of all public secondary schools in Kansas to determine if block scheduling had been implemented, of which type and for how long. To be considered a block schedule, the school had to implement periods of 80 minutes or more. The author found that 130 out of 345 schools had implemented some kind of block scheduling. The study measured the impact of this block scheduling upon students' achievement in Mathematics. This was measured by the Kansas State Mathematics Assessment Test, which was believed to be a valid and reliable measure of mathematical achievement. A 'Power Score' was calculated based on the subscales of the test: reasoning, communication and problem-solving. Data was collected for five years from all schools in Kansas State, representing test results taken for approximately 150,000 tenth-grade students. The author concluded that there was no significant measurable difference in the impact of block scheduling compared with traditional scheduling on the Mathematics assessment test.

4.3 Quality and relevance of outcome studies

Two reviewers independently extracted data from individual studies by answering questions about the aims and rationale of the study, execution of method, sampling strategy, internal and external validity, results and conclusions and generalisability of the findings. Two coding tools were used to extract data from each study and these can be found in Appendix 2.4. The weight of evidence (WoE) contributed by each study was assessed through careful assessments and re-reading of the study and all the answers provided. Each data extraction was carried out blind and answers to all the data-extraction questions and WoEs were compared. Discussion about any discrepancies was made at this time and final judgements were agreed. The majority of studies were considered to be of medium weight of evidence overall for this review (N=12) and two were considered to be of low quality evidence. None of the studies was considered to be of high quality on weight of evidence D.

4.4 Synthesis of evidence

4.4.1 Introduction

The synthesis examines the effect of block scheduling on academic achievement. The findings are organised according to the type of outcome measured, and are reported in the following order: academic achievement, Mathematics, Science, English and grades. For each of these outcome measures, the synthesis is organised according to the type of intervention (block scheduling), considering the effect of all types of block scheduling, only 4 x 4 and then only A/B block scheduling.

Each included study compared the effect of block scheduling on academic achievement against a control group, students who had undergone a traditional schedule. In most cases, this referred to 6-8 classes per day for 40-60 minutes per class spread across year. The control group was, therefore, comparable across all studies.

4.1.1.1 Studies/outcome measures excluded from the synthesis

The studies by Schroth and Dixon (1995) and Veal (1999) have been excluded from all the meta-analyses because we could not calculate a standardised effect size from the data available in the report. This means that 12 of the 14 included studies were used in the meta-analyses. However, the findings of the two individual studies have been included in the in-depth review, where appropriate.

Table 4.2: Weights of evidence of studies included in the in-depth review

Study	Quality of execution (WoE A)	Appropriateness of study design (WoE B)	Relevance (WoE C)	Overall weight of evidence (WoE D)
Cobb et al. (1999)	High	Medium	Medium	Medium
DiRocco (1997)	Medium	Medium	Low	Medium
Hughes (2004)	Medium	Low	Low	Low
Lapkin et al. (1997)	High	Medium	Low	Medium
Lewis et al. (2003)	Medium	Medium	Low	Medium
Marchette (2003)	Medium	Medium	High	Medium
McCreary and Hausman (2001)	Medium	Medium	Medium	Medium
Rice et al. (2002)	Medium	Medium	High	Medium
Schreiber et al. (2001)	Medium	Medium	Low	Medium
Schroth and Dixon (1995)	Medium	Low	Low	Low
Texas Education Agency (1999)	Medium	Medium	High	Medium
Veal (1999)	Medium	Low	Medium	Medium
Walker (2000)	Medium	Medium	High	Medium
Zhang (2001)	Medium	Medium	High	Medium

4.1.1.2 Selecting the effect size from each study

One effect size was selected from each study, for each meta-analysis. Selection was based on the most common type of block scheduling/outcome measure used across the studies. This meant that, from each individual study, we selected effect sizes in the following order:

4 x 4 block schedule: a. Mathematics, b. Science, c. English, d. School grades

A/B block schedule: a. Mathematics, b. Science, c. English, d. School grades

We used the same order for selecting which subjects to include in the meta-analysis when answering the sub-questions on the effect of the 4 x 4 block schedule and the effect of the A/B block schedule on academic achievement. When combining studies that vary in outcome and by intervention, this can lead to high levels of heterogeneity.

4.1.2.3 Selection of fixed and random effects models

For this review, we decided to use the random effects model to answer both the in-depth review question and sub-questions because it is more appropriate for combining heterogeneous groups and places less emphasis on the larger studies than the fixed effects model.

4.4.2 The effect of block scheduling on academic achievement

4.4.2.1 What is the effect of block scheduling on academic achievement?

Twelve studies looked at the effect of blocked scheduling on academic achievement. The studies included in this meta-analysis vary by type of block scheduling (4 x 4, A/B, hybrid, individual block lesson) and outcome measures (Mathematics, Science, English and School grades). Of the 12 studies included in this meta-analysis, 11 were judged to be of medium weight of evidence and one study was judged to be of low weight of evidence overall (WoE D) for this review.

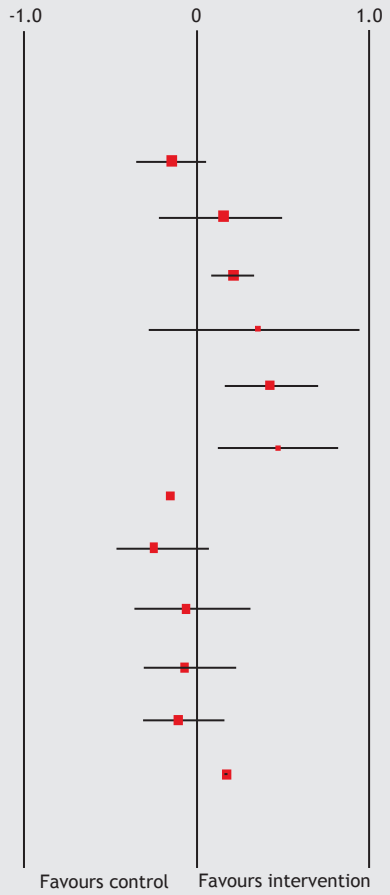
Figure 4.1 shows the effect sizes for these studies. There is no consistent pattern of effect and the visual impression is matched by the statistical indicators that suggest there is a substantial degree of heterogeneity between the studies (I2 = 85.7%). In order to explore the results, further sensitivity analysis was undertaken by completing a number of mini-meta-analyses exploring the effect of different types of block scheduling, different subject areas and different outcomes.

4.4.2.2 What is the effect of the 4 x 4 block scheduling on academic achievement?

Seven studies evaluated whether 4 x 4 block scheduling was more effective than traditional school schedules at improving academic achievement (Cobb et al., 1999; Hughes, 2004; Lewis et al., 2003; Marchette, 2003; Schreiber et al., 2001; Texas Education Agency, 1999; Zhang,

Figure 4.1: Effect sizes (Hedges' g) of block scheduling on academic achievement

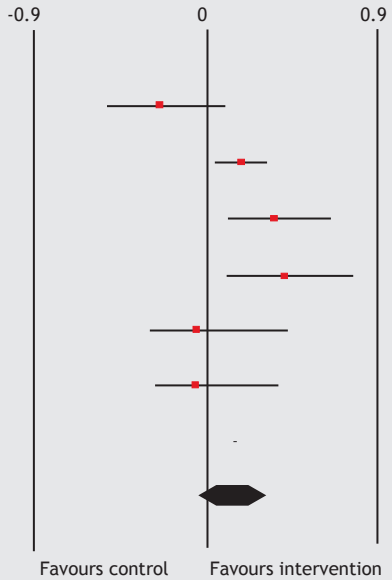
Item	Effect (CI)	Weight %	Size
All B/S			
Cobb et al (1999)	-0.14 (-0.33, 0.04)	10.1	472
Dirocco (1997)	0.18 (-0.14, 0.49)	6.1	159
Hughes (2004)	0.18 (0.05, 0.31)	12.1	937
Lapkin et al (1997)	0.39 (-0.17, 0.96)	2.6	49
Lewis (2003)	0.45 (0.15, 0.74)	6.5	186
Marchette (2003)	0.48 (0.10, 0.85)	4.9	121
McCreary et al (2001)	-0.10 (-0.15, -0.04)	14.4	4900
Rice et al (2002)	-0.25 (-0.54, 0.04)	6.7	4569
Schreiber et al (2001)	-0.03 (-0.35, 0.30)	5.8	191
Texas Education Agency (1999)	-0.04 (-0.31, 0.24)	7.0	200
Walker (2000)	-0.10 (-0.32, 0.12)	8.8	345
Zhang (2001)	0.10 (0.09, 0.11)	15.0	640050



Heterogeneity statistic Q = 76.8 df = 11 p = 6.24E-12 I2 = 85.7%
Test statistic (combined effect) z = 1.02 p = 0.309

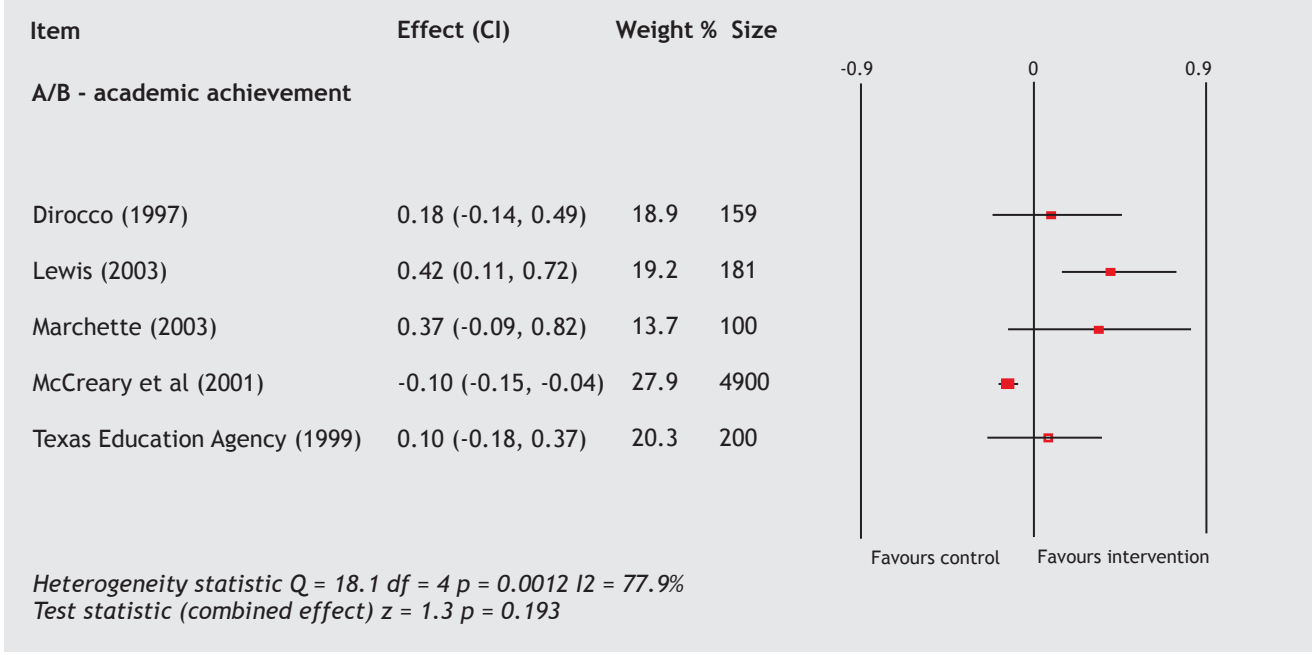
Figure 4.2: Effect sizes (Hedges' g) of 4 x 4 block scheduling on academic achievement

Item	Effect (CI)	Weight %	Size
4x4 Academic Achievement			
Cobb et al (1999)	-0.14 (-0.33, 0.04)	16.4	472
Hughes (2004)	0.18 (0.05, 0.31)	20.5	937
Lewis (2003)	0.42 (0.11, 0.72)	9.3	181
Marchette (2003)	0.48 (0.10, 0.85)	7.0	121
Schreiber et al (2001)	-0.03 (-0.35, 0.30)	8.5	191
Texas Education Agency (1999)	-0.04 (-0.31, 0.24)	105	200
Zhang (2001)	0.10 (0.09, 0.11)	27.7	640050
	0.11 (-0.01, 0.22)		



Heterogeneity statistic Q = 18.3 df = 6 p = 0.00555 I2 = 67.2%
Test statistic (combined effect) z = 1.85 p = 0.0645

Figure 4.3: Forest plot of effectiveness of interventions (random effect model), for studies with medium/high WoE



2001). Of the seven studies, six were judged to be medium and one (Hughes, 2004) was judged to be low overall on WoE D for this review. Using the criteria set out in section 2.4, the most common measure of academic achievement in this group of studies was Mathematics, and this was chosen as the first outcome to be included in this meta-analysis (N=4), followed by Science (N=2) and student grades (N=1).

Figure 4.2 shows the effect sizes for studies in this group. The individual effect sizes were meta-analysed to produce a weighted average effect size of g = 0.11(C.I. -0.01 to 0.22).

Four studies favour block scheduling and three studies favour the traditional schedule. The forest plot (Figure 4.2) also shows that the three studies which favour the control have wide confidence intervals which cross the line of 'no effect'.

4.4.2.3 What is the effect of A/B block scheduling on academic achievement?

Five studies evaluated the effect of A/B block schedules on academic achievement compared with traditional schedules (DiRocco, 1997; Lewis et al., 2003; Marchette, 2003; McCreary and Hausman, 2001; Texas Education Agency, 1999). All five studies were judged to be medium overall on WoE D for this review. Similar to the 4 x 4 block scheduling studies, the most common measure of academic achievement was Mathematics and this was chosen as the first outcome to be included in this meta-analysis (N=3, DiRocco 1997, McCreary and Hausman 2001, Texas Education Agency 1999) followed by Science (Lewis 2003, Marchette 2003).

Figure 4.3 shows the effect sizes for studies in this group. Four studies favour the intervention (block scheduling) and one study favours the control (traditional schedule). However, three of the four studies which favour the intervention have confidence intervals which cross the line of 'no effect'. Similar to the meta-analysis on the effect of block scheduling on academic achievement, this group of studies produced a high level of statistical heterogeneity (I2 = 77.9%). Therefore, they cannot be combined to produce a valid summary effect size; there is no pooled estimate of effect shown in Figure 4.3.

Mathematics used as an equivalent for grade 10 exams (Texas Education Agency, 1999). Studies also vary according to the type of block schedule (4 x 4 and A/B). Four studies were judged as medium and one study was judged as low overall on WoE D for this review.

Figure 4.4 shows the effect sizes for studies in this group. The individual effect sizes were meta-analysed to produce a weighted average effect size of g=0.18 (C.I. 0.06 to 0.30). Three out of four studies favour the intervention. Only one study has a wide confidence interval crossing the line of 'no effect'.

Removing Hughes (2004) from the synthesis, as a 'low' weighted study, creates a synthesis of all 'medium' weighted studies (see Figure 4.5). This provides the same pooled summary effect size, g= 0.18 but with a wider confidence interval (CI -0.04, 0.39).

Figure 4.4: Effect sizes (Hedges' g) of block scheduling on student exam/grade scores

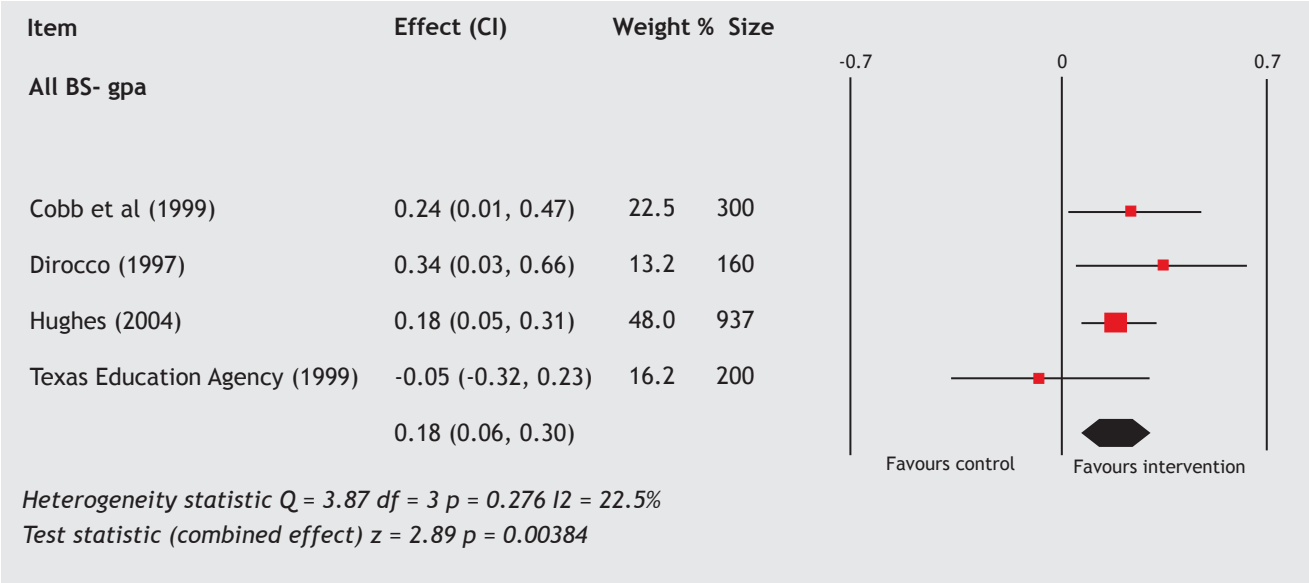
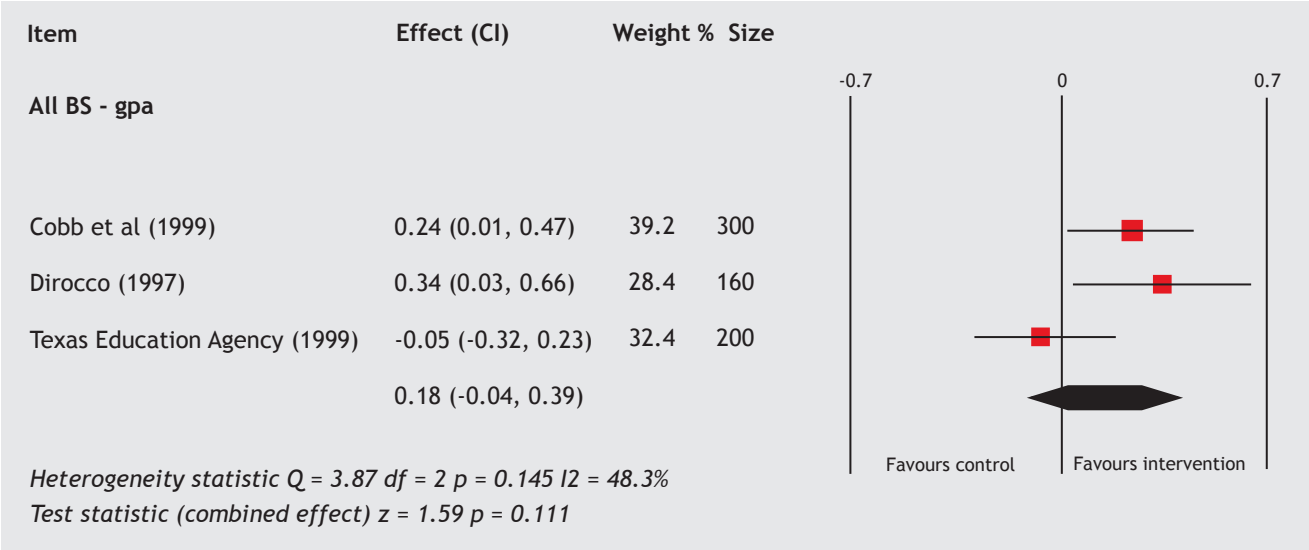


Figure 4.5: Effect sizes (Hedges' g) of block scheduling on student grades excluding Hughes (2004)



4.4.2.5 What is the effect of 4 x 4 block scheduling on student exam/grade scores?

Three studies measured the effect of 4x4 block scheduling on student exam/grade Scores, (Cobb et al., 1999; Hughes, 2004, Texas Educational Agency, 1999). Two studies were judged to be medium and one study was judged to be low overall on WoE D for this review.

Figure 4.6 shows the effect sizes for studies in this group. The individual effect sizes were meta-analysed to produce a weighted average effect size of g=0.15 (C.I. 0.02 to 0.29, p=0.002). Two out of three studies favour the intervention. Only one study has a wide confidence interval crossing the line of 'no effect'.

4.4.2.6 What is the effect of A/B block scheduling on student exam/grade scores?

Two studies look at the effect of A/B block scheduling on student grades (DiRocco 1997, Texas Education Agency, 1999) and were both weighted medium overall on WoE D for this review.

Figure 4.7 shows the effect sizes for studies in this group. The individual effect sizes were meta-analysed to produce a weighted average effect size of g=0.22 (C.I.- 0.01 to 0.44). Both studies favour the intervention and one study has a wide confidence interval crossing the line of 'no effect'.

Figure 4.6: Effect sizes (Hedges' g) of 4 x 4 block scheduling on student exam/grade scores

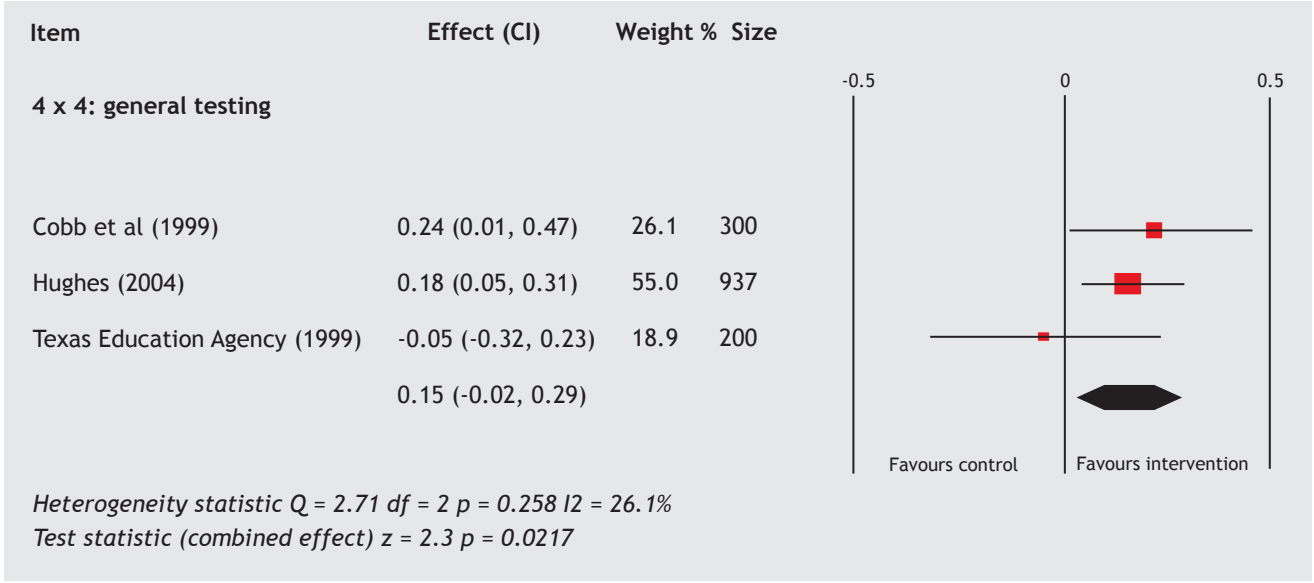
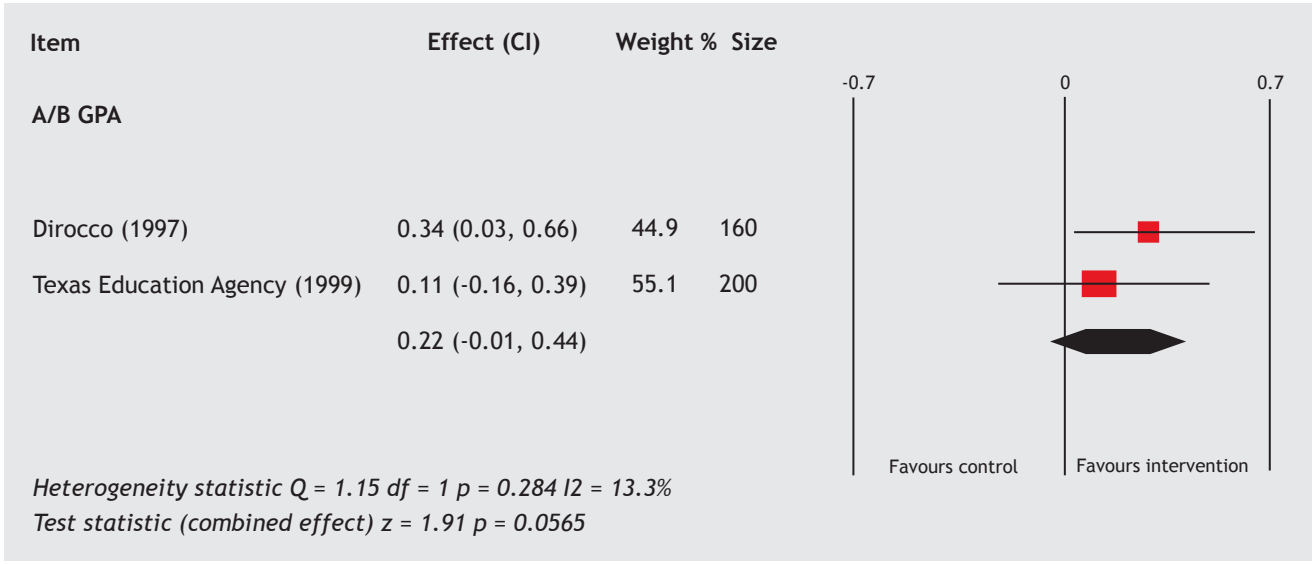


Figure 4.7: Effect sizes (Hedges' g) of A/B block scheduling on student exam/grade scores



4.4.2.7 Additional study

The study by Veal (1999), which was excluded from the meta-analysis because effect sizes could not be calculated, also investigated the effect of a hybrid schedule (which consists of three traditional and two block classes each day) on student exam/grade scores. They found that students had higher student average exam/grade scores compared with those enrolled in traditional schedules.

However, it is important to consider whether an average academic score across subjects could in fact be concealing positive effects for block scheduling in some subjects at the same time as negative effects in others. The analysis of the effects of block scheduling within single subjects above suggests that this might be a possibility.

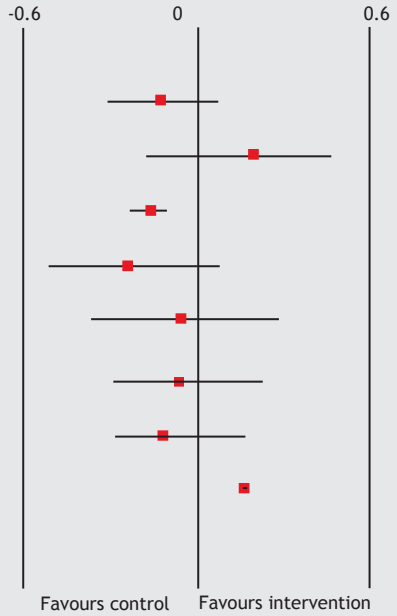
4.4.3 The effect of block scheduling on different subjects

4.4.3.1 What is the effect of block scheduling on mathematics achievement?

Eight studies examined the effects of block scheduling on mathematics achievement compared with traditional schedules (Cobb et al., 1999; DiRocco, 1997; McCreary and Hausman, 2001; Rice et al., 2002; Schreiber et al., 2001; Texas Education Agency, 1999; Walker, 2000; Zhang, 2001). These studies vary by type of block scheduling (4 x 4, A/B, hybrid and block lesson) and the measurement used to assess achievement in Mathematics (from standardised state-wide tests to school-specific testing). All the studies were judged to be 'medium' overall on WoE D for this review.

Figure 4.8: Effect sizes (Hedges' g) of block scheduling on Mathematics achievement

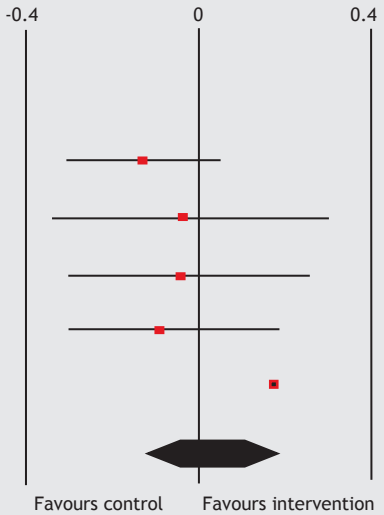
Item	Effect (CI)	Weight %	Size
All BS - maths			
Cobb et al (1999)	-0.14 (-0.33, 0.04)	13.7	472
Dirocco (1997)	0.18 (-0.14, 0.49)	8.4	159
McCreary et al (2001)	-0.10 (-0.15, -0.04)	19.2	4900
Rice et al (2002)	-0.25 (-0.54, 0.04)	9.1	4569
Schreiber et al (2001)	-0.03 (-0.35, 0.30)	8.0	191
Texas Education Agency (1999)	-0.04 (-0.31, 0.24)	9.6	200
Walker (2000)	-0.10 (-0.32, 0.12)	12.0	345
Zhang (2001)	0.10 (0.09, 0.11)	20.0	640050



Heterogeneity statistic Q = 64.9 df = 7 p = 1.6E-11 I2 = 89.2%
Test statistic (combined effect) z = 0.725 p = 0.469

Figure 4.9: Effect sizes (Hedges' g) of 4 x 4 block scheduling on Mathematics achievement

Item	Effect (CI)	Weight %	Size
4x4: mathematics			
Cobb et al (1999)	-0.14 (-0.33, 0.04)	21.4	472
Schreiber et al (2001)	-0.03 (-0.35, 0.30)	11.5	191
Texas Education Agency (1999)	-0.04 (-0.31, 0.24)	14.1	200
Walker (2000)	-0.10 (-0.32, 0.12)	18.2	345
Zhang (2001)	0.10 (0.09, 0.11)	34.7	640050
	-0.02 (-0.16, 0.11)		



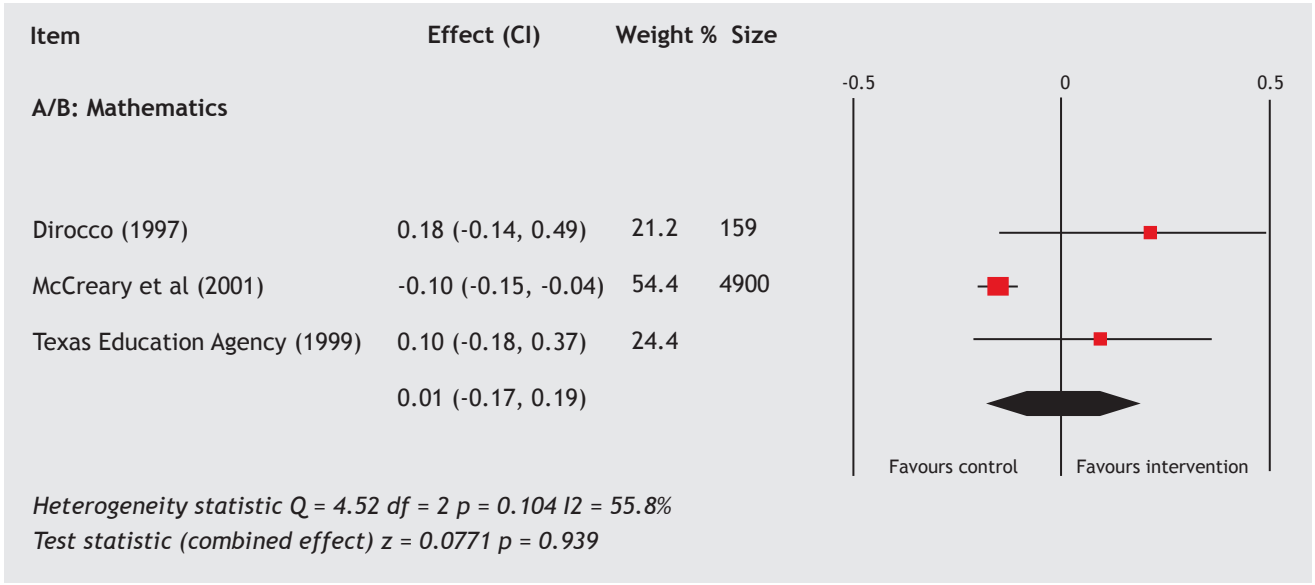
Heterogeneity statistic Q = 11.8 df = 4 p = 0.0192 I2 = 66%,
Test statistic (combined effect) z = 0.33 p = 0.741

Figure 4.8 shows the effect sizes for studies in this group. The majority of studies favoured the traditional schedule (N=6), showing a negative effect of block scheduling on achievement in Mathematics. However, all the studies, except for the two larger ones (McCreary and Hausman, 2001; Zhang, 2001) cross the line of 'no effect'. Similar to the meta-analysis on the effect of block scheduling on academic achievement, this group of studies produced a high level of statistical heterogeneity (I2= 89.2%) and cannot be combined to produce a valid summary effect size; there is no pooled estimate of effect shown in Figure 4.8.

4.4.3.2 What is the effect of 4 x 4 block scheduling on Mathematics achievement?

Five studies evaluated whether 4 x 4 block scheduling improves mathematics achievement compared with traditional schedules (Cobb et al., 1999; Schreiber et al., 2001; Texas Education Agency, 1999; Walker Zhang, 2001). All four studies

Figure 4.10: Effect sizes (Hedges' g) of A/B block scheduling on Mathematics achievement



were judged to be medium overall on WoE D for this review.

When the studies examining the effect of 4 x 4 block scheduling were pooled (see Figure 4.9) the effect size favoured the traditional schedule rather than a 4 x 4 schedule (g =-0.02, 95% C.I. -0.16 to 0.11). However, there is a high degree of heterogeneity between the studies and the confidence interval of the pooled estimate of effect does not cross zero; therefore, there is no clear evidence on the impact of scheduling (4 x 4 or traditional) on achievement in Mathematics.

The study by Schroth and Dixon (1995) (excluded from the meta-analysis because effect sizes could not be calculated) also found that students' achievement in Mathematics did not improve as a result of attending 90-minute, rather than the traditional 50-minute, lessons.

4.4.3.3 What is the effect of A/B block scheduling on mathematics achievement?

Three studies look at the effect of A/B block scheduling on Mathematics teaching compared with traditional schedules (DiRocco, 1997; McCreary and Hausman, 2001; Texas Education Agency, 1999) and were all weighted medium overall on WoE D for this review.

When the studies examining the effect of A/B block scheduling were pooled (see Figure 4.10), these studies favoured A/B block scheduling rather than the traditional schedule (g=0.01 95%C.I. -0.17 to 0.19). However, the two studies that favour the intervention have wide confidence intervals crossing the line of 'no effect', thus excluding the possibility that A/B block scheduling could impact positively or negatively on achievement in Mathematics.

4.4.3.4 What is the effect of block scheduling on science achievement?

Five studies considered the effect of block scheduling on student achievement in Science compared with traditional schedules. The studies include 4 x 4 and A/B types of block scheduling but none of the studies considered the effect of a hybrid schedule or block lessons on achievement in Science (DiRocco, 1997; Lewis et al., 2003; Marchette, 2003; McCreary and Hausman 2001; Zhang, 2001). Outcome measures include Science as a whole and only Biology. All studies have been judged as 'medium' overall on WoE D for this review.

Figure 4.11 shows the effect sizes for studies in this group. All the studies favour block scheduling (N=6). Only one study (DiRocco, 1997) has a confidence interval that crosses the line of 'no effect', suggesting the majority of studies found a difference between block and traditional scheduling. However, similar to the meta-analysis on the effect of block scheduling on academic achievement, this group of studies produced a high level of statistical heterogeneity (I2 = 84.4 %) and cannot be combined to produce a valid summary effect size; there is no pooled estimate of effect shown in Figure 4.11.

4.4.3.5 What is the effect of 4 x 4 block scheduling on Science achievement?

Three studies evaluated whether 4 x 4 block schedules were more effective than traditional schools schedules at improving science achievement (Lewis et al., 2003; Marchette, 2003; Zhang, 2001). All three studies were judged to be medium overall on WoE D for this review.

Figure 4.11: Effect sizes (Hedges' g) of block scheduling on Science achievement

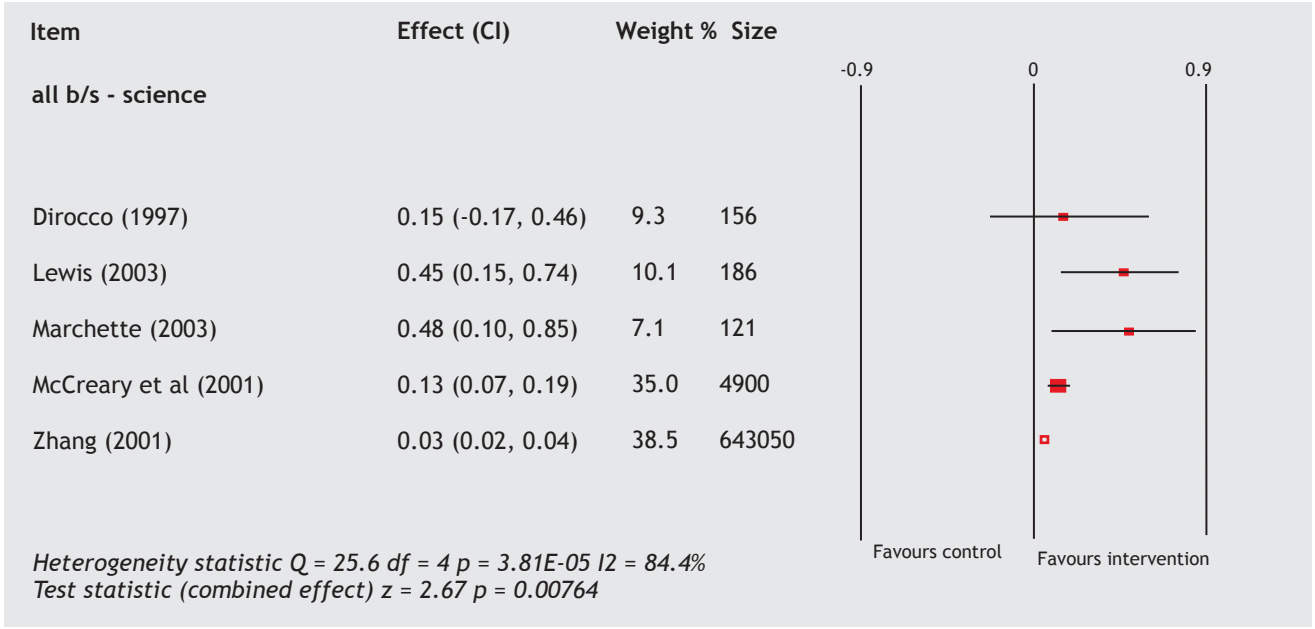


Figure 4.12: Effect sizes (Hedges' g) of 4 x 4 block scheduling on science achievement

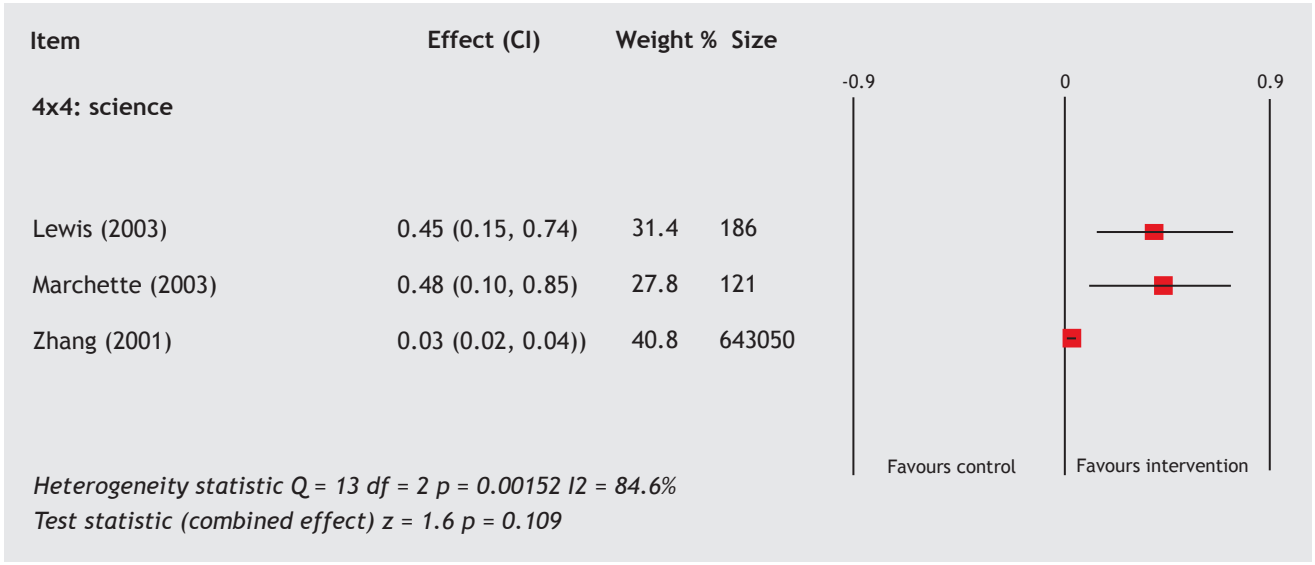


Figure 4.12 shows the effect sizes for studies in this group. All the studies favour block scheduling (N=3) and none of the studies has confidence intervals which cross the line of 'no effect'. However, similar to the meta-analysis on the effect of block scheduling on academic achievement, this group of studies produced a high level of statistical heterogeneity (I2 = 84.6 %) and cannot be combined to produce a valid summary effect size; there is no pooled estimate of effect shown in Figure 4.12.

4.4.3.6 What is the effect of A/B block scheduling on science achievement?

Four studies evaluate the effect of A/B block scheduling in science achievement compared with traditional schedules (DiRocco, 1997; Lewis et al., 2003; Marchette, 2003; McCreary and Hausman, 2001). All four studies were judged as medium overall on WoE D for this review.

Figure 4.13 shows the effect sizes for studies in this group. The pooled estimate of effect was positive for the impact of A/B block scheduling on achievement in science (g=0.20, 95% C.I. 0.06 to 0.33). All four studies favour the intervention, of which only two have wide confidence intervals crossing the line of 'no effect'.

4.4.3.7 What is the effect of block scheduling on English achievement?

Five studies look at the effect of block scheduling on English achievement. The majority of this set of studies examines the effect of 4 x 4 block scheduling on English achievement. Only one study (DiRocco,

Figure 4.13: Effect sizes (Hedges' g) of A/B block scheduling on science achievement

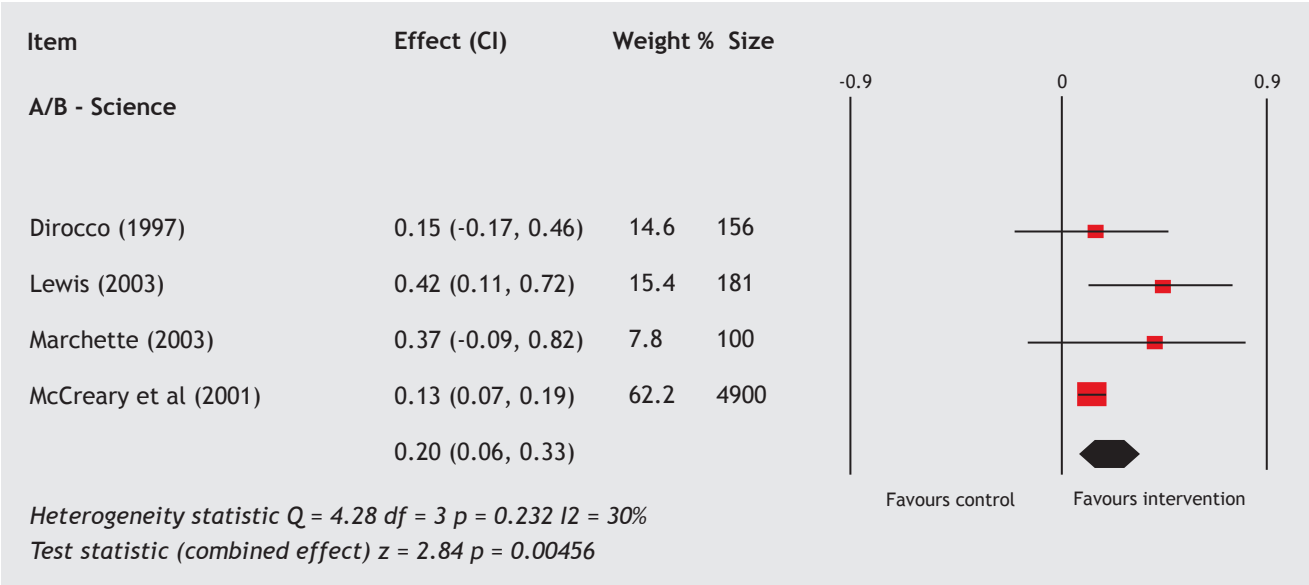
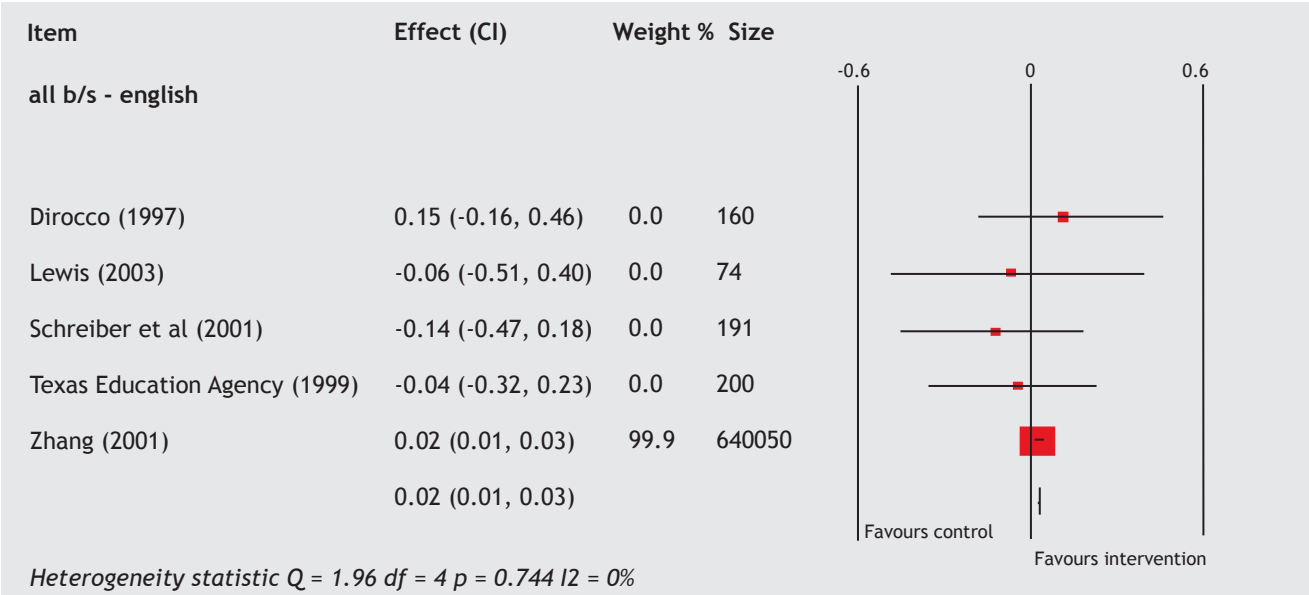


Figure 4.14: Effect sizes (Hedges' g) of block scheduling on English achievement



1997) examines the effect of A/B block scheduling. The measures of English achievement vary across the studies from English as a whole, to reading and language arts. All studies have been judged as 'medium' overall on WoE D for this review.

Figure 4.14 shows the effect sizes for studies in this group. Three of the five studies favoured the traditional schedule and the remaining two studies favoured block scheduling. The results of the study by Zhang (2001) favoured block scheduling; however, it was so large (600, 00 plus students) that its results have in effect become the pooled effect size for this category (g=0.02 95% C.I. 0.01 to 0.03). Without the Zhang (2001) study the effect size became negative and the 95% confidence interval did not exclude zero. Our interpretation of this result was that, for all practical purposes, there was no difference in effect between the block scheduling and traditional schedules.

4.4.3.8 What is the effect of 4 x 4 block scheduling on English achievement?

Four studies evaluated the effect of 4 x 4 block scheduling on English achievement (Lewis et al., 2003, Language Arts; Schreiber et al., 2001, reading and writing; Texas Education Agency, 1999, reading; Zhang, 2001, English). All the studies were judged to be medium overall (WoE D) for this review.

The pooled estimate of effect for 4 x 4 block scheduling compared with the traditional curriculum in English (when the large study, Zhang 2001, referred to above was removed) favoured the traditional schedule (g= -0.08 95% C.I. -0.27 to

Figure 4.15: Effect sizes (Hedges' g) of 4 x 4 block scheduling on English achievement (excluding Zhang, 2001)

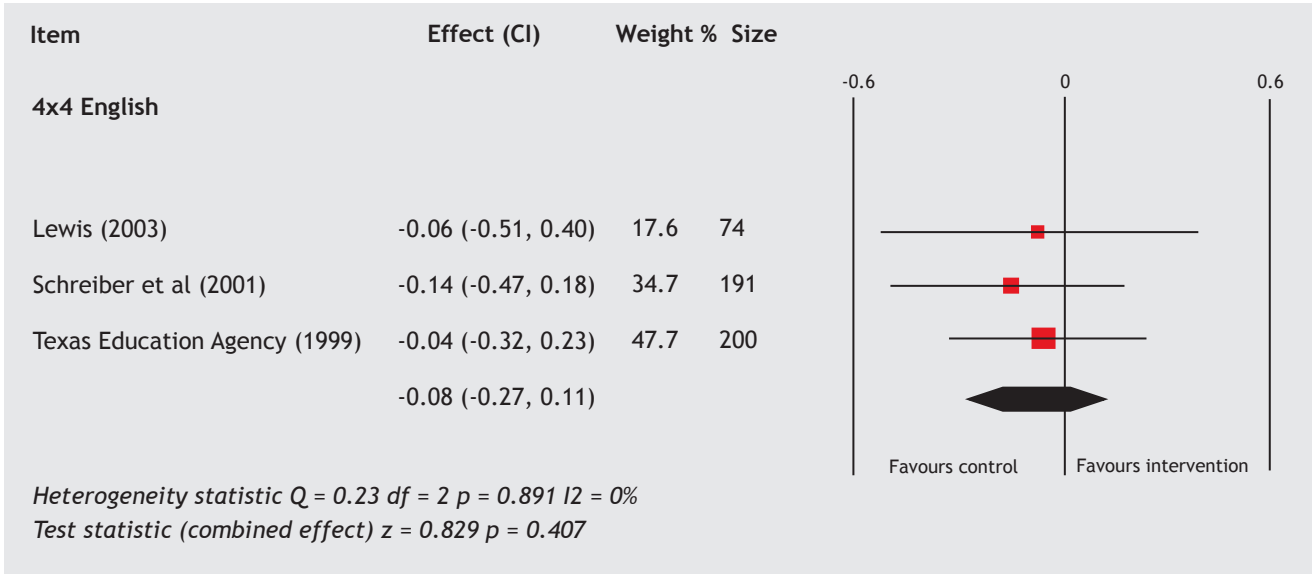
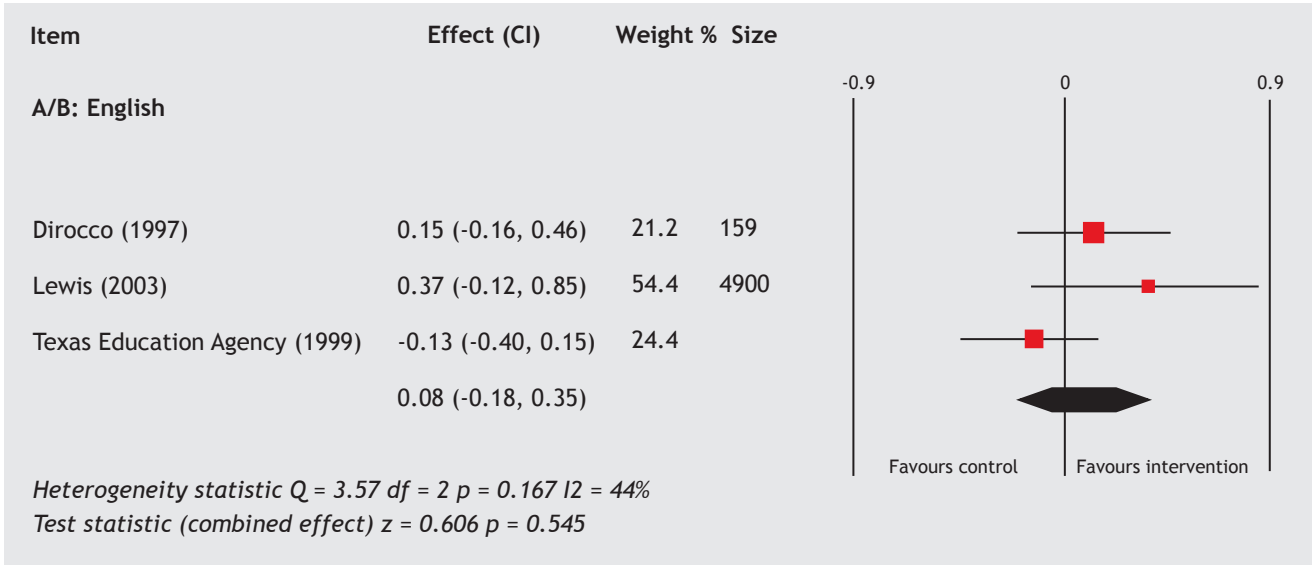


Figure 4.16: Effect sizes (Hedges' g) of A/B block scheduling on English achievement



0.11). However, as the 95% confidence interval crosses zero, the result does not exclude a possible positive effect.

4.4.3.9 What is the effect of A/B block scheduling on English achievement?

Three studies look at the effect of A/B block scheduling on English achievement compared with traditional schedules (DiRocco, 1997; Lewis et al., 2003; Texas Education Agency, 1999) and were all weighted medium for this review.

Figure 4.16 shows the effect sizes for studies in this group. The pooled estimate of effect for A/B block scheduling compared with the traditional curriculum favoured block scheduling (g=0.08 95% C.I. -0.18 to 0.35). However as the 95% confidence interval crosses zero the result does not exclude a possible negative effect.

4.5 Quality-assurance results

4.5.1 Data extraction for in-depth review

As outlined in Chapter 2, data extraction and assessment of the weight of evidence for this review were conducted by pairs of Review Group members (KD and KB; KB and NK; KD and NK) working independently and then comparing their decisions, before coming to a consensus. Some disagreement occurred in categories for study method. This was thought to be due to the difficulty of categorising studies as either experiments with non-random allocation, or cohort studies or case control studies. Some disagreement also occurred when deciding if the authors had used a sampling frame, what the sampling frame was, and how authors had selected participants to be included in the study.

CHAPTER FIVE

Implications

5.1 Approach and summary

Our approach to identifying possible implications for policy and practice used an interpretation framework to group interventions according to the strength of evidence and direction of effect of each set of results. (See Appendix 2.6 for more details. None of the block scheduling subject/outcome groupings analysed in the in-depth review met the criteria to be classified as 'strong' evidence of in/effectiveness. Where studies investigating a particular type of block scheduling and/or a particular outcome and/or a particular subject could not be combined due to high levels of statistical heterogeneity, the interpretation given was 'insufficient evidence'.

5.2 Implications for policy and practice

5.2.1 Evidence of effectiveness (positive or negative)

There were a number of block scheduling type/subject/outcome groupings for which there was evidence of effectiveness.

4 x 4 block scheduling on average cross-subject achievement

There was evidence to suggest that 4 x 4 block scheduling had a (g=0.15) positive effect compared with traditional schedules on student average cross-subject achievement. This finding is based on seven studies with a total population of 642,152 students.

A/B block scheduling on science achievement

There is evidence to suggest that, when compared with traditional schedules, A/B block scheduling also had a positive effect (g=0.20) and therefore could improve academic performance in Science subjects. This finding is based on a total population of 5,337 students, aged 11 -16 from across a range of US middle and high schools. Students attended Science classes of 70-90 minutes every other day, and their performance was measured by state-wide or school-specific testing.

5.2.2 Limited evidence (positive or negative)

Although there is limited evidence to suggest that block scheduling improves cross-subject achievement and student cross-subject exam/grade scores, when exploring the effect of particular types of block scheduling in individual subjects, such as Mathematics and English, the findings were inconsistent. The point estimates of effect suggest that, if students are going to achieve higher results, in Mathematics, it is more likely be a consequence of enrolment in traditional, not block, scheduling. For English, there was a small positive effect in favour of block scheduling. However, this was attributable to just one study; when that study was removed, similar to the case of Mathematics, the traditional schedule was more effective. However, these findings must be treated with caution as the pooled estimate did not exclude a positive effect of block scheduling. Thus, further research needs to be conducted to be able to draw conclusions about the effectiveness of block scheduling in either of these subjects.

5.2.3 Weak evidence (positive or negative)

There is weak evidence on the effectiveness of A/B block scheduling on achievement in English and on average cross-subject student grades. Although both sets of findings appear to suggest that participating in A/B block scheduling can have a positive impact, the findings do not exclude the possibility of no difference between the two groups. The evidence is also weakened by the fact that the combined student population for each meta-analysis is fewer than 500 students.

Strong evidence of effectiveness (positive or negative)

None of the meta-analyses met the criteria to be considered strong evidence of effectiveness.

Evidence of effectiveness (positive or negative) on the effect of:

- 4 x 4 block scheduling on average cross subject achievement
- A/B block scheduling on science achievement

Limited evidence: potential effects (positive or negative) on the effect of:

- Block scheduling on student grade/grade exam scores
- 4 x 4 scheduling on student grade/exam scores
- 4x4 block scheduling on Mathematics achievement
- A/B block scheduling on Mathematics achievement
- block scheduling on English achievement
- 4 x 4 block scheduling on English achievement

Weak evidence: on the effect of:

- A/B block scheduling on student grade/exam scores
- A/B block scheduling on English achievement

Insufficient evidence: on the effect of:

- block scheduling on average cross-subject achievement
- A/B block scheduling on average cross-subject achievement
- block scheduling on Mathematics
- block scheduling on Science
- 4 x 4 block scheduling on Science

5.2.4 Insufficient evidence

In cases in which tests revealed significant statistical heterogeneity (I2= <70%) between the studies, data was not combined to produce a pooled summary effect size. This meant we had insufficient evidence on the effect of block scheduling and A/B block scheduling on academic achievement (cross subject), general block scheduling on achievement in Mathematics, and general and 4 x 4 block scheduling achievement in Science.

5.2.5 Conclusion

Overall, the evidence on the effectiveness of block scheduling compared with the traditional model of timetabling is inconclusive. Although there was some evidence to suggest that different types of school schedules have an impact, in some subjects, the positive effects are not strong enough for making recommendations. In most cases, there was no clear evidence to suggest that organising lessons into blocks (4 x 4 or A/B) would be more or less effective than structuring the school day using existing traditional schedules.

The majority of studies included in the in-depth review used retrospective study designs to evaluate the effect of block scheduling. In most cases, block scheduling was implemented at the school level and applied to all subjects being taught in schools. However, studies only reported outcomes for one or more of the core subjects, such as Mathematics (the most common outcome measured), English or Science. Only a few studies reported outcomes in subjects such as History or Social Sciences; however, this was rare. Therefore, from the studies we have synthesised, it has not been possible to ascertain whether the authors were selective in the outcomes they chose to report. Nor is it possible to judge whether implementing block scheduling could show positive results in one subject, but actually cause harm in other subjects.

It is important to be cautious about the findings for the effect of block scheduling on cross- subject academic achievement exam/grade scores. There are concerns with this way of measuring 'average' academic achievement because it could be concealing success in some subjects and failures in others. Further exploration of the effect of block scheduling in individual subject highlighted this potential problem: although overall effects were positive for Science, they were mixed for Mathematics and English. Interestingly, both studies excluded from the meta-analysis (Schroth and Dixon, 1999; Veal, 1999) confirm this interpretation. Veal (1999) found that students enrolled in hybrid block schedules had higher cross-subject exam/grades scores, yet Schroth and Dixon (1995) found that test results for students enrolled in block Mathematics lessons did not improve.

Thus, this review does not provide conclusive evidence to support the introduction of policy guidance on the use of block scheduling in secondary schools for all subject areas. Although the findings do not indicate that participating in block schedules would produce negative outcomes for pupils across subjects, policy initiatives should not advocate the use of block scheduling as an effective approach to improving academic achievement without due caution. These findings are not dissimilar to the meta-analysis conducted by Lewis et al. (2003) on block scheduling in high schools. They concluded that the findings were not strong enough (small effect sizes) and therefore did not drawn any practice implications.

5.3 Implication for research

The following areas of future research have been identified from this review:

A systematic review of the evidence on the implementation of block scheduling in schools and the identification of the mechanisms that might contribute to its effectiveness or lack of effectiveness

A systematic review of the evidence on teachers' and students' perception of block scheduling and consideration of whether there is a relationship between different types of school schedules and the experience of teaching and learning

Primary research on the impact of the length of lessons and curriculum coverage on teaching and learning in the UK, and consideration of whether findings vary by age, gender, ethnicity, social class or ability of students.

The effect sizes identified might appear small but, if implemented on a national scale, could have a large practical consequence: for example, in terms of increasing the proportion of students who passed their exams at grade C. However, the quality of the studies is such that the positive effect seen may be entirely an artefact of the study design. It would therefore be important to conduct primary research on the effectiveness of block scheduling using prospective randomisation of pupils to block and non-block schedules.

5.4 Strengths and limitations of this systematic review

The main strengths of the review are as follows:

The review process is transparent, as well as replicable and updateable. The explicit reporting of the methods allow for this review to be replicated and its findings to be critically appraised.

The review's extensive search strategy aimed to pick up academic and grey literature.

Quality assurance is paramount in the review process. During the screening stage, judgements were made and agreed by two, sometimes three, reviewers for 6% of all identified studies. Key items of the data extraction for the mapping stage and all items for the in-depth review were double-coded. The quality appraisal and synthesis stages were undertaken by two reviewers, with outstanding issues discussed with a third reviewer.

The involvement of the commissioners of the review helped to make the review more policy-relevant.

Careful consideration was given to the quality of the evidence. Each study was subject to thorough assessment, being judged independently by two reviewers with a final judgement being agreed through deliberation. Each study in the in-depth review was carefully judged according to three dimensions:

Trustworthiness of the findings. An adapted version of the Home Office Quality Assessment Tool was used to judge the validity, bias and appropriateness of each study design for measuring effectiveness.

Appropriateness of the study design for addressing the review question. Studies were judged according to how well the sample selection and analysis procedures controlled for differences between intervention and control groups.

Relevance of the study for the review, based on its sample size and representativeness

The presentation of the study results as effect sizes facilitates direct comparison and synthesis of results across similar interventions.

The main limitations of the review are as follows:

There were only a few high quality studies that measured effectiveness of block scheduling. The conclusions of this review are based upon studies whose overall quality has been judged to be 'medium'. The absence of randomised controlled trials or high quality experimental study designs

from this literature means that the body of evidence is not as robust as we would like when trying to answer a question about effectiveness.

Studies were limited to those published in English.

The focus of the included studies in the map and in-depth review was on the US. A high proportion of the studies included in the map were conducted in the US, with all but one study in the in-depth review conducted in the US. The in-depth review is therefore limited to the focus of this research. Reflecting the policy initiatives of the US, this set of studies focuses on certain interventions and topics (e.g. block scheduling, extended school year). This, therefore, limits the scope of the synthesis and has implications for the findings of this review:

There are limitations on the transferability of the review findings to the UK policy context. Alternate block scheduling (where students take different academic subjects on alternate days), for example, may be difficult to apply to UK primary schools in a policy context that requires English/Literacy classes daily (Literacy Hour).

It was not always clear precisely how the experimental interventions differed from the control interventions, so we cannot be sure that an experimental intervention in one study was not identical to a control intervention, nor do we know to what extent control programmes in the studies match current policy and practice in the UK context.

User involvement was limited to the commissioners of the review, which excluded those who would be directly impacted by the findings of this review, namely students and teachers.

CHAPTER SIX

References

6.1 Studies included in map and synthesis

In-depth review

Cobb RB, Abate S, Baker D (1999) Effects on students of a 4 x 4 junior high school block scheduling program. Education Policy Analysis Archives 7(3).

DiRocco MD (1997) Effects of alternating day block scheduling on student academic performance, attendance and discipline (Alternative day scheduling, intensive scheduling). Dissertation thesis. Pennsylvania State University.

Hughes WW (2004) Blocking student performance in high school? *Economics of Education Review* **23**(6): 663-667.

Lapkin S, Harley B, Hart D (1997) Block scheduling for language study in middle grades: a summary of the Carleton case study. *Learning Languages* **2**(3): 4-8.

Lewis CW, Cobb BR, Winokur MA, Leech N, Viney M (2003) The Effects of Full and Alternative Day Block Scheduling on Language Arts and Science Achievement in a Junior High School. Education Policy Analysis Archives 11(41)

Marchette F (2003) Impacts of scheduling configurations on Mississippi biology subject area testing. Paper presented at: Annual Meeting of the Mid-South Educational Research Association, Biloxi, MS.

McCreary J, Hausman C (2001) Differences in student outcomes between block, semester, and trimester schedules. Salt Lake City: University of Utah.

Rice JK, Croninger RG, Roellke CF (2002) The effect of block scheduling high school mathematics courses on student achievement and teachers' use of time: implications for educational productivity. *Economics of Education Review* 21: 599–607.

Schreiber JB, Veal WR, Flinders DJ, Churchill S (2001) Second year analysis of a hybrid schedule high school. Education Policy Analysis Archives, 9(46). http://epaa.asu.edu/ojs/article/view/375 (accessed 25 January 2010).

Schroth G, Dixon J (1995) The effects of block scheduling on seventh grade mathematics students. Dallas, TX: East Texas State University.

Texas Education Agency (1999) Block scheduling in Texas public high schools. Policy Research Report No. 13. Austin, TX: TEA.

Veal WR (1999) What could define block scheduling as a fad? *American Secondary Education* **27**(4): 3-12.

Walker G (2000) The effect of block scheduling on mathematics achievement in high and low SES secondary schools. Doctoral dissertation, University of Kansas. Dissertation Abstracts International 61: 4638.

Zhang G (2001) Academic performance differences between students in block and traditionally scheduled high schools 1993-2000. Paper presented at: The Annual Meeting of the American Educational Research Association, Seattle: 10-14 April.

Studies included in the map

Ahl A (2007) Equality and the Freedom to Choose the 'What and When' of Schooling: Students with Special Educational Needs in Timetable-Free Settings. *European Journal of Special Needs Education* **22**(2):183–197.

Alber-Kelsay K (1998) Full-day kindergarten vs half-day kindergarten: the outcome of first grade reading achievement. Unpublished Master's thesis. Ohio State Department of Education.

Allinder RM, Bolling R, Oats R, Gagnon W (1998) An analysis of alternative means of measuring student achievement prior to a scheduled school break. Special Services in the Schools 14: 51-62.

Ammons TL, Booker JL, Killmon CP (1995) The effects of time of day on student attention and achievement. Unpublished manuscript. University of Virginia, Curry School of Education.

Ananda N (1997) Positive effects of intersession tutoring in a year-round school. Paper presented at: The Annual Meeting of the American Educational Research Association, Chicago: 24-28 March.

Anderson J (1994) Alternative approaches to organizing the school day and year: a national commission examines new structures for improving student learning. The School Administrator 15: 8-11.

Appalachia Educational Laboratory, United States Department of Education (1997) Let's ask the students: Kentucky, Tennessee, Virginia, and West Virginia Students talk about schools and change. Washington, DC: Office of Educational Research and Improvement, Educational Resources Information Centre.

Arnold DE (2002) Block schedule and traditional schedule achievement: a comparison. NASSP Bulletin 86: 42-53.

Axelrad-Lentz SF (1996) Michigan extended school year programs 1992-1995: an evaluation of a state grant initiative. Bloomfield, MI: Greentree Research and Development.

Barrett T (1992) Results of the year-round education parent, staff and student surveys. California: Department of Educational Accountability.

Benson JG, Borman GD (2007) Family and contextual socioeconomic effects across seasons: when do they matter for the achievement growth of young children? WCER Working Paper No. 2007-5, August.

Betts JR, Zau AC, King K (2005) From blueprint to reality: San Diego's education reforms. San Francisco, CA: Public Policy Institute of California.

Bishop J, Worner W, Weber L (1988) Extending the school day: an evaluation of a seven-period class schedule. Studies in Educational Evaluation 14: 361–380.

Bryant JG Jr, Claxton DB (1996) Physical education and the four-by-four schedule. *Physical Educator* **53**(4): 203-209.

Buckman DC, King BB, Ryan S (1995) Block scheduling: a means to improve school climate. NASSP Bulletin 79: 9-18.

Byrd JK (2001) Student achievement: is equality really necessary? Paper presented at: Annual Meeting of the American Educational Research Association, Seattle, WA, 10-14 April.

Callan RJ (1999) Effects of matching and mismatching students' time-of-day preferences. *Journal of Educational Research* **92**: 295-299.

Calvery R, Sheets G, Bell D (1999) Modified block scheduling: an assessment of teachers and students perception. Paper presented at: Meeting of the Mid-South Educational Research Association, Point Clear, AL.

Carroll JM (1994) The Copernican Plan evaluated: the evolution of a revolution. Phi Delta Kappan 76(2): 104-110.

Cobb RB, Abate S, Baker D (1999) Effects on students of a 4 x 4 junior high school block scheduling program. Education Policy Analysis Archives 7(3).

Corley EL (2003) A quantitative look at student attitudes/perceptions about block scheduling. Paper presented at: Annual Meeting of the Mid-Western Educational Research Association, Columbus, OH.

Creamean SL, Horvath RJ (2000) The effectiveness of block scheduling. Master of Arts Action Research Project, Saint Xavier University Field-Based Masters Program.

Daneshvary N, Clauretie TM (2001) Efficiency and costs in education: year-round versus traditional schedules. Economics of Education Review 3: 279-287.

Davis-Wiley P, Cozart A (1996) Block scheduling in the secondary arena part ii: perceptions from the inside. Paper presented at: Annual Meeting of the Mid-Western Educational Research Association, Tuscaloosa, AL.

DiRocco MD (1997) Effects of alternating day block scheduling on student academic performance, attendance and discipline (Alternative day scheduling, intensive scheduling). Dissertation thesis. Pennsylvania State University.

Dossett D, Munoz M (2000) Year-round education in a reform environment: the impact on student achievement and cost-effectiveness analysis. Louisville, KY: Department of Leadership, Foundations and Human Resource Education, University of Louisville.

Eineder DV, Bishop HL (1997) Block scheduling the high school: the effects on achievement, behavior, and student-teacher relationships. NASSP Bulletin 81(589): 45-54.

Engin A (2006) Does the weekly lesson program affect the language learning? http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/27/f4/63.pdf (accessed 25 January 2010).

EPPI-Centre (2002) Core keywording strategy: data collection for a register of educational research. Version 0.9.7. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Eren OM, Daniel L (2007) Time to learn? The organizational structure of schools and student achievement. Empirical Economics 2-3: 301-332.

Evans J (1996) Children's attitudes to recess and the changes taking place in Australian primary schools. Research in Education 56: 49-61.

Evans W, Tokarczyk J, Rice S, McCray A (2002) Block scheduling: an evaluation of outcomes and impact. Clearing House 75(6): 319-323.

Farbman D (2005) Time for a change: the promise of extended-time schools for promoting student achievement. Boston, MA: Massachusetts 2020. http://www.mass2020.org/full_report.pdf (accessed 25 January 2010).

Fardig D (1992) Year-round education: program evaluation report. Orlando, FL: Orange County Public Schools.

Finn JD, Pannozzo GM (2004) Classroom organization and student behavior in kindergarten. *Journal of Educational Research* **2**: 79-92.

Fletcher RK, Jr (1997) A study of the block scheduling movement in six high schools in the Upper Cumberland Region of Tennessee. Revision of paper presented at: Annual Meeting of the Tennessee Academy of Science. Sewanee: 22 November

Fouts JT (1994) A school within a school: evaluation results of the first year of a restructuring effort. Seattle, WA: School of Education, Seattle Pacific University.

Frazier JA, Morrison FJ (1998) The influence of extended-year schooling on growth of achievement and perceived competence in early elementary school. *Child Development* **69**(2): 495-497.

From J, Holmgren C, Hakan A, Ahl A (2003) Are timetable-free schools possible? European Educational Research Journal 2: 547-558.

Geismar TJ, Pullease BG (1996) The trimester: a competency based model of block scheduling. Research Brief. NASSP Bulletin 80: 95-105.

Green CA (1998) The extended school year program consolidated report: achievement test scores and survey findings. Detroit, MI: Office of Research, Evaluation and Assessment, Detroit Public Schools.

Griffin LL, Nicholson JW, Jr (2002) An evaluation of the block schedule in two high schools. Paper presented at: Meeting of the Mid-South Educational Research Association, Chattanooga, TN.

Guskey TR, Kifer E (1995) Evaluation of a High School Block Schedule Restructuring Program. Paper presented at: Annual Meeting of the American Educational Research Association, San Francisco, CA

Hackmann DG, Waters DL (1998) Breaking away from tradition: the Farmington High School restructuring experience. NASSP Bulletin 82(596): 83-92.

Harmston MT, Pliska A, Ziomeck RL, Hackman DG (2003) The relationship between schedule type and ACT assessment scores: a longitudinal study. ACT Research Report Series. http://www.act.org/research/reports/pdf/ACT_RR2003-3.pdf (accessed 25 January 2010).

Holmes CT, McConnell BM (1990) Full-day versus half-day kindergarten: an experimental study. Paper presented at: Annual Meeting of the American Educational Research Association, Boston, MA.

Hough D (1990) Cost-Effects analysis of year-round education programs. Paper presented at: Annual Meeting of the American Educational Research Association, Boston, MA: 16-20 April.

Houston Independent School District (1997) Core curriculum, 1996-97: research report on an educational program. Austin, TX: Texas Department of Research and Accountability.

Hughes WW (2004) Blocking student performance in high school? Economics of Education Review 23(6): 663-667.

Idol L (1998) Optional extended year program feedback. Austin, TX: Austin Independent School District.

Klein J (2001) Attention, scholastic achievement and timing of lessons. *Scandinavian Journal of Educational Research* **45**(3): 301-309.

Klein J (2004) Planning middle school schedules for improved attention and achievement. *Scandinavian Journal of Educational Research* **48**(4): 441-450.

Knudson RE (1995) Year-round school: are there student differences? Paper presented at: Annual Meeting of the American Educational Research Association, San Francisco, CA.

Koopmans M (1991) A study of the longitudinal effects of all-day kindergarten attendance on achievement. Newark, NJ. Newark Board of Education, Office of Research, Evaluation and Testing.

Kubitschek WN, Hallinan MT, Arnett SM, Galipeau KS (2005) High school schedule changes and the effect of lost instructional time on achievement. *High School Journal* **89**(1): 63-71.

Kubow PK, Wahlstrom KL, Bemis AE (1999) Starting time and school life: reflections from educators and students. Phi Delta Kappan 80(5): 366-371.

Lapkin S, Harley B, Hart D (1997) Block scheduling for language study in middle grades: a summary of the Carleton case study. Learning Languages 2(3): 4-8.

Le V-N, Kirby SN, Barney H, Setodji CM, Gershwin D (2006) School readiness, full-day kindergarten, and student achievement: an empirical investigation, Washington DC: Rand Corporation. http://www.rand.org/pubs/monographs/MG558/

Lee VE, Burkam DT, Honigman J, Ready DD, Meisels SJ (2006) Full-day vs half-day kindergarten: in which program do children learn more? American Journal of Education 112: 163-208.

Lewis CW, Cobb BR, Winokur MA, Leech N, Viney M (2003) The Effects of Full and Alternative Day Block Scheduling on Language Arts and Science Achievement in a Junior High School. Education Policy Analysis Archives 11(41)

Lister RJ (1997) Block scheduling at Portsmouth High School: a status report. Portsmouth, NH: Portsmouth School Administrative Unit.

Liu J, Dye JF (1998) Teacher and student attitudes toward block scheduling in a rural school district. American Secondary Education 26(3): 1-7.

Lockwood S (1995) Semesterizing the high school schedule: the impact on student achievement in algebra and geometry. NASSP Bulletin 79(575): 102-110.

Maltese AV, Dexter KM, Tai RH, Sadler PM (2007) Breaking from tradition: unfulfilled promises of block scheduling in science. Science Educator 16: 1-7.

Marchette F (2003) Impacts of scheduling configurations on Mississippi biology subject area testing. Paper presented at: Annual Meeting of the Mid-South Educational Research Association, Biloxi, MS.

Mattox K, Hancock DR, Queen JA (2005) The effect of block scheduling on middle school students' mathematics achievement. NASSP Bulletin 89(642): 3-13.

McCoy LP (1995) Studies in teaching: 1995 research digest. Paper presented at: Annual Research Forum, Winston-Salem, North Carolina.

McCoy M, Helen S (1998) Block scheduling: does it make a difference? A high school case study. Paper presented at: Annual Meeting of the Southwest Educational Research Association, Houston, TX.

McCreary J, Hausman C (2001) Differences in student outcomes between block, semester, and trimester schedules. Salt Lake City, UT: University of Utah.

Meehan ML, Cowley KS, Schumacher D, Hauser B (2003) Evaluation of Kentucky's Extended School Services Program. Paper presented at: Annual CREATE National Evaluation Institute, Louisville, KY.

Meyer LA, Hastings CN, Wardrop J, Linn RL (1989) How entering ability and instructional settings and not the length of the school day mediate kindergartners' reading performance. Technical Report No. 466. http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1e/8b/a1.pdf (accessed 25 January 2010).

Mitchell RE, Mitchell DE (1999) Student segregation and achievement tracking in year-round schools. Paper presented at: Annual Meeting of the American Sociological Association, Chicago, IL, 6 August.

Morris RN (2002) A case study on the perspectives of an optional k-5 year-round/multi-age program in Virginia. http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/4a/22.pdf (accessed 25 January 2010).

Mutter DW (1996) 4 X 4 block schedule evaluation. Chesapeake, VA: Chesapeake Public Schools, Office of Program Evaluation.

Muyskens P, Ysseldyke JE (1998) Student academic responding time as a function of classroom ecology and time of day. *Journal of Special Education* **4**: 411-424.

Neukam B, Gilman DA (1996) Language arts instruction time. http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/90/2b.pdf (accessed 25 January 2010).

New York City Board of Education (2000) Analyses of performance of extended-time and non-extended time SURR schools. http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/16/e8/2d.pdf (accessed 25 January 2010).

Nichols JD (2005) Block-scheduled high schools: impact on achievement in English and language arts. *Journal of Education Research* **98I**(5): 299-309.

North Carolina State Department of Public Instruction (1994) Block scheduling in North Carolina High Schools. Raleigh, NC: Division of Innovation and Development Services.

Nunnelley J (1996) The impact of half-day versus full-day kindergarten programs on student outcomes: a pilot project. Project report. http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/96/4d.pdf (accessed 25 January 2010).

Nyroos M, Rönnberg L, Lundahl L (2004) A matter of timing: time use, freedom and influence in school from a pupil perspective. *European Educational Research Journal* **3**(4) 743-758.

Ohio Department of Education (1992) A longitudinal research study of the effects of preschool attendance and kindergarten schedule: kindergarten through grade four. Columbus, OH: Division of Early Childhood Education, Ohio Department of Education.

Payne D, Jordan M (1996) The evaluation of a high school block schedule: convergence of teacher and student data. *American Secondary Education* **25**(2): 16-19.

Pellegrini AD, Huberty PD, Jones I, (1995) The effects of recess timing on children's playground and classroom behaviors. *American Educational Research Journal* **32**(4): 845-864.

Pisapia J, Westfall AL (1997) Alternative high school scheduling: a view from the student's desk. Richmond, VA: Metropolitan Educational Research Consortium

Pittman RB, Herzog MJR (1998) Evaluation of a year-round schedule in a rural school district. *Journal of Research in Rural Education* **14**(1): 15-25.

Policy Studies Associates, Inc. (1994) The uses of time for teaching and learning. Volume II: case studies. Washington, DC: Studies of Education Reform.

Prohm B, Baenen N (1996) Are multi-track year round schools effective? Wake County Public Schools, NC: Evaluation and Research Department.

Randler C, Kranich K , Eisele MAF, Randler CK, Konstanze EM (2008) Block scheduled versus traditional biology teaching: an educational experiment using the water lily. *Instructional Science* **36l**(1): 17-25.

Reece, JL (2000) Retention of academic skills over the summer months in alternative and traditional calendar school. *Journal of Research and Development in Education* **33**:166-174.

Reed, D (2002) Description of success: a four-teacher instructional model. Sam Houston State University. Houston, TX.

Reid, L (1995) Perceived effects of block scheduling on the teaching of English. Fort Collins, CO: Colorado State University.

Rice JK, Croninger RG, Roellke, CF (2002) The effect of block scheduling high school mathematics courses on student achievement and teachers' use of time: implications for educational productivity. *Economics of Education Review* **21**(6): 599-607.

Ritter C (1992) The effects of year-round school calendar on gifted and talented students. Unpublished master's thesis. Sam Houston State University. Houston, TX.

Robinson-Lewis G (1991) Full day kindergarten program: 1988-1989, 1989-1990, 1990-1991: summative evaluation. Kansas, MO: Kansas City Public Schools.

Roth JL, Brooks-Gunn J, Linver MR, Hofferth SL (2003) What happens during the school day? Time diaries from a national sample of elementary school teachers. Teachers College Record 105(3): 317-343.

Roth WM, Tobin K, Ritchie SM (2008) Time and temporality as mediators of science learning. *Science Education* **92**(1): 115-140.

Russell JF (1997) Relationships between the implementation of middle-level program concepts and students' achievement. *Journal of Curriculum and Supervision* **12**(2): 169-185.

Saunders WM, Foorman BR, Carlson CD (2006) Is a separate block of time for oral English language development in programs for English learners needed? *Elementary School Journal* **107**(2): 181-198.

Schreiber JB, Veal WR, Flinders DJ, Churchill S (2001) Second year analysis of a hybrid schedule high school. Education Policy Analysis Archives 9(46). http://epaa.asu.edu/epaa/v9n46/ (accessed 25 January 2010).

Schroth G, Dixon J (1995) The effects of block scheduling on seventh grade mathematics students. Dallas. TX: East Texas State University.

Sharman RG (1989) Student satisfaction and achievement related to organizational structure: a study on semestering in junior high. *Education Canada* **29**(3): 28-35.

Shields C, Oberg SL (1999) What can we learn from the data? Toward a better understanding of the effects of multitrack year-round schooling. *Urban Education* **34**(2): 125-154.

Sjosten-Bell W (2005) Influence of time-of-day on student performance on mathematical algorithms, Division of Education. Unpublished master's thesis. University of California.

Smith DL, McNelis MJ (1990) The impact of a seven-period school day on teachers and students. Paper presented at: Annual Meeting of the American Educational Research Association, Boston, MA: 16-20 April.

Snyder D (1997) 4-block scheduling: a case study of data analysis of one high school after two years. Paper presented at: Annual Meeting of the Midwest Educational Research Association, Chicago, IL: 15-18 October.

Southern Regional Education Board (1997a) Case study: Hoke County High School, Raeford, North Carolina. Atlanta, GA: SREB.

Southern Regional Education Board (1997b) Case study: North Laurel High School, London, Kentucky. Atlanta, GA: SREB.

Southern Regional Education Board (1997c) Case study: Sussex Technical High School Georgetown, Delaware. Atlanta, GA: SREB.

Spellman JE, Oliver JS (2001) The relationship between attitude toward science with enrollment in a 4 x 4 block schedule. Paper presented at: Annual Meeting of the Association for the Education of Teachers in Science, Costa Mesa, CA: 18-21 January.

Stader DL (2001) Block scheduling in small high schools: perceptions from the field. *Rural Educator* **22**(3): 37-41.

Stenvall JT, Stenvall MJ (2001) An analysis of 2000 API scores for California public schools on traditional and year-round calendars at the elementary, middle and high school levels. San Diego, CA: National Association for Year Round Education

Sumpner C, Blatchford P (1998) What do we know about breaktime? Results from a national survey of breaktime and lunchtime in primary and secondary schools. *British Educational Research Journal* **24**(1): 79-94.

Texas Education Agency (1999) Block scheduling in Texas public high schools. Policy Research Report No. 13. Austin, TX: TEA.

Thomas C, O'Connell RW (1997) Student perceptions of block scheduling in a New York State public high school. Paper presented at: Annual Meeting of the Northeastern Educational Research Association, Ellenville, NY, 22-24 October.

Trenta L, Newman I (2002) Effects of a high school block scheduling program on students: a four-year longitudinal study of the effects of block scheduling on student outcome variables. *American Secondary Education* **31**(1): 54-71.

Ullrich WJ, Yeamen JT (1999) Using a modified block schedule to create a positive learning environment. *Middle School Journal* **31**(1): 14-20.

VanSciver JH (1999) Extending the school year into the summer. *Principal* **78**(4): 60-63.

Veal WR (1999) What could define block scheduling as a fad? *American Secondary Education* **27**(4): 3-12.

Walker, G (2000) The effect of block scheduling on mathematics achievement in high and low SES secondary schools. Unpublished doctoral dissertation, University of Kansas. Dissertation Abstracts International 61: 4638.

Walston JT, West J (2004) Full-day and half-day kindergarten in the United States: findings from the Early Childhood Longitudinal Study, kindergarten class of 1998–99 (NCES 2004–078). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Washington W (2000) Optional extended year program feedback, Austin, TX: Austin Independent School District.

Westlund I (2007) Reconstructed task orientation and local time governance in compulsory schools: the Swedish case. *British Journal of Sociology of Education* **6**: 783–796.

Wilson J, Stokes LC (2000) Students' perceptions of the effectiveness of block versus traditional scheduling. *American Secondary Education* **28**(3): 3-12.

Yan W, Lin Q (2005) Effects of class size and length of day on kindergartners' academic achievement: findings from early childhood. Longitudinal study. Early Education and Development 16(1): 49-68.

Zakaluk BL, Straw SB (2002) The efficacy of an extended-day kindergarten program: a report for the St. James School Division (1999-2000, 2000-2001). Winnipeg: Faculty of Education, University of Manitoba.

Zhang G (2001) Academic performance differences between students in block and traditionally scheduled high schools 1993-2000. Paper presented at: The Annual Meeting of the American Educational Research Association, Seattle: 10-14 April.

Zvoch K, Reynolds RE, Parker RP (2008) Full-day kindergarten and student literacy growth: does a lengthened school day make a difference? *Early Childhood Research Quarterly* **23**(1): 94-107.

6.2 Other references used in the text of the technical report

Callan R (1998) Giving students the (right) time of day. *Educational Leadership* **55** (4): 84-87.

DCSF (2007) The Children's Plan: building brighter futures. http://www.dcsf.gov.uk/childrensplan/downloads/The_Childrens_Plan.pdf (accessed 25 January 2010).

DCSF (2008a) Guidance for local authorities on setting education performance targets: Part 1: LA statutory targets for Key Stages 2, 3, 4, early years' outcomes, children in care, black and minority ethnic groups, attendance. http://www.standards.dfes.gov.uk/ts/docs/part1guidance08.pdf (accessed 25 January 2010).

DCSF (2008b) Guidance for local authorities on setting education performance targets: Part 2: Target setting with schools. http://www.standards.dfes.gov.uk/ts/docs/part2guidance08.doc (accessed 25 January 2010).

DfES (2002) Designing and timetabling the primary curriculum: a practical guide for Key Stages 1 and 2. Coventry: QCA.

DfES (2004) A condensed Key Stage 3: designing a flexible curriculum. Coventry: QCA.

DfES (2006) A condensed Key Stage 3: designing a flexible curriculum 2006 update. Coventry: QCA.

Dunn R, Griggs S (1990) Research on the learning style characteristics of selected racial and ethnic groups. *Reading and Writing Quarterly* **6**(3): 261-280.

Eurybase (2008) The education system in England, Wales, Northern Ireland 2007/2008. Brussels: European Commission.

Harp T, Orsak L (1990) One administrator's challenge: implementing a learning style program at the secondary level. Reading and Writing Quarterly 6(3): 335-341.

Independent Commission on the Organisation of the School Year (2002) The rhythms of schooling: detailed recommendations for school terms and holidays in 2003-04 http://www.esinet.norfolk.gov.uk/cadmin/misheets/MI61_02att3.htm (accessed 25 January 2010).

Lam-Phoon S (1986) A comparative study of the learning styles of Southeast Asian and American Caucasian college students of two Seventh-day Adventist campuses. Unpublished doctoral dissertation. Andrews University.

LGA (2004) Standard school year statement. http://www.lga.gov.uk/lga/core/page.do?pageId=119164 (accessed 25 January 2010).

LGA (2007) Standard school year: implementation update http://www.lga.gov.uk/lga/aio/119198

Lynch PK (1981) An analysis of the relationships among academic achievement, attendance and the individual learning style time preferences of eleventh and twelfth grade students identified as initial or chronic truants in a suburban New York school district. Unpublished doctoral dissertation. St. John's University.

Milgram R, Dunn R, Price G (eds) (1993) Teaching gifted and talented learners for learning style: an international perspective. New York: Praeger.

Rothenberg D (1995) Full-day kindergarten programme. http://ceep.crc.uiuc.edu/eecearchive/digests/1995/drkind95.html (accessed 25 January 2010).

Scroggins GV, Karr-Kidwell PJ (1995) Implementation of block scheduling in a four-year high school: a literary review and a handbook for administrators, teachers and parents. http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/34/22.pdf (accessed 25 January 2010).

Shields CM, LaRocque LJ (1996) Literature review on year round schooling (with an annotated bibliography). Vancouver: The University of British Columbia.

Thomas J, Brunton, J (2006) EPPI-Reviewer 3.0: analysis and management of data for research synthesis. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Trenta L, Newman I (2001) Evaluation of an on-going block scheduling programme. Paper presented at: Annual Meeting of the Mid-Western Educational Research Association. Chicago, IL, 24-27 October.

Virostko, J (1983) An analysis of the relationships among academic achievement in mathematics and reading, assigned instructional schedules, and the learning style time preferences of third, fourth, fifth, and sixth grade students. Unpublished doctoral dissertation, St. John's University.

Worthen BR, Zsiray SW (1994) What twenty years of educational studies reveal about year-round education. Chapel Hill, NC: The University of North Carolina at Chapel Hill.

Wheeler, G (1995) A study of half day vs all day pre-kindergarten readiness. Unpublished manuscript. Huntsville, TX: Sam Houston State University.

Appendix 1.1: Authorship of this report

This report was written by Kelly Dickson, Karen Bird, Mark Newman and Naira Kalra at EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

For further information about the work of the EPPI-Centre, please contact:

EPPI-Centre
Social Science Research Unit
Institute of Education, University of London
18 Woburn Square

Tel: +44 (0)20 7612 6397
Fax: +44 (0)20 7612 6800
Email: EPPIAdmin@ioe.ac.uk

Project consultants

Elizabeth McNess, Graduate School of Education, University of Bristol

Marilyn Osborn, Graduate School of Education, University of Bristol

Acknowledgements

The work described in this report was undertaken by the EPPI-Centre, which received funding from the Department for Children, Schools and Families (DCSF) to undertake the work. The views expressed in the report are those of the authors and not necessarily those of the DCSF.

Appendix 2.1: Inclusion and exclusion criteria

Include studies that are:

1.about the structure of time within the school day or year

2.on children and young people aged 5 to 16 in mainstream, maintained and independent schools

3.report empirical data

4.report data on pupils or the school

5.published or reported in English

6.published or reported between 1988 and 2008

Exclude studies that are:

1. not on the structure of time within the school day or year

2. not reporting on children and young people aged 5 to 16

3. not on mainstream, maintained and independent school

4. of the following study type:

a) descriptive

b) methodology

c) editorial, commentary, book review

d) policy document

e) resource, textbook

f) bibliography

g) theoretical paper

h) position paper

i) reviews (systematic and non-systematic)

5. not reporting data on pupils or the school

6. not published or reported in English

7. not published or reported between 1988 and the present

Appendix 2.2: Search strategy for electronic databases

AEI (Australian Education Index)

Via Dialog DataStar

Search date: 12 May 2008

Number of hits: 346

"((pupil$ OR learner$) OR (Children.W..DE.) OR (student NEAR school)) AND ((Organisation.W..DE. OR Class-Organisation.DE. OR Course-Organisation.DE. OR School-Organisation.DE.) OR (timing OR length OR timetabl$ OR schedul$) OR (Time.W..DE. OR Time-On-Task.DE. OR Time-Management.DE. OR Time-Blocks.DE. OR Time-Sharing.DE. OR Time-Factors-Learning.DE.)) AND ((Schools.W..DE.) OR (Classrooms.W..DE.) OR (lesson$ OR recess OR break ADJ time$ OR playtimes OR exam$ OR school ADJ year OR school ADJ term OR school ADJ day) OR (Curriculum.W..DE. OR Curriculum-Development.DE. OR Core-Curriculum.DE. OR Curriculum-Design.DE. OR Curriculum-Mapping.DE. OR Primary-School-Curriculum.DE. OR Secondary-School-Curriculum.DE. OR Student-Centred-Curriculum.DE.))" for information added since 19880101

BEI (British Education Index)

Via Dialog DataStar

Search date: 12 May 2008

Number of hits: 92

"((Schools.W..DE. OR Independent-Schools.DE. OR Comprehensive-Schools.DE. OR Elementary-Schools.DE. OR Grant-Maintained-Schools.DE. OR Secondary-Schools.DE. OR Middle-Schools.DE. OR Junior-Schools.DE. OR Junior-Secondary-Schools.DE. OR Maintained-Schools.DE. OR Primary-Schools.DE. OR Private-Schools.DE.) OR (classrooms) OR (lesson$ OR recess OR break ADJ time$ OR playtimes OR exam$ OR school ADJ year OR school ADJ term OR school ADJ day) OR (Curriculum.W..DE. OR Curriculum-Development.DE. OR Core-Curriculum.DE. OR Curriculum-Design.DE. OR Learner-Centred-Curriculum.DE. OR Primary-School-Curriculum.DE. OR School-Based-Curriculum.DE. OR Secondary-School-Curriculum.DE.)) AND ((Time.W..DE. OR Time-On-Task.DE. OR Time-Management.DE. OR Time-Factors-Learning.DE.) OR (organisation OR timing OR length OR timetabl$ OR schedul$)) AND ((student NEAR school) OR (Pupils.W..DE. OR Middle-School-Pupils.DE. OR Primary-School-Pupils.DE. OR Secondary-School-Pupils.DE.) OR (learner$) OR (Adolescents.W..DE.))"

ERIC (Education Resources Information Centre)

Via CSA

Search date: 9 May 2008

Number of hits: 4192

(DE=(time or organization) or KW=(timing or length or timetabl*) or

KW=schedul*) and(DE=(schools or classrooms) or KW=(lesson* or recess or

(break times)) or KW=(playtimes or exam* or curriculum) or KW=((school

day) or (school year) or (school term))) and(DE=children or KW=(pupil* or

(student* near school) or learner*))

International Bibliography of Social Sciences (IBSS)

Via EBSCO

Search date: 13 May 2008

Number of hits: 28

1 (students or pupil* or learner*).mp. [mp=abstract, title, book title, original title, heading word, subject heading, geographic heading]

2 (timing or length or timetabl* or schedul*).mp. [mp=abstract, title, book title, original title, heading word, subject heading, geographic heading]

3 (schools or classroom* or lesson* or recess or “break times” or playtimes or “school year” or “school day” or “school term” or curriculum or exam or exams or examinations).mp. [mp=abstract, title, book title, original title, heading word, subject heading, geographic heading]

4 1 and 2 and 3

PsycInfo

Search date: 12 May 2008

Number of hits: 3,250

1.students/ or *classmates/ or *elementary school students/ or *high school students/ or *junior high school students/ or *reentry students/ or *special education students/ or *transfer students/ or *vocational school students/

2 organizational behavior/ or exp organizational learning/ or exp organizational structure/

3 exp learning schedules/ or learning strategies/

4 exp learning environment/

5 schools/ or *elementary schools/ or high schools/ or *junior high schools/ or *middle schools/

6 education/ or *elementary education/ or *high school education/ or *middle school education/ or *private school education/ or *public school education/ or *secondary education/ or *special education/

7 *classroom environment/ or exp classroom management/

8 lesson plans/

9 *recreation/

10 exp curriculum based assessment/ or exp curriculum development/

11 timetabl$.mp.

12 timing.mp

13 pupils.mp

18 1 or 5 or 6 or 13

19 2 or 3 or 4 or 7 or 8 or 9 or 11 or 12

20 18 and 19

Social Science Citation Index (SSCI)

Via Web of Knowledge

Search date: 13 May 2008

Number of hits: 148

TS=(student AND school) OR TS=pupil* OR TS=learner* AND Language=(English)

Databases=SSCI Timespan=1988-2008

AND

TS=timing OR TS=length OR TS=timetabl* OR TS=schedul* AND Language=(English)

Databases=SSCI Timespan=1988-2008

AND

TS=schools OR TS=classroom* OR TS=lesson* OR TS=recess OR TS="break times" OR TS=playtimes OR TS="school year" OR TS="school day" OR TS="school term" OR TS=curriculum OR TS=exam OR TS=exams OR TS=examinations AND Language=(English)

Databases=SSCI Timespan=1988-2008

Appendix 2.3: Journals/websites handsearched

British Education Index Update

Number of hits: 0

CERUK (Current Education and children's services Research)

Number of hits: 0

DCSF (Department for Children, Schools and Families)

Number of hits: 0

Free Text

Timing OR Schooling OR Time OR Organisation OR Timetable OR Schedule OR Break OR Play

DCSF keywords

Curriculum OR Evaluation OR Schools

Institute for Public Policy Research (IPPR)

Number of hits: 1

National Foundation for Educational Research (NFER)

Number of hits: 2

Organisation for Economic Co-operation and Development (OECD)

Number of hits: 0

Google Scholar

Search date: 24 June 2008

Number of hits: 3

(pupil OR pupils OR student OR students) + (timing OR length OR timetable OR timetabling OR schedule OR schedules OR scheduling) + (school OR schools OR lesson OR lessons OR recess OR "break time" OR playtime OR year OR term OR exam OR exams OR examination).

Citation searching

Number of hits: 4

Appendix 2.4: Data extraction and review-synthesis coding tool

EPPI-Centre data extraction and coding tool for education studies V2.0

Section A: Administrative details

Use of these guidelines should be cited as: EPPI-Centre (2007) Review Guidelines for Extracting Data and Quality Assessing Primary Studies in Educational Research. Version 2.0 London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

A.1 Name of the reviewer	A.1.1 Details
A.2 Date of the review	A.2.1 Details
A.3 Please enter the details of each paper which reports on this item/ study and which is used to complete this data extraction.	A.3.1 Paper (1) Fill in a separate entry for further papers as required. A.3.2 Unique Identifier: A.3.3 Authors: A.3.4 Title: A.3.5 Paper (2) A.3.6 Unique Identifier: A.3.7 Authors: A.3.8 Title:
A.4 Main paper	A.4.1 Unique Identifier:

A.5 Please enter the details of each paper which reports on this study but is NOT being used to complete this data extraction.	A.5.1 Paper (1) Fill in a separate entry for further papers as required. A.5.2 Unique Identifier: A.5.3 Authors: A.5.4 Title: A.5.5 Paper (2) A.5.6 Unique Identifier: A.5.7 Authors: A.5.8 Title:
A.6 If the study has a broad focus and this data extraction focuses on just one component of the study, please specify this here.	A.6.2 Specific focus of this data extraction (Please specify.)
A.7 Identification of report (or reports) Please use AS MANY KEYWORDS AS APPLY.	A.7.1 Citation A.7.2 Contact A.7.3 Handsearch A.7.4 Unknown A.7.5 Electronic database
A.8 Status Please use ONE keyword only.	A.8.1 Published A.8.2 Published as a report or conference paper A.8.3 Unpublished
A.9 Language (Please specify.)	A.9.1 Details of language of report

Section B: Study aims and rationale

B.1 What are the broad aims of the study?	B.1.1 Explicitly stated (Please specify.) B.1.2 Implicit (Please specify.) B.1.3 Not stated/unclear (Please specify.)
B.2 What is the purpose of the study?	B.2.1 A: Description B.2.2 B: Exploration of relationships B.2.3 C: What works? B.2.4 D: Methods development B.2.5 E: Reviewing/synthesising research
B.3 If the study addresses a 'what works' question, does it focus on outcomes or process?	B.3.1 Not applicable (not a study focusing on 'what works') B.3.2 Outcomes B.3.3 Process
B.4 Why was the study done at that point in time, in those contexts and with those people or institutions?	B.4.1 Explicitly stated (Please specify.) B.4.2 Implicit (Please specify.) B.4.3 Not stated/unclear (Please specify.)
B.5 Was the study informed by, or linked to, an existing body of empirical and/or theoretical research?	B.5.1 Explicitly stated (Please specify.) B.5.2 Implicit (Please specify.) B.5.3 Not stated/unclear (Please specify.)
B.6 Which of the following groups were consulted in working out the aims of the study, or issues to be addressed in the study?	B.6.1 Researchers (Please specify.) B.6.2 Funder (Please specify.) B.6.3 Head teacher/Senior management (Please specify.) B.6.4 Teaching staff (Please specify.) B.6.5 Non-teaching staff (Please specify.) B.6.6 Parents (Please specify.) B.6.7 Pupils/students (Please specify.) B.6.8 Governors (Please specify.) B.6.9 LEA/Government officials (Please specify.) B.6.10 Other education practitioner (Please specify.) B.6.11 Other (Please specify.) B.6.12 None/Not stated B.6.13 Coding is based on: Authors' description B.6.14 Coding is based on: Reviewers' inference
B.7 Do authors report how the study was funded?	B.7.1 Explicitly stated (Please specify.) B.7.2 Implicit (Please specify.) B.7.3 Not stated/unclear (Please specify.)
B.8 When was the study carried out?	B.8.1 Explicitly stated (Please specify.) B.8.2 Implicit (Please specify.) B.8.3 Not stated/unclear (Please specify.)
B.9 What are the study research questions and/or hypotheses?	B.9.1 Explicitly stated (Please specify.) B.9.2 Implicit (Please specify.) B.9.3 Not stated/unclear (Please specify.)

Section C: Study policy or practice focus

C.1 What is/are the topic focus/foci of the study?	C.1.1 Assessment (Please specify.) C.1.2 Classroom management (Please specify.) C.1.3 Curriculum (See next question below.) C.1.4 Equal opportunities (Please specify.) C.1.5 Methodology (Please specify.) C.1.6 Organisation and management (Please specify.) C.1.7 Policy (Please specify.) C.1.8 Teacher careers (Please specify.) C.1.9 Teaching and learning (Please specify.) C.1.10 Other (Please specify.) C.1.11 Coding is based on: Authors' description C.1.12 Coding is based on: Reviewers' inference
C.2 What is the curriculum area, if any?	C.2.1 N/A (not on a specific curriculum area) C.2.2 Art C.2.3 Business Studies C.2.4 Citizenship C.2.5 Cross-curricular C.2.6 Design and Technology C.2.7 Environment C.2.8 General C.2.9 Geography C.2.10 Hidden C.2.11 History C.2.12 ICT C.2.13 Literacy - first languages C.2.14 Literacy - further languages C.2.15 Literature C.2.16 Mathematics C.2.17 Music C.2.18 PSE C.2.19 Physical Education C.2.20 Religious Education C.2.21 Science C.2.22 Vocational C.2.23 Other C.2.24 Coding is based on: Authors' description C.2.25 Coding is based on: Reviewers' inference

C.3 What is/are the educational setting(s) of the study?	C.3.1 Kindergarten C.3.2 Community centre C.3.3 Correctional institution C.3.4 Government department C.3.5 Higher education institution C.3.6 Home C.3.7 Independent school C.3.8 Local education authority C.3.9 Nursery school C.3.10 Other early years setting C.3.11 Post-compulsory education institution C.3.12 Primary school C.3.13 Pupil referral unit C.3.14 Residential school C.3.15 Secondary school C.3.16 Special needs school C.3.17 Workplace C.3.18 Other educational setting C.3.19 Coding is based on: Authors' description C.3.20 Coding is based on: Reviewers' inference
C.4 In which country or countries was the study carried out? Provide further details where relevant (e.g. region or city).	C.4.1 Explicitly stated (Please specify.) C.4.2 Not stated/unclear (Please specify.)

Section D: Actual sample

If there are several samples or levels of sample, please complete for each level

D.1 Who or what is/are the sample in the study?	D.1 Who or what is/are the sample in the study? D.1.1 Learners D.1.2 Senior management D.1.3 Teaching staff D.1.4 Non-teaching staff D.1.5 Other educational practitioners D.1.6 Government D.1.7 Local education authority officers D.1.8 Parents D.1.9 Governors D.1.10 Other sample focus (Please specify.)
D.2 What was the total number of participants in the study (the actual sample)?	D.2.1 Not applicable (e.g. study of policies, documents, etc.) D.2.2 Explicitly stated (Please specify.) D.2.3 Implicit (Please specify.) D.2.4 Not stated/Unclear (Please specify.)

D.3 What is the proportion of those selected for the study who actually participated in the study?	D.3.1 Not applicable (e.g. review) D.3.2 Explicitly stated (Please specify.) D.3.3 Implicit (Please specify.) D.3.4 Not stated/Unclear (Please specify.)
D.4 Which country/countries are the individuals in the actual sample from?	D.4.1 Not applicable (e.g. study of policies, documents, etc.) D.4.2 Explicitly stated (Please specify.) D.4.3 Implicit (Please specify.) D.4.4 Not stated/Unclear (Please specify.)
D.5 If the individuals in the actual sample are involved with an educational institution, what type of institution is it?	D.5.1 Not applicable (e.g. study of policies, documents, etc.) D.5.2 Community centre (Please specify.) D.5.3 Post-compulsory education institution (Please specify.) D.5.4 Government Department (Please specify.) D.5.5 Independent school (Please specify age range and school type.) D.5.6 Nursery school (Please specify.) D.5.7 Other early years setting (Please specify.) D.5.8 Local education authority (Please specify.) D.5.9 Higher education institution (Please specify.) D.5.10 Primary school (Please specify.) D.5.11 Correctional institution (Please specify.) D.5.12 Pupil referral unit (Please specify.) D.5.13 Residential school (Please specify.) D.5.14 Secondary school (Please specify age range.) D.5.15 Special needs school (Please specify.) D.5.16 Workplace (Please specify.) D.5.17 Other educational setting (Please specify.) D.5.18 Coding is based on: Authors' description D.5.19 Coding is based on: Reviewers' inference
D.6 What ages are covered by the actual sample?	D.6.1 Not applicable (e.g. study of policies, documents, etc.) D.6.2 0-4 D.6.3 5-10 D.6.4 11-16 D.6.5 17 to 20 D.6.6 21 and over D.6.7 Not stated/Unclear (Please specify.) D.6.8 Coding is based on: Authors' description D.6.9 Coding is based on: Reviewers' inference

D.7 What is the sex of participants?	D.7.1 Not applicable (e.g. study of policies, documents, etc.) D.7.2 Single sex (Please specify.) D.7.3 Mixed sex (Please specify.) D.7.4 Not stated/Unclear (Please specify.) D.7.5 Coding is based on: Authors' description D.7.6 Coding is based on: Reviewers' inference
D.8 What is the socio-economic status of the individuals within the actual sample? If more than one group is being compared, please describe for each group.	D.8.1 Not applicable (e.g. study of policies, documents, etc.) D.8.2 Explicitly stated (Please specify.) D.8.3 Implicit (Please specify.) D.8.4 Not stated/Unclear (Please specify.)
D.9 What is the ethnicity of the individuals within the actual sample?	D.9.1 Not applicable (e.g. study of policies, documents, etc.) D.9.2 Explicitly stated (please specify) D.9.3 Implicit (Please specify.) D.9.4 Not stated/Unclear (Please specify.)
D.10 What is known about the special educational needs of individuals within the actual sample?	D.10.1 Not applicable (e.g. study of policies, documents, etc.) D.10.2 Explicitly stated (Please specify.) D.10.3 Implicit (Please specify.) D.10.4 Not stated/Unclear (Please specify.)
D.11 Please specify any other useful information about the study participants.	D.11.1 Details

Section E: Programme or intervention description

E.1 If a programme or intervention is being studied, does it have a formal name?	E.1.1 Not applicable (no programme or intervention) E.1.2 Yes (Please specify.) E.1.3 No (Please specify.) E.1.4 Not stated/Unclear (Please specify.)
E.2 Theory of change	E.2.1 Details
E.3 Aim(s) of the intervention	E.3.1 Not stated E.3.2 Not explicitly stated (Write in, as worded by the reviewer.) E.3.3 Stated (Write in, as stated by the authors.)
E.4 Year intervention started Where relevant	E.4.1 Details
E.5 Duration of the intervention	E.5.1 Not stated E.5.2 Not applicable E.5.3 Unclear E.5.4 One day or less (Please specify.) E.5.5 One day to one week (Please specify.) E.5.6 1 week (and 1 day) to 1 month (Please specify.) E.5.7 1 month (and 1 day) to 3 months (Please specify.) E.5.8 3 months (and 1 day) to 6 months (Please specify.) E.5.9 6 months (and 1 day) to 1 year (Please specify.) E.5.10 1 year (and 1 day) to 2 years (Please specify.) E.5.11 2 years (and 1 day) to 3 years (Please specify.) E.5.12 3 years (and 1 day) to 5 years (Please specify.) E.5.13 more than 5 years (Please specify.) E.5.14 Other (Please specify.)
E.6 Person providing the intervention (tick as many as appropriate)	E.6.1 Not stated E.6.2 Unclear E.6.3 Not applicable E.6.4 Counsellor E.6.5 Health professional (Please specify.) E.6.6 Parent E.6.7 Peer E.6.8 Psychologist E.6.9 Researcher E.6.10 Social worker E.6.11 Teacher/lecturer E.6.12 School E.6.13 Other (Please specify.)

E.7 Number of people recruited to provide the intervention (and comparison condition) (e.g. teachers or health professionals)	E.7.1 Not stated E.7.2 Unclear E.7.3 Reported (Please include the number for the providers involved in the intervention and comparison groups, as appropriate.)
E.8 How were the people providing the intervention recruited? (Please write in.) Also, give information on the providers involved in the comparison group(s), as appropriate.	E.8.1 Not stated E.8.2 Stated (Please write in.)
E.9 Was special training given to people providing the intervention?	E.9.1 Not stated E.9.2 Unclear E.9.3 Yes (Please specify.) E.9.4 No

Section F: Results and conclusions

In future, this section is likely to incorporate material from EPPI-Reviewer to facilitate reporting numerical results.

F.1 How are the results of the study presented?	F.1.1 Details
F.2 What are the results of the study as reported by the authors?	F.2.1 Details
F.3 What do the author(s) conclude about the findings of the study?	F.3.1 Details

Section G: Study method

G.1 Study timing	G.1.1 Cross-sectional G.1.2 Retrospective G.1.3 Prospective G.1.4 Not stated/Unclear (Please specify.)
G.2 When were the measurements of the variable(s) used as outcome measures made, in relation to the intervention	G.2.1 Not applicable (not an evaluation) G.2.2 Before and after G.2.3 Only after G.2.4 Other (Please specify.) G.2.5 Not stated/Unclear (Please specify.)

G.3 What is the method used in the study?	G.3.1 Random experiment with random allocation to groups G.3.2 Experiment with non-random allocation to groups G.3.3 One group pre-post test G.3.4 One group post-test only G.3.5 Cohort study G.3.6 Case-control study G.3.7 Cross-sectional study G.3.8 Views study G.3.9 Ethnography G.3.10 Systematic review G.3.11 Other review (non-systematic) G.3.12 Case study G.3.13 Document study G.3.14 Action research G.3.15 Methodological study G.3.16 Secondary data analysis

Section H: Methods - groups

H.1 If comparisons are being made between two or more groups*, please specify the basis of any divisions made for making these comparisons.	H.1.1 Not applicable (not more than one group) H.1.2 Prospective allocation into more than one group (e.g. allocation to different interventions, or allocation to intervention and control groups) H.1.3 No prospective allocation but use of pre-existing differences to create comparison groups (e.g. receiving different interventions or characterised by different levels of a variable such as social class) H.1.4 Other (Please specify.) H.1.5 Not stated/Unclear (Please specify.)
H.2 How do the groups differ?	H.2.1 Not applicable (not in more than one group) H.2.2 Explicitly stated (Please specify.) H.2.3 Implicit (Please specify.) H.2.4 Not stated/unclear (Please specify.)
H.3 Number of groups	H.3.1 Not applicable (not more than one group) H.3.2 One H.3.3 Two H.3.4 Three H.3.5 Four or more (Please specify.) H.3.6 Other/unclear (Please specify.)

H.4 If prospective allocation into more than one group, what was the unit of allocation?	H.4.1 Not applicable (not more than one group) H.4.2 Not applicable (no prospective allocation) H.4.3 Individuals H.4.4 Groupings or clusters of individuals (e.g. classes or schools) (Please specify.) H.4.5 Other (e.g. individuals or groups acting as their own controls) (Please specify.) H.4.6 Not stated/unclear (Please specify.)
H.5 If prospective allocation into more than one group, which method was used to generate the allocation sequence?	H.5.1 Not applicable (not more than one group) H.5.2 Not applicable (no prospective allocation) H.5.3 Random H.5.4 Quasi-random H.5.5 Non-random H.5.6 Not stated/Unclear (Please specify.)
H.6 If prospective allocation into more than one group, was the allocation sequence concealed?	H.6.1 Not applicable (not more than one group) H.6.2 Not applicable (no prospective allocation) H.6.3 Yes (Please specify.) H.6.4 No (Please specify.) H.6.5 Not stated/Unclear (Please specify.)
H.7 Study design summary	H.7.1 Details

Section I: Methods - sampling strategy

I.1 Are the authors trying to produce findings that are representative of a given population?	I.1.1 Explicitly stated (Please specify.) I.1.2 Implicit (Please specify.) I.1.3 Not stated/Unclear (Please specify.)
I.2 What is the sampling frame (if any) from which the participants are chosen?	I.2.1 Not applicable (Please specify.) I.2.2 Explicitly stated (Please specify.) I.2.3 Implicit (Please specify.) I.2.4 Not stated/Unclear (Please specify.)
I.3 Which method does the study use to select people, or groups of people (from the sampling frame)?	I.3.1 Not applicable (no sampling frame) I.3.2 Explicitly stated (Please specify.) I.3.3 Implicit (Please specify.) I.3.4 Not stated/unclear (Please specify.)

I.4 Planned sample size	I.4.1 Not applicable (Please specify.) I.4.2 Explicitly stated (Please specify.) I.4.3 Not stated/unclear (Please specify.)
I.5 How representative was the achieved sample (as recruited at the start of the study) in relation to the aims of the sampling frame?	I.5.1 Not applicable (e.g. study of policies, documents, etc.) I.5.2 Not applicable (no sampling frame) I.5.3 High (Please specify.) I.5.4 Medium (Please specify.) I.5.5 Low (Please specify.) I.5.6 Unclear (Please specify.)
I.6 If the study involves studying samples prospectively over time, what proportion of the sample dropped out over the course of the study?	I.6.1 Not applicable (e.g. study of policies, documents, etc.) I.6.2 Not applicable (not following samples prospectively over time) I.6.3 Explicitly stated (Please specify.) I.6.4 Implicit (Please specify.) I.6.5 Not stated/Unclear (Please specify.)
I.7 For studies that involve following samples prospectively over time, do the authors provide any information on whether, and/or how, those who dropped out of the study differ from those who remained in the study?	I.7.1 Not applicable (e.g. study of policies, documents, etc.) I.7.2 Not applicable (not following samples prospectively over time) I.7.3 Not applicable (no drop-outs) I.7.4 Yes (Please specify.) I.7.5 No
I.8 If the study involves following samples prospectively over time, do authors provide baseline values of key variables, such as those being used as outcomes, and relevant socio-demographic variables?	I.8.1 Not applicable (e.g. study of policies, documents, etc.) I.8.2 Not applicable (not following samples prospectively over time) I.8.3 Yes (Please specify.) I.8.4 No

Section J: Methods - recruitment and consent

J.1 Which methods are used to recruit people into the study?	J.1.1 Not applicable (Please specify.) J.1.2 Explicitly stated (Please specify.) J.1.3 Implicit (Please specify.) J.1.4 Not stated/Unclear (Please specify.) J.1.5 Please specify any other details relevant to recruitment and consent.
J.2 Were any incentives provided to recruit people into the study?	J.2.1 Not applicable (Please specify.) J.2.2 Explicitly stated (Please specify.) J.2.3 Not stated/Unclear (Please specify.)

J.3 Was consent sought? Please comment on the quality of consent, if relevant.	J.3.1 Not applicable (Please specify.) J.3.2 Participant consent sought J.3.3 Parental consent sought J.3.4 Other consent sought J.3.5 Consent not sought J.3.6 Not stated/Unclear (Please specify.)

Section K: Methods - data collection

K.1 Which variables or concepts, if any, does the study aim to measure or examine?	K.1.1 Explicitly stated (Please specify.) K.1.2 Implicit (Please specify.) K.1.3 Not stated/Unclear
K.2 Please describe the main types of data collected and specify if they were used (a) to define the sample, or (b) to measure aspects of the sample as findings of the study.	K.2.1 Details
K.3 Which methods were used to collect the data?	K.3.1 Curriculum-based assessment K.3.2 Focus group interview K.3.3 One-to-one interview (face to face or by phone) K.3.4 Observation K.3.5 Self-completion questionnaire K.3.6 self-completion report or diary K.3.7 Examinations K.3.8 Clinical test K.3.9 Practical test K.3.10 Psychological test (e.g. I.Q test) K.3.11 Hypothetical scenario including vignettes K.3.12 School/college records (e.g. attendance records, etc.) K.3.13 Secondary data such as publicly available statistics K.3.14 Other documentation K.3.15 Not stated/Unclear (Please specify.) K.3.16 Please specify any other important features of data collection. K.3.17 Coding is based on: Author's description K.3.18 Coding is based on: Reviewers' interpretation
K.4 Details of data-collection instruments or tool(s).	K.4.1 Explicitly stated (Please specify.) K.4.2 Implicit (Please specify.) K.4.3 Not stated/Unclear (Please specify.)

K.5 Who collected the data?	K.5.1 Researcher K.5.2 Head teacher/Senior management K.5.3 Teaching or other staff K.5.4 Parents K.5.5 Pupils/students K.5.6 Governors K.5.7 LEA/Government officials K.5.8 Other educational practitioner K.5.9 Other (Please specify.) K.5.10 Not stated/Unclear K.5.11 Coding is based on: Author's description K.5.12 Coding is based on: Reviewers' inference
K.6 Do the authors' describe any ways they addressed the repeatability or reliability of their data-collection tools/methods?	K.6.1 Details
K.7 Do the authors describe any ways they have addressed the validity or trustworthiness of their data-collection tools/methods?	K.7.1 Details
K.8 Was there a concealment of which group that subjects were assigned to (i.e. the intervention or control) or other key factors from those carrying out measurement of outcome, if relevant?	K.8.1 Not applicable (please say why) K.8.2 Yes (please specify) K.8.3 No (please specify)
K.9 Where was the data collected?	K.9.1 Educational Institution (Please specify.) K.9.2 Home (Please specify.) K.9.3 Explicitly stated (Please write in as worded by the author.) K.9.4 Not stated/Unclear (Please specify.)

Section L: Methods - data analysis

L.1 What rationale do the authors give for the methods of analysis for the study?	L.1.1 Details
L.2 Which methods were used to analyse the data?	L.2.1 Explicitly stated (Please specify.) L.2.2 Implicit (Please specify.) L.2.3 Not stated/Unclear (Please specify.) L.2.4 Please specify any important analytic or statistical issues.
L.3 Which statistical methods, if any, were used in the analysis?	L.3.1 Details
L.4 Did the study address multiplicity by reporting ancillary analyses, including sub-group analyses and adjusted analyses, and do the authors report on whether these were pre-specified or exploratory?	L.4.1 Yes (Please specify.) L.4.2 No (Please specify.) L.4.3 Not applicable

L.5 Do the authors describe strategies used in the analysis to control for bias from confounding variables?	L.5.1 Yes (Please specify.) L.5.2 No L.5.3 Not applicable
L.6 For evaluation studies that use prospective allocation, please specify the basis on which data analysis was carried out.	L.6.1 Not applicable (not an evaluation study with prospective allocation) L.6.2 'Intention to intervene' L.6.3 'Intervention received' L.6.4 Not stated/Unclear (Please specify.)
L.7 Do the authors describe any ways they have addressed the repeatability or reliability of data analysis?	L.7.1 Details
L.8 Do the authors describe any ways that they have addressed the validity or trustworthiness of data analysis?	L.8.1 Details
L.9 If the study uses qualitative methods, how well has diversity of perspective and content been explored?	L.9.1 Details
L.10 If the study uses qualitative methods, how well has the detail, depth and complexity (i.e. the richness) of the data been conveyed?	L.10.1 Details
L.11 If the study uses qualitative methods, has analysis been conducted such that context is preserved?	L.11.1 Details

Section M: Quality of study - reporting

M.1 Is the context of the study adequately described?	M.1.1 Yes (Please specify.) M.1.2 No (Please specify.)
M.2 Are the aims of the study clearly reported?	M.2.1 Yes (Please specify.) M.2.2 No (Please specify.)
M.3 Is there an adequate description of the sample used in the study and how the sample was identified and recruited?	M.3.1 Yes (Please specify.) M.3.2 No (Please specify.)
M.4 Is there an adequate description of the methods used in the study to collect data?	M.4.1 Yes (Please specify.) M.4.2 No (Please specify.)
M.5 Is there an adequate description of the methods of data analysis?	M.5.1 Yes (Please specify.) M.5.2 No (Please specify.)
M.6 Is the study replicable from this report?	M.6.1 Yes (Please specify.) M.6.2 No (Please specify.)
M.7 Do the authors state where the full, original data are stored?	M.7.1 Yes (Please specify.) M.7.2 No (Please specify.)

M.8 Do the authors avoid selective reporting bias (e.g. do they report on all variables they aimed to study, as specified in their aims/research questions)?	M.8.1 Yes (Please specify.) M.8.2 No (Please specify.)

Section N: Quality of the study - weight of evidence

N.1 Are there ethical concerns about the way the study was done?	N.1.1 Yes, some concerns (Please specify.) N.1.2 No (Please specify.)
N.2 Were users/relatives of users appropriately involved in the design or conduct of the study?	N.2.1 Yes, a lot (Please specify.) N.2.2 Yes, a little (Please specify.) N.2.3 No (Please specify.)
N.3 Is there sufficient justification for why the study was done the way it was?	N.3.1 Yes (Please specify.) N.3.2 No (Please specify.)
N.4 Was the choice of research design appropriate for addressing the research question(s) posed?	N.4.1 Yes, completely (Please specify.) N.4.2 No (Please specify.)
N.5 Have sufficient attempts been made to establish the repeatability or reliability of data-collection methods or tools?	N.5.1 Yes, good (Please specify.) N.5.2 Yes, some attempt (Please specify.) N.5.3 No, none (Please specify.)
N.6 Have sufficient attempts been made to establish the validity or trustworthiness of data-collection tools and methods?	N.6.1 Yes, good (Please specify.) N.6.2 Yes, some attempt (Please specify.) N.6.3 No, none (Please specify.)
N.7 Have sufficient attempts been made to establish the repeatability or reliability of data analysis?	N.7.1 Yes (Please specify.) N.7.2 No (Please specify.)
N.8 Have sufficient attempts been made to establish the validity or trustworthiness of data analysis?	N.8.1 Yes, good (Please specify.) N.8.2 Yes, some attempt (Please specify.) N.8.3 No, none (Please specify.)
N.9 To what extent are the research design and methods employed able to rule out any other sources of error/bias which would lead to alternate explanations for the findings of the study?	N.9.1 A lot (Please specify.) N.9.2 A little (Please specify.) N.9.3 Not at all (please specify)
N.10 How generalisable are the study results?	N.10.1 Details
N.11 In light of the above, do the reviewers differ from the authors over the findings or conclusions of the study?	N.11.1 Not applicable (no difference in conclusions) N.11.2 Yes (Please specify.)
N.12 Have sufficient attempts been made to justify the conclusions drawn from the findings, so that the conclusions are trustworthy?	N.12.1 Not applicable (results and conclusions inseparable) N.12.2 High trustworthiness N.12.3 Medium trustworthiness N.12.4 Low trustworthiness

N.13 Weight of evidence A: Taking account of all quality assessment issues, can the study findings be trusted in answering the study question(s)?	N.13.1 High trustworthiness N.13.2 Medium trustworthiness N.13.3 Low trustworthiness
N.14 Weight of evidence B (Appropriateness of research design and analysis for addressing the question, or sub-questions, of this specific systematic review)	N.14.1 High N.14.2 Medium N.14.3 Low
N.15 Weight of evidence C (Relevance of particular focus of the study (including conceptual focus, context, sample and measures) for addressing the question, or sub-questions, of this specific systematic review)	N.15.1 High N.15.2 Medium N.15.3 Low
N.16 Weight of evidence D (Overall weight of evidence)	N.16.1 High N.16.2 Medium N.16.3 Low

Map-specific keywords

A.1 Use of time	A.1.1 Length of the school day (e.g. half day, full day, 6 hours or more) A.1.2 Length of lessons (e.g. block scheduling, time allocated to specific subjects) A.1.3 Organisation of break times A.1.4 Length of terms/semesters A.1.5 Multiple-shift schooling (e.g. one cohort of students attends in the morning and a different cohort in the afternoon) A.1.6 Length of school year e.g. year round schooling A.1.7 Organisation of teaching specific subjects A.1.8 Time of day subjects are taught (e.g. timing of lessons, such as whether taking place am or pm) A.1.9 Timing of exams A.1.10 Other A.1.11 Pupils' use of time A.1.12 Pupils' organising use of time in school day

A.2 Type of pupil outcomes reported	A.2.1 Literacy (e.g. reading scores, writing, etc.) A.2.2 Numeracy (e.g. Mathematics scores, numerical ability A.2.3 Student grades (e.g. grade point average) A.2.4 exam results (e.g. GCCE, high school diploma, graduation rates) A.2.5 Pupil attendance rates A.2.6 Pupil drop-out rates A.2.7 Pupil self-esteem A.2.8 Pupil relationships outcomes (Please specify.) A.2.9 Pupil psycho-social well-being A.2.10 Pupil motivation A.2.11 Pupil attention rates A.2.12 Pupil memory rates A.2.13 Other
A.3 Type of school outcomes reported	A.3.1 School climate Organisation, ethos, well-being, cohesion, status A.3.2 Leadership A.3.3 Teaching Curriculum, teaching methods, assessment A.3.4 Costs A.3.5 Other
A.4 Type of pupil views reported	A.4.1 Satisfaction with intervention A.4.2 Perception of intervention A.4.3 Perception of impact / effectiveness of intervention A.4.4 Pupils' perception of the implementation of an intervention A.4.5 Other
A.5 Type of teacher/school staff data reported	A.5.1 Teachers use of time A.5.2 Teacher/school staff perception of the implementation of an intervention A.5.3 Other

Home Office quality assessment tool – adapted for use in current review

Section A: Home Office QAT 1 Sample

A.1 Sample size and power	B.1.1 Sample is sufficient to detect the estimated effect size at 80% power (1) Please state the effect size estimate. B.1.2 Sample is not sufficient to detect the estimated effect size at 80% power (3) Please state the effect size estimate. B.1.3 Not reported (5)
A.2 Method of study	B.2.1 Whole population or random samples (1) B.2.2 Purposive samples with potential impact adequately controlled for statistically (2) B.2.3 Purposive samples with potential impact not adequately controlled for statistically, or not controlled for at all (3) B.2.4 Not reported (5)
A.3 Method of selection	B.3.1 Control and experimental groups comparable (1) B.3.2 Control and experimental groups not comparable, but differences adequately controlled for statistically (2) B.3.3 Control and experimental groups not comparable, and differences not adequately controlled for statistically, or not controlled at all (3) B.3.4 Not reported (5)

Section B: Home Office QAT 2 Bias

B.1 Response/refusal bias	C.1.1 No bias (1) C.1.2 Some bias but adequately controlled for statistically (2) C.1.3 Some bias and not adequately controlled for statistically, or not controlled for at all (3) C.1.4 Not reported (5)
B.2 Attrition bias	C.2.1 No/very little (<10%) attrition (1) C.2.2 Some attrition but adequately controlled for statistically (2) C.2.3 Some attrition but not adequately controlled for statistically, or not controlled for at all (3) C.2.4 Not reported (5)
B.3 Performance bias	C.3.1 Groups treated equally and observers blinded (or not relevant) 1 Blinding is not relevant where outcome is based on official statistics (e.g. police records) C.3.2 Differences in way group treated and/or no blinding - minor effect 2 C.3.3 Differences in way groups treated and/or no blinding - major effects (3) C.3.4 Not reported (5)

Section C: Home Office QAT 3 Data collection

C.1 Data-collection method	D.1.1 Very appropriate (1) D.1.2 Appropriate (2) D.1.3 Not appropriate (3) D.1.4 Not reported (5)
C.2 Outcome measurement timing	D.2.1 Very appropriate (1) D.2.2 Appropriate (2) D.2.3 Not appropriate (3) D.2.4 Not reported (5)
C.3 Validation of outcome measures	D.3.1 Very appropriate (1) D.3.2 Appropriate (2) D.3.3 Not appropriate (3) D.3.4 Not reported (5)

Section D: Home Office QAT 4. Data analysis

D.1 Appropriate data analysis techniques / reporting	E.1.1 Very appropriate (1) E.1.2 Appropriate (2) E.1.3 Not appropriate (3) E.1.4 Not reported (5)

Appendix 2.5: Weight of evidence details

Weight of evidence A: Quality of the execution (internal methodological coherence), based upon the study only

Calculated from the Home Office QAT questions as follows:

Scores for questions (((A1+A2*+A3)/3) + (B4+B5+B6)/3) = (C7+C8+C9)/3)) + D10

1. High = total score of 6 or less

2. Medium = total score of 7 or 8

3. Low = total score of 9 or more

*Question A2 was given a score of 3 for all studies. This was decided because, unlike A.1 and A.3, it addresses external rather than internal validity.

Weight of evidence B: Appropriateness of research design and analysis for addressing the question, or sub-questions, of this specific systematic review

1. High = Randomised controlled trials

2. Medium = Matched sample or sample variation statistically controlled for

3. Low = None of the above

Weight of evidence C: Relevance of particular focus of the study (including conceptual focus, context, sample and measures) for addressing the question, or sub-questions, of this specific systematic review

1. High = Large number of schools and high number of students (over 1,000)

2. Medium = high number of students (over 1,000)

3. Low = small number of schools (1-3) and low number of students (below 1,000)

Weight of evidence D: Overall weight of evidence

For WoE A, B and C the following scoring was given for each judgement:

1 = High, 2 = Medium, 3 = Low, 3-4 = High, 5-7 = Medium, 8-9 = Low

Appendix 2.6: Effectiveness interpretation framework

Strong evidence of effectiveness (positive or negative)

This category will have the following:

At least two or more studies that score high on weight of evidence D with an individual or combined sample size of >500.

Where the result (weighted mean) shows

Positive

a positive effect size (favouring the intervention) of at least +0.20 and where the lower 95% confidence interval does not cross the 'line of no effect'

or a Negative

effect size (favouring the control) of at least -0.20 and where the upper confidence interval does not cross the 'line of no effect'

Evidence of effectiveness (positive or negative)

This category will have at least the following:

Two or more studies that score medium or high on weight of evidence D with an individual or collective sample size of >500 students

Two or more studies

Positive

where the effect size (weighted mean or single effect size, is smaller than +0.20)shows a positive (favouring the intervention) and the lower (respectively) 95% confidence interval does not cross the 'line of no effect'

Or Negative

Where effect size (favouring the control) is smaller than -0.20 and the upper (respectively) 95% confidence interval does not cross the 'line of no effect'

Limited evidence: (positive or negative)

This category will have

at least one study that scores medium or high on weight of evidence D with an individual or combined sample size of >500 students; and

the weighted mean crosses the 'line of no effect' (95% confidence interval)

Limited evidence: weak evidence

This category will have

no studies that score medium or high on weight of evidence D; or

at least one study that does score medium or high on weight of evidence D but combined set of studies that have a sample size of <500 students

Insufficient evidence

The weighted mean average measures I2<70%.

Adapted from:

http://www.bestevidence.org

Appendix 3.1: Details of studies included in the in-depth review

Study	Study aim	Evidence statements
Cobb et al. (1999)	The study evaluated the effects of a 4 x 4 block scheduling programme in a middle school.	In a US evaluation of one middle school, using a post-test matched pairs design, Cobb et al. found that block scheduled students performed significantly less on a standardised Mathematics test compared with students taking part in a traditional schedule. The authors did find a statistically significant interaction between block scheduling on gender and age, with block scheduling having a more positive semester GPA effect on male students compared with female students, and for the tenth and eleventh graders compared with eighth and ninth graders. However, these interactions did not hold for the cumulative GPA.
DiRocco (1997)	The study examined the effects of alternate day block scheduling on student achievement, attendance and discipline.	In a US thesis of seventh and eighth graders from one middle school, Dirrocco evaluated the effects of alternate day block scheduling. He found that both eventh grade and eighth grade final course averages in Social Studies, Reading and final GPA were higher, and that eighth grade final course averages were higher in Mathematics and Science as a result of alternate block scheduling.
Hughes (2004)	The study examined the impact of a change of schedule on overall academic performance as measured by a student's GPA.	In a US case study of a middle school, Hughes found that changing from a traditional schedule to a block schedule had a positive effect on student performance, with the class of 1995 GPA average being less than the class of 1999. However, the author concludes that there is an issue with the validity of the study because of the adjustment made to the grading scale due to an absence of data from the original student population. This meant that the maximum GPA for the class of 1999 has been deflated and is much less than that of the class of 1995 and is more in line with the mean GPA of 1995.

Study	Study aim	Evidence statements
Lapkin et al. (1997)	The study investigated the effects of two alternate, compact models of French programme delivery compared with the current model.	In a Canadian study, Lapkin et al. use quasi-randomisation methods to allocate junior high school students to two intervention groups (half-day and 80-minute classes) and a comparison group (40-minute classes). They found there were no significant differences between the groups in French listening comprehension and speaking, but there were significant differences in reading and writing with students in the block schedules performing better. This improvement was maintained in the following school year when students reached eighth grade.
Lewis et al. (2003)	The study investigated the effects of full and alternate day block scheduling on Language Arts and Science achievement in a junior high school.	In a US study of 102 junior high schools students, Lewis et al. look at the impact of both 4 x 4 block scheduling and alternate block scheduling. They found that students in both forms of block scheduling outperformed students in traditional scheduling, and that A/B block scheduling has the largest positive impact on low-achieving students.
Marchette (2003)	The study assessed the impacts of different types of scheduling on student testing performance in Biology.	In a US evaluation of 147 Mississippi high schools, Marchette et al. looked at the impact of different types of school schedules on student achievement in Biology. They found that students in the alternate block schedule had the highest aggregate mean score, the 4 x 4 block schedule had the second highest aggregate mean score, and those in a traditional schedule had the lowest aggregate mean score; however, when analysed using an ANCOVA, there were no statistically significant differences in the mean scores between the different types of school schedules.
McCreary and Hausman (2001)	The study assessed the differences in student outcomes between students in high schools using block, semester or trimester schedules in one urban school system over a four-year period. The study specifically tested for the differences in student annual GPA scores on SATs, credits attempted and absentee rates.	In a US study of three high schools from one urban district, each implementing a different type of school schedule McCreary and Hausman examined student annual GPAs, scores on the Stanford Achievement Test 9 and found that students in a traditional semester schedule had higher GPAs than those in the alternate block schedules or trimester schedules. Students also attempt more credits under block and trimester schedules, but they are not enrolling in more core courses.
Rice et al. (2002)	The study drew on data from the National Education Longitudinal Study (1988) to estimate the impact of block-scheduled Mathematics courses on tenth-grade student achievement and teachers' use of class time.	In a US evaluation, Rice et al. draw on data from the National Education Longitudinal Study to estimate the impact of block-scheduled Mathematics courses on tenth-grade student achievement. They found no difference in students' eighth-grade or tenth-grade Mathematics achievement scores for students in the two types of block scheduled classes. When the authors controlled for other factors, they found that being enrolled in a block-scheduled tenth-grade Mathematics class has a significant, but negative, impact on student achievement scores.
Schreiber et al. (2001)	The study assessed the effect of type of schedule (block, hybrid or traditional) on student achievement, as measured by a standardised test.	In a US evaluation of block scheduling, Schreiber et al. collected data from two independent tenth-grade cohorts from one high school. Overall, they found that block scheduling did not improve achievement in Reading and Language; it only improved mathematic computation for students who had experienced block scheduling for two years (cohort 1) rather than one (cohort 2). Block scheduling did not interact with gender or higher or lower achieving students for either cohorts.

Study	Study aim	Evidence statements
Schroth and Dixon (1995)	To measured the effects of block scheduling on 7th grade mathematics students	In a US study of two Texan middles schools, Scroth et al. analysed students' scores using the Texas Assessment of Academic Skills test in 1994 (pre-block scheduling), and again in 1995, one year after block scheduling. They found that, when lower achieving students attend Mathematic classes for a longer period of time (90 minutes), test scores did not increase more than scores of students in the traditional daily 50 minute classes. In addition, higher achieving students showed virtually no improvement from being enrolled in block scheduled classes.
Texas Education Agency (1999)	The study investigated the effects of high school scheduling practices on student performance	An evaluation by the Texas Education Agency of high school students from 1,070 schools found that high schools using modified 'alternative block' schedules reported the highest average percentage (74.5%) of tenth-grade students who passed all academic tests in Reading, Mathematics and Writing. The smallest average percentage (64.3%) was observed among schools using accelerated block schedules. However, when the authors controlled for various contextual factors, they concluded that school schedules did not explain or account for the variation in overall academic performance.
Veal (1999)	The aim is to assess the effectiveness of a hybrid schedule which consists of three traditional and two block classes each day.	In a US study of one junior high school, the authors found that students enrolled in hybrid block schedules had high GPAs compared with students enrolled in traditionally scheduled lessons. They also found that females performed significantly better than males in pre and post block scheduling.
Walker (2000)	To compare the academic performance of block scheduled with traditionally scheduled schools as measured by scores on the Kansas State assessment of Mathematics.	In a US thesis of 345 high schools and approximately 150,000 students, Walker found that all schools achieved significantly higher Power scores on the Kansas State Mathematics Assessment over the period 1994-1999. There was no significant measurable difference in the impact of block scheduling or traditional scheduling on the Mathematics assessment test, but socio-economic status was found as the best predictor of achievement in Mathematics.
Zhang (2001)	to examine academic performance differences between students in block and traditionally scheduled high schools	In a US study or 214 high schools and approximately 640,000 students Zhang reported that, in Algebra, there was a significant difference between students in the 4 x 4 and traditionally scheduled schools, with students in the 4 x 4 schools significantly outperforming students in the traditional schools across four years from 1997-2000. When testing in English, Biology and US History, there were no significant differences between students in the 4 x 4 and traditionally scheduled schools.

Appendix 3.2 Summary of effects

Table A4.1: 4 x 4 block scheduling

Study measure		4 x 4 Block scheduling			Control			
		Sample size	Mean	SD	Sample size	Mean	SD	g
Cobb et al. (1999)	Cumulative GPA	150	2.980	0.760	150	2.800	0.740	0.239
	Standardised Mathematics	236	61.630	26.180	236	65.360	25.240	-0.145
Hughes (2004)	Grade Point Average	513	2.593	0.742	424	2.453	0.796	0.182
Lewis et al. (2003)	Language Arts: M, L/A	8	222.500	4.440	8	212.000	7.870	1.554
	Language Arts: F, L/A	8	218.500	8.180	8	220.130	7.490	-0.197
	Language Arts: M, H/A	9	230.780	7.310	9	232.560	7.330	-0.232
	Language Arts: F, H/A	12	227.420	5.760	12	231.170	5.690	-0.632
	Science: M, L/A	19	218.740	15.110	24	210.830	9.300	0.637
	Science: F, L/A	22	218.590	10.810	20	207.850	6.170	1.182
	Science: M, H/A	30	223.300	11.500	26	222.540	8.360	0.074
	Science: F, H/A	17	220.940	9.590	28	219.680	8.410	0.140
	Language Arts: Combined	37			37			-0.056
	Science: Combined	88			98			0.445
Marchette (2003)	Biology Tests	46	334.270	17.170	75	326.260	16.440	0.476
Schreiber et al. (2001)	4 x 4 ISTEP Mathematics	49	70.080	21.280	142	70.590	17.450	-0.027
	4 x 4 ISTEP English	49	66.920	17.440	142	69.440	17.270	-0.145
Walker (2000)	Mathematics power scores							
Zhang (2001)	Algebra 1	436,686	48.200	10.000	203,364	47.200	10.000	0.100
	Biology	436,686	49.200	10.000	203,364	49.600	10.000	0.030
	English	436,686	49.800	10.000	203,364	49.600	10.000	0.020
	Economic, Legal and Political Systems	436,686	50.300	10.000	203,364	50.000	10.000	0.030
	US History	436,686	49.700	10.000	203,364	49.800	10.000	-0.010

Table A4.2: A/B Block scheduling

Study measure		A/B Block scheduling			Control			
		Sample size	Mean	SD	Sample size	Mean	SD	g
DiRocco (1997)	GPA	78	86.910	7.640	82	84.190	8.110	0.343
	Reading	78	67.700	27.300	82	63.700	26.300	0.149
	Language	78	62.900	28.400	82	56.500	31.800	0.211
	Mathematics	77	61.400	27.900	82	56.400	28.800	0.175
	Science	75	74.800	26.300	81	70.900	25.700	0.149
	Social Science	77	72.900	26.400	81	60.200	30.200	0.445
Lewis et al. (2003)	Language Arts: M, L/A	8	225.380	4.140	8	212.000	7.870	2.012
	Language Arts: F, L/A	8	226.500	7.090	8	220.130	7.490	0.826
	Language Arts: M, H/A	9	230.000	7.260	9	232.560	7.330	-0.334
	Language Arts: F, H/A	12	230.830	7.710	12	231.170	5.690	-0.048
	Science: M, L/A	17	214.760	13.950	24	210.830	9.300	0.337
	AB Science: F, L/A	21	217.290	6.070	20	207.850	6.170	1.513
	Science: M, H/A	15	221.330	8.150	26	222.540	8.360	-0.143
	Science: F, H/A	30	222.000	9.620	28	219.680	8.410	0.253
	Language Arts: Combined	37			37			0.367
	Science: Combined	83			98			0.419
Marchette (2003)	Biology test	25	338.300	59.110	75	326.260	16.440	0.367
McCreary and Hausman (2001)	Mathematics SAT	2,400	714.300	35.600	2,500	717.800	35.600	-0.098
	Science SAT	2,400	694.200	36.900	2,500	689.400	36.900	0.130
Texas Educational Agency	All TAAS tests results	100			100			0.114
	Reading	100			100			-0.126
	Mathematics	100			100			0.096
	Writing	100			100			0.135

Table A4.3: Single block lessons

Study		Single block lesson			Control			
		Sample size	Mean	SD	Sample size	Mean	SD	g
Lapkin et al. (1997)	Listening	25	9.00	3.100	19	7.900	3.700	0.321
80-minute lessons	Reading	21	11.100	4.600	22	6.900	4.500	0.321
	Dictation	26	36.00	17.100	23	38.300	13.900	-0.144
	Composition	26	4.500	2.400	23	3.500	2.600	0.394
	Oral test	8	22.200	8.900	9	19.800	6.000	0.304
	Role play	8	4.100	2.600	9	4.900	2.100	-0.324
Rice et al. (2002)	Mathematics	46	-	35.600	4,523		35.600	-0.246
>70-minute lessons								

The results of this systematic review are available in four formats:

SUMMARY	Explains the purpose of the review and the main messages from the research evidence
REPORT	Describes the background and the findings of the review(s) but without full technical details of the methods used
TECHNICAL REPORT	**Includes the background, main findings, and full technical details of the review**
DATABASES	Access to codings describing each research study included in the review

These can be downloaded or accessed at
http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2473

First produced in 2010 by:

Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)
Social Science Research Unit
Institute of Education, University of London
18 Woburn Square
London WC1H 0NR

Tel: +44 (0)20 7612 6367

http://eppi.ioe.ac.uk/
http://www.ioe.ac.uk/ssru/

ISBN: 978-1-907345-02-9

The **Evidence for Policy and Practice Information and Co-ordinating Centre** (EPPI-Centre) is part of the Social Science Research Unit (SSRU), Institute of Education, University of London.

The EPPI-Centre was established in 1993 to address the need for a systematic approach to the organisation and review of evidence-based work on social interventions. The work and publications of the Centre engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Founded in 1990, the Social Science Research Unit (SSRU) is based at the Institute of Education, University of London. Our mission is to engage in and otherwise promote rigorous, ethical and participative social research as well as to support evidence-informed public policy and practice across a range of domains including education, health and welfare, guided by a concern for human rights, social justice and the development of human potential.

This document is available in a range of accessible formats including large print. Please contact the Institute of Education for assistance:

telephone: +44 (0)20 7947 9556 email: info@ioe.ac.uk