EPPI CENTRE

IOE LONDON

Leading education and social research
Institute of Education
University of London

Accountability and children's outcomes in high-performing education systems

Analytical maps of approaches to measuring children's education, health and well-being outcomes in high-performing educational systems

Report written by Chris Husbands, Ann Shreeve and Natalia R. Jones

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

Report no. 1617 · December 2008

SSRU
Social Science Research Unit

REPORT

Accountability and children's outcomes in high-performing education systems: Analytical maps of approaches to measuring children's education, health and well-being outcomes in high-performing educational systems

REPORT

Report by Chris Husbands (Institute of Education,University of London)
Ann Shreeve (Institute of Education,University of London)
Natalia R.Jones (Institute of Education,University of London)

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 1617R.

This report should be cited as:

Husbands C, Shreeve A, Jones NR (2008) Accountability and children's outcomes in high-performing education systems: analytical maps of approaches to measuring children's education, health and well-being outcomes in high-performing education systems. In: *Research Evidence in Education Library.* London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

CONTENTS

List of abbreviations

DCSF	Department for Children, Schools and Families
COOL	Educational Careers Cohort Survey
ECM	Every Child Matters
EPPI-Centre	Evidence for Policy and Practice Information and Co-ordinating Centre
Eurydice	The information network on education in Europe
IEA	International Association for the Evaluation of Educational Achievement
INCA	International Review of Curriculum and Assessment Frameworks Internet Archive
OECD	Organisation for Economic Co-operation and Development
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
TIMMS	Trends in International Mathematics and Science Study
UNICEF	United Nations Children's Fund

Abstract

What do we want to know?

The research team wished to establish what approaches to measuring children's education, health and well-being outcomes existed in countries with high-performing education systems and how these measurements were used.

Who wants to know and why?

The study aimed to provide an international perspective for policy makers and strategic planners involved in reviewing the types and uses of child outcome indicators as part of an initiative to ensure system-wide improvements in children's outcomes in England.

What did the research team find?

Education indicators were most commonly collected. In the main these covered attainment and participation in education training and employment. Social and emotional development and environmental indicators were also collated, albeit less frequently.

Health indicators were varied in type but found infrequently. Measures included aspects of general public health and healthy lifestyles.

Well-being indicators were also varied in type but not often found. They encompassed perceptions of well-being; family environment; relationships and social participation; education, employment and income; housing, homelessness and environment; and criminal activity.

Outcome indicators were often used for the purposes of monitoring and accountability, although the nature and purposes of these functions varied markedly between systems. Educational indicators were used for monitoring schools and national standards. They were sometimes used as a means of holding individual schools and the education system to account; they also informed the development of policies, as well as school and area improvement. Only two countries reported a broad range of outcomes in a holistic way.

What are the implications?

The reporting of a broad range of child outcome indicators in a meaningful way that allows policymakers and planners to use them is a serious challenge for most governments; there is considerable policy interest in how this can best be done. Localised accountability at school-level may not be the most productive way to improve the school system as a whole.

How were these results obtained?

The research team identified 13 countries with high-performing education systems. Through systematic searches of the internet

and recommendations from contacts, 109 documents were found which contributed to summaries of each country's approaches. From this, descriptive accounts were developed, outlining whether, which and how indicators were used.

CHAPTER ONE

Background

Aims and rationale for the study

The researchers set out to provide an analytical map of approaches to the measurement and monitoring of children's outcomes across education, health and well-being in high-performing education systems. They were commissioned to examine international perspectives on approaches to measuring and using information about children's outcomes, as well as to contribute to the emerging knowledge about the characteristics of high-performing education systems (see Barber and Mourshed, 2007).

Definitional and conceptual issues

High performance

Although there is considerable policy interest in comparative educational performance, there is as yet no reliable or agreed framework for identifying 'high-performing educational systems' nor any consensus about which outcomes should be used for judging performance. Some countries repeatedly perform well, but few tests cover the same range of countries and there is no evidence on 'value-added performance' – only crude outcome measures of the performance of a sample of young people. For this reason, one cannot be sure that countries which perform well are indeed high-performing, nor whether their high educational performance derives from the organisation, management and delivery of their education system rather than, say, cultural factors to do with the status of education.

In order to examine practices in education systems which perform well across a number of measures – which are referred to by the portmanteau term 'high-performing education systems' - the researchers developed a definition of high performance based on combining data across the UNICEF and cognitive performance tests. For the purposes of this study, *high-performing education systems were defined as those which perform well against both educational attainment indicators (PISA and TIMSS) and the UNICEF report that also included health and well-being indicators*. This is probably the first time that a research study has attempted to provide a clear framework for selecting and identifying high-performing systems.

Defining outcomes

The 'outcomes' of education are multiple. A persistent argument used in much critical policy work is that outcomes-based approaches to the assessment of education themselves narrow the scope of educational practices. The key markers of a civilised and educated society – tolerance, open-mindedness, creativity and so on – cannot be measured in terms of 'measurable' outcomes, and, indeed, the outcomes of

education go far beyond what can be measured (Pring, 2005). However, education is a large component of any government's public spending and treasuries typically look for evidence of returns on investment. In many countries, therefore, measures of the returns to educational investment are sought. Most typically, the return to educational spending has been measured in terms of *short-term cognitive* measures, such as examination results; *short-term non-cognitive* measures, including completion rates, participation rates in higher education or youth unemployment; and long-term returns to individuals in terms of lifetime earnings (Wolf, 2002; Goldin and Katz, 2008).

As the Every Child Matters (ECM) experience shows, however, what is especially challenging for the Government is to draw together relevant data in order to inform policy and practice, and it is these data in which the researchers are interested. They are not concerned with mapping *all* the data collected by governments on children (which would be a massive undertaking in itself). Instead, they set out to understand the data related to the outcomes of education and other children's services that governments draw on in their engagement with these services and service areas. We have therefore included material relating to children's outcomes in educational attainment (general cognitive or specific to reading, writing, mathematics or science), health and other measures of well-being which are seen by governments as relevant to their engagement with children's services providers. In addition, they examined material related to measuring, recording, reporting and assessing outcomes in relation to the use of these data as performance indicators.

Research question

What indicators are deployed to measure children's education, health and well-being outcomes in high-performing educational systems and how are they used?

In answering this question, the researchers of this report have sought to identify what knowledge exists about the following:

- whether indicators are deployed to measure children's education, health and well-being outcomes in high-performing educational systems
- which such indicators are used across a range of high-performing educational systems
- how chosen indicators are deployed

Policy and practice background

Policy

There is increased concern in and beyond this country about children's well-being, which has resulted in the English Government's ECM agenda, and the international UNICEF report (2007) on children's well-being. This policy strand is closely intertwined with another: the concern to compare system-wide performance in education between countries, which has been a feature of increasingly influential comparative studies such as Progress In International Reading Study (PIRLS), PISA and TIMSS. As a result, there is interest in how, if at all, high-performing systems measure indicators of children's well-being. This has led the Department for Children, Schools and Families (DCSF) to commission this study of practices in the identification, audit and monitoring of measures of children's education health and well-being, in order to obtain an international perspective against which to plan future research, develop policy, and review the system for monitoring children's outcomes in England.

Practice

In England, the Government has overseen an extensive programme of reform in children's service provision, including the construction of 150 local authority Children's Services Departments, the establishment

of Children's Trusts and the introduction of joint area reviews (Bachman et al., 2007). This programme culminated in the reshaping of central government administration and the publication of the Children's Plan with the aspiration to make England 'the best place in the world for children to grow up' (DCSF, 2007) and ensure system-wide focus on improving children's outcomes. Currently, the Government lacks an effective performance management system to test accountability across the wider range of outcomes, although it has recently published proposals to broaden the focus of school inspection to include wider outcomes.

Implementation of ECM highlighted stark difficulties in mapping children's outcomes. In England, accountability for educational outcomes, as measured in short-term cognitive indicators in literacy, mathematics and science (and a wider range of subjects at 16), has largely been at school-level. School-level accountability has been a powerful feature of education policy and practice, although widespread concern has been raised about the validity, reliability and fitness for purpose of the available measures (e.g. Statistics Commission, 2004). What is less clear is how schools and others might be held accountable for wider outcomes; the relationship between accountability at school-level and at local area level; and the availability of routinely collected indicators beyond the short-term cognitive measures provided by test results.

Research background

Children's outcomes

The Government is aiming to improve outcomes for all children and narrow the gap between the highest and lowest performing groups of children (Kendall et al., 2008). While there is evidence to suggest that children's cognitive attainment has improved in England over the last decade and a half, there remains concern about how this attainment compares with that of children in other developed countries (DCSF, 2007).

Research evidence for the relationship between national policy intervention and improved child outcomes is difficult to find. The most recent synthesis of the impact of welfare reform on children's outcomes concluded that impacts of reforms differ with the stage of a child's development, but most are relatively short-term (Grogger et al., 2002). Much international evidence on children's outcomes suggests that outcomes are driven by long-term structural features of children's lives, with exposure to poverty an overwhelming determinant (Jones et al., 2002; Plewis et al., 2001).

One set of concerns has related to the nature of appropriate indicators which might be used to explore children's outcomes. Most of the measures against the five ECM outcomes are negative indicators: measures of children's illness are more readily available than their health, of those occasions when they do not make a positive contribution (e.g. crime statistics) rather than of measures when they do. The measurement and assessment of children's outcomes is therefore a policy challenge for the Government.

Research methodology

The DCSF and the EPPI-Centre realised that a substantial amount of what is known about practice rests in policy and review documents rather than in the research literature, and sought an analytical map of this practice. So, although this study was commissioned by the EPPI-Centre, and deploys elements of the centre's conventional systematic review methodology, it also adopts a distinctive approach to the collection, collation, appraisal and presentation of evidence. Whereas conventional systematic reviews draw on a range of peer-reviewed research evidence to establish the current state of knowledge about a defined research question, this study uses policy documents and reports to develop analytical maps which describe a range of current policy practices in different polities. The approach here was to acquire information from relevant websites in a systematic way using defined search strategies, and to verify

the outcomes of those search strategies wherever possible by using informants in policy roles. The research team did not seek to reach conclusions about the effectiveness or impact of practices but to provide, on the basis of systematic enquiry, an analysis of approaches to the measurement of children's outcomes in high-performing education systems.

The study takes the form of a 'scoping map', as conceptualised by the EPPI-Centre. The model for this approach is the process of 'descriptive mapping' during a systematic review, which is designed to answer questions about what research is available on a given topic and uncover gaps.[1] A scoping map is intended to describe the characteristics of relevant literature rather than weigh the empirical evidence that exists in relation to the effectiveness or otherwise of different interventions. As a result, this report does not evaluate the methodological rigour of studies or synthesise their findings.

Funders, users and authors of the study

This mapping exercise was commissioned by the EPPI-Centre at the Institute of Education, University of London, on behalf of the DCSF for England. It is intended for policymakers and strategic planners in England who are reviewing the types and range of indicators currently in use for monitoring children's outcomes and holding children's services providers to account for their performance.

The maps may also be of use to policymakers, strategic planners and government bodies to whom children's service providers are accountable in other countries who are reviewing approaches to measuring outcomes for children and young people.

The research team was based at the Institute of Education, University of London. The team had considerable experience of using indicators of children's outcomes in education, health and well-being, having all previously evaluated Children's Trust Pathfinders. Professor Chris Husbands has a background in policy analysis and advice, particularly in relation to schooling; Ann Shreeve has practical experience of using indicators to monitor the education system in England; Dr Natalia R Jones has a background in quantitative research studies, particularly in the health sector, as well as working as an evaluator and researcher; and Professor Chris Husbands (Bills et al., 2007; Bills et al., 2008) and Ann Shreeve (Bills et al., 2007) have previous experience of using EPPI-Centre procedures for conducting systematic reviews.

1 http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=175

CHAPTER TWO

Methods of the review

Description of the methods

The research team began by establishing a set of criteria for identifying and sampling 'high-performing education systems', using data from four international studies of outcomes for children and young people:

- UNICEF 2007: Child poverty in perspective: an overview of child well-being in rich countries
- PISA 2007a, PISA 2007b: International standardised assessment of 15-year-olds
- TIMSS 2003 (Gonzales et al., 2004): Mathematics scale scores of eighth-grade students (13-14 year-olds)
- TIMSS 2003 (Gonzales et al. 2004): Science scale scores of eighth-grade students (13-14 year-olds)

Firstly, countries were identified which were listed in the top 10 of the UNICEF list and on at least one other list. This produced seven countries: Belgium, Denmark, Finland, Ireland, the Netherlands, Sweden and Switzerland. Secondly, countries were identified which were listed in the top 20 of all three of the non-UNICEF lists (PISA 2006; TIMSS 2003 science; TIMSS 2003 mathematics). This produced a further five countries: Australia, Hungary, Japan, Korea and New Zealand. Alongside these 12 countries, Singapore was also added as it was top of both the TIMSS 2003 mathematics and science lists. A summary of the characteristics of the selected countries can be found in the Technical Report, Appendix 2.1.

User involvement

Contacts in each country were identified by representatives of the DCSF and the EPPI-Centre; they were mainly located in the ministry responsible for education, although, in one case, they were at a university. The contacts included senior advisers, analysts, statisticians, planners and an academic who were involved in research, planning, information-sharing or international affairs. An email was sent asking them to arrange access for the research team to any recent government publications.

Identifying and describing material

Websites were searched systematically during May 2008 for relevant research, policy, legislation and statistics (see Technical Report, Appendix 2.2 for addresses of ministry websites). Inclusion and exclusion criteria were applied to titles, abstracts and contents (see Technical Report, Appendix 2.3). Further material was provided by ministry contacts by the cut-off date of September 17 2008.

Inclusion and exclusion criteria

To ensure that only relevant government publications and official or semi-official reports were included, an explicit list of inclusion and exclusion criteria was developed to exclude material that was inappropriate. During May 2008, the criteria were applied hierarchically to screen titles and abstracts beginning with inclusion criterion 1.

Inclusion criterion 1: Material relating to countries identified in the first stage of the research, regardless of the country of publication

Inclusion criterion 2: Material published within the last eight years. Justification: The researchers were concerned with current practice, rather than past practice. Eight years was a relatively arbitrary cut-off, but captures all approaches established and used since PISA 2000.

Inclusion criterion 3: Material relating to children and young people's (aged 0-19) outcomes in educational attainment (general cognitive, or specific to reading, writing, mathematics or science), children's health and other measures of children's well-being. Justification: to exclude further and higher education; to be more explicit about the age range of children/young people.

Inclusion criterion 4: Material related to measuring, recording, reporting and assessing children and young people's outcomes in relation to the use of this data as performance indicators

Inclusion criterion 5: Material that was in the English language

The documents and websites which made it through the title and abstract screening were screened in full during June 2008, using the original inclusion and exclusion criteria and the additional criterion:

Inclusion criterion 6: The most recent published report where it was part of a regular review cycle (e.g. annual report for 2007)

Analytic maps

The material remaining after the application of the inclusion and exclusion criteria were keyworded, using a study specific keywording sheet (see the Technical Report, Appendix 2.5) adapted from the EPPI-Centre Core Keywording Strategy (EPPI-Centre, 2002). In addition, researchers extracted relevant information about each country's approach using an information retrieval coding tool (see the Technical Report, Appendix 2.6) designed to gather evidence to answer:

- whether, and which, indicators were used by government(s)
- how indicators were used

Summaries of each country's approaches were complied, shared with contacts and revised. The revised summaries were analysed to produce the analytical maps of which and how indicators were used.

CHAPTER THREE

Results and findings

The initial screening yielded 114 papers potentially relevant to the analytic map. A further 54 papers were identified through handsearching ministries' websites. Allowing for papers that were unattainable because URLs did not work (7) and duplicates (9), 152 documents and an extra 21 documents that were identified during information retrieval stage went through to full-text screening, making a total of 173 documents. At the second stage of screening, a further 64 papers were excluded, again most commonly on the grounds that they did not meet the third criterion. This resulted in a final total of 109 papers that met the criteria for inclusion in the systematic map.

Volume and range of materials

A good amount of information (over 20 publications) was found for three countries (Australia, Singapore and Sweden) and a reasonable amount (between 10 and 20 publications) for a further five (Finland, Hungary, Ireland, Japan and the Netherlands). Fewer than ten publications were found for five countries: Belgium, Denmark, Korea, New Zealand and Switzerland.

Summary of types and use of indicators

The researchers established whether indicators were used for measuring education, health and well-being outcomes, and how they were used; they summarised findings for each country in the Technical Report, Appendix 3.1. Using these summaries, tables were compiled, recording the type and frequency of indicators and how they were used (see Technical Report, Appendix 3.2). The results of this analysis are summarised below.

Education outcome indicators found were mainly measures of attainment and participation in education and employment (all countries but limited information for Switzerland). Some measures of equity, schools equipment and teachers qualifications were also found. The Netherlands had recently introduced indicators of social and emotional development and the home and school environments.

Health outcome indicators were typically general public health or healthy life style measures and occurred infrequently across eight countries (Australia, Finland, Ireland, Japan, Korea, New Zealand, Singapore and Sweden). The most common outcome indicators of general public health were mental health, including suicide (which was also used as a measure of well-being), mortality, oral health, morbidity, injury and poisoning, sexual health and substance misuse. Most frequently used healthy lifestyle indicators were of physical activity and physical development.

Well-being outcome indicators were also varied in type and found infrequently in ten

Table 3.1: Map of types of education child outcome indicators

Type of education indicator	Details
Attainment	Subjects: language, mathematics and science, citizenship, etc. Data sources: national standardised tests, voluntary tests, periodic surveys, international surveys Timings: entering school, within the primary phase, on completion of primary school, within the secondary phase, on completing compulsory secondary school, after leaving compulsory school Sample groups: specific groups of children: boys and girls, ethnic groups, indigenous people, immigrants, second generation immigrants, bilingual pupils Analysis: progress, value added
Participation in education and employment	Enrolment Attendance Home schooling Suspensions Exclusion from school School completion Destination on leaving school Return to education after dropping out Dropout rates in higher education Youth unemployment Employment of graduates
Social and emotional development	Psychosocial aspects of pupils' development
Environmental factors	Home School

countries (Australia, Belgium, Finland, Ireland, Japan, Korea, the Netherlands, New Zealand, Singapore and Sweden). The most common indicators were socio-economic indicators of education, employment and income, followed by family environment, and relationships and social participation. Children's perceptions of their own well-being; housing, homelessness and environment; and criminal activity indicators were collected less frequently.

Map of educational outcome indicators

In this section are mapped the main education outcomes indicators which were used. Attainment indicators were common; these were mainly performance in subjects. In terms of measures of attainment, comments are made about the data sources used, when in a child or young person's life attainment was measured, which groups were sampled and how data was analysed. As well as attainment, participation in education and employment indicators was often measured, while a few countries collected indicators of social and emotional development and environmental factors.

Attainment in subjects

From the evidence base, it was found that almost all countries used attainment as a child outcome indicator, although not all the material found about attainment referred to specific subjects. However, there was evidence of literacy and competence in the national language being measured in many countries (Australia, Belgium, Hungary, Ireland, Japan, Korea, the Netherlands, New Zealand and Sweden) and second language learners' use of the national language in Australia and Sweden. Use of the indigenous

language by native people was measured in New Zealand. Numeracy and/or mathematics were measured in nine countries (Australia, Belgium, Hungary, Ireland, Japan, Korea, the Netherlands, New Zealand and Sweden). Competence in English was measured in Japan, Korea, the Netherlands and Sweden. The Netherlands also measured use of a second foreign language. Performance in science was measured in Australia, Ireland, Japan, Korea and New Zealand. Citizenship was measured in Australia and Hungary, and social studies in Japan and Korea. Information technology attainment was measured in Australia, Korea and the Netherlands. Belgium and the Netherlands measured performance in biology. Other subjects referred to only once, by the Netherlands, were physics, chemistry, environmental studies, history, geography, economics, technology, life skills, visual arts, music, drama and dance; they also measured intelligence. Australia, Denmark and Finland measured performance in vocational subjects. (See the Technical Report, Appendix 3.2, Table 3.2.1.)

International comparisons

Some countries referred to their performance in comparative tests: twelve countries mentioned PISA 2006 (all countries except Korea), four TIMMS (Belgium, Hungary, New Zealand and Singapore, and five PIRLS (the Netherlands, New Zealand, Singapore, Sweden and Switzerland). Some countries used indicators for making international comparisons, comparing their own performance with that of other countries. Evidence was found of the use of findings from international comparative studies of attainment of pupils. Four (Denmark, Finland, Ireland and Japan) of the 12 countries in the sample that participated in PISA 2006 took the results into account, including using them to understand better why they were successful, to compare their results against other countries, to identify trends or to identify areas for improvement. In Ireland, the report contextualising PISA 2006 results was produced by the school inspectorate. International comparators are also used by governments (Australia, Belgium, Hungary, Ireland, the Netherlands, New Zealand, Singapore, Sweden and Switzerland) as an external check on both the nationally school system and outcomes for children, with a view to identifying areas for investigation and (in Sweden) to target resources. (See Technical Report, Appendix 3.2, Table 3.2.2.)

Timings of measurement of attainment indicators

Attainment was most commonly measured at the end of compulsory schooling (Australia, Belgium, Hungary, Ireland, Japan, the Netherlands, New Zealand, Singapore and Sweden). In four countries, outcome indicators for attainment were measured within all three of the school phases: primary, lower secondary and upper secondary (Australia, Belgium, Hungary, and New Zealand). Attainment was measured at the end of primary schooling (Australia, Belgium, Ireland, Japan and the Netherlands) and at the end of the lower secondary phase (Australia, the Netherlands, Singapore and Sweden). A few countries measured children's attainment on entry to school (Australia, Hungary, the Netherlands and New Zealand). (See Technical Report, Appendix 3.2, Table 3.2.4.)

Participation in education and employment

Children's enrolment in school was commonly measured (Australia, Ireland, Japan, Korea, the Netherlands, New Zealand and Singapore) and occasionally pre-school participation (Australia and Ireland). Other indicators of participation were also used, albeit less often. These included pupils' actual attendance (Australia, Belgium, Ireland and Japan), suspensions and exclusions from school (New Zealand), truancy (New Zealand), school refusals (Belgium), retention in later years of schooling (Australia), home schooling (New Zealand) and grade repetition (Belgium). (See Technical Report, Appendix 3.2, Table 3.2.5.)

The school phase with the largest number of indicators attached to it was secondary.

Indicators were clustered around participation in education and employment outcomes for young people. Educational participation indicators included dropout rates in upper secondary school (Belgium, Finland, Korea, the Netherlands and Sweden), age on leaving school (New Zealand) and school completion rates (Finland). Post-secondary school education was measured: for example, destination on leaving school (New Zealand), second level education (Ireland), transfer to higher education (Belgium) and results at the end of the first year of higher education (Belgium). The return to education after dropping out was measured (Denmark and Sweden). There was also interest in collecting outcome indicators for dropout rates in higher education (Belgium). Employment indicators included youth unemployment one year after leaving school (the Netherlands and Finland), the unemployment gap between people in different levels of education (Sweden), and employment of graduates (Korea). (See Technical Report, Appendix 3.2, Table 3.2.6.)

Resource allocation

Although not specifically educational outcomes for children, evidence was found of indicators being collected that related to finance, resources, staffing and demographic patterns, which were used for planning school places and the overall education system (Denmark, Ireland, Japan, Korea, New Zealand and Sweden). Evidence was also found of resource or input indicators, such as numbers of students receiving financial help or training in Singapore and the numbers of computers per pupil in Denmark and Korea (see Technical Report, Appendix 3.2, Table 3.2.7).

Equity indicators

In some countries, specific groups were measured in order to monitor equality (Australia, Denmark, Finland, New Zealand and Singapore). In Denmark, Finland and Singapore, gender differences were examined, with other groups of pupils also scrutinised, including ethnic groups (Singapore), indigenous peoples (Australia and New Zealand), immigrants (Denmark) and bilingual pupils (Denmark). Denmark collected indicators about pupils with special educational needs. (See Technical Report, Appendix 3.2, Table 3.2.8.)

Further use of educational data

Progress was measured in Australia, Belgium, Singapore and Sweden. Sweden collected value added measures that used regression analysis developed from research showing that the socio-economic and national background of students, together with the gender composition of students, explained a large proportion of the statistical variance between the performance of pupils in different schools. The calculated residual effect was used as a measurement of the relative achievement of the school, as an approximation of the value added by the school. (See Technical Report, Appendix 3.2, Table 3.2.9.)

Map of health outcome indicators

The researchers note that there was overlap between some health and well-being indicators, and to a lesser extent with education. Health indicators were classified as general public health and healthy life styles. Countries classified indicators in different ways: for instance, Australia closely linked health with well-being and education outcomes.

No evidence was found of child health outcomes being collected in five of the countries in our study. Of the eight remaining countries, general public health outcomes were mental health, including suicide (which was also used as a measure of well-being) (Australia, Finland, Ireland, New Zealand and Singapore), mortality (Australia, Ireland and Singapore), oral health (Australia, Ireland and Japan), injury and poisoning (Australia, Ireland and Singapore), sexual health (Australia, Ireland and Singapore) and substance misuse (Australia, Finland and Ireland). Other general health indicators, occurring in at least two

countries, were morbidity (Australia and Ireland), disability (Australia and Ireland) chronic diseases (Australia and Ireland), auditory health (New Zealand and Sweden) and immunisation (Ireland and Singapore). The most common healthy lifestyle indicators were physical activity (Australia, Finland, Ireland, Japan, Korea and Singapore) and physical development (Finland, Ireland, Japan, Korea and Sweden). A less frequent measure of healthy lifestyles was diet and nutrition (Australia and Ireland), while perceptions of life expectancy were only measured in Australia. (See Technical Report, Appendix 3.2, Table 3.2.10.)

A composite list of health indicators with actual measures is provided in the Technical Report, Appendix 4.1, Table 4.1.1.

Table 3.2: Map of types of health child outcome indicators

General public health	Healthy life style
Life expectancy	Well-being
Mortality	Diet and nutrition
Morbidity	Physical activity
Disability	Physical development
Injury and poisoning	
Mental health	
Sexual health and reproductive health	
Chronic diseases	
Oral health	
Auditory health	
Substance misuse	
Immunisation	

Map of well-being outcome indicators

In order to provide policymakers with a broad range of well-being indicators, indicators of well-being were collated into six categories: well-being; family environment; relationships and social participation; education, employment and income; housing, homelessness and environment; and criminal activity.

Eight countries gathered well-being indicators on outcomes in education, employment and income (Australia, Finland, Ireland, Japan, Korea, New Zealand, Singapore and Sweden), while five measured relationships and social participation (Australia, Ireland, Japan, the Netherlands and Sweden) and family environment (Australia, Finland, Ireland, Singapore and Sweden). Four countries collected measures of general well-being – usually young people's perceptions (Australia, Belgium, Ireland and New Zealand). Three countries measured housing and homelessness (Australia, Ireland and Sweden) and Ireland collected data about young people's perceptions of their environment, such as safety and good places to go in their areas. Ireland, Japan and Singapore collected indicators of criminal activity. (See Technical Report, Appendix 3.2, Table 3.2.11.)

Australia, Ireland and Sweden had the most comprehensive range of well-being outcome indicators; they covered most of the categories in Box 3.3.

Box 3.3: Map of types of child outcome well-being indicators

- Children and young people's perceptions of well-being
- Characteristics of the family environment
- Peer and family relationships and social participation
- Education, employment and income factors that affect well-being
- Housing, homelessness and environmental factors
- Criminal activity

The data for the indicators were collected through surveys of children and young people such as 'perceptions of well-being' as well as routinely collected data such as 'young people subject to care and child protection orders'. Some of the indicators were of positive outcomes, such as 'participation rates in voluntary activities'.

A detailed table of well-being indicators and the actual wording of measures drawn mainly from these three countries can be found in the Technical Report, Appendix 4.1. The best examples were *Australia's young people their health and well-being* (Al-Yaman et al., 2003) (see Technical Report, Appendix 4.2, section 4.2.1) and the report *State of the Nation's Children: Ireland 2006* (Ireland, Office of the Minister for Children, 2006)(see Technical Report, Appendix 4.2, section 4.2.2). Only one well-being indicator was found for Belgium, Korea and the Netherlands.

Uses of indicators

Evidence was found of high-performing systems using educational outcome data on children and young people for monitoring both national standards (all countries except Switzerland) and schools (Australia, Belgium, Denmark, Finland, Hungary, Ireland, Japan, New Zealand and Sweden).

Indicators were used as a means of holding to account both individual schools (Australia, Belgium, Denmark, Finland, Hungary, Ireland, Japan, the Netherlands and Sweden) and the education system (Australia, Belgium, Finland, Hungary, Ireland, Japan, Korea, the Netherlands, New Zealand and Singapore). They were also used for informing the development of policies (Australia, Belgium, Finland, Hungary, Ireland, Japan, Korea, the Netherlands, New Zealand and Sweden) and for the purpose of informing individual school improvements (Australia, Belgium, Finland, Hungary, Ireland, Japan, the Netherlands and New Zealand). There was a little evidence that data was used for informing national improvement programmes (Australia, Belgium and New Zealand), directing resources (Belgium, Finland, Ireland, New Zealand and Sweden) and for holding states, local authorities or municipalities to account for child outcomes (Australia, Denmark and Sweden). Singapore used indicators to monitor the Convention of the Rights of the Child (see Technical Report, Appendix 3.2, Table 3.2.12).

Japan produced monitoring reports for education outcomes annually that also included health and well-being indicators (see Technical Report, Appendix 4.2, section 4.2.3). Ireland produced the report *State of the Nation's Children – Ireland 2006* as a baseline against which to examine future trends (see Technical Report, Appendix 4.2, Section 4.2.2).

Box 3.4: Map of uses of indicators

- Monitoring performance: child outcomes, national children's services, economic
- Accountability purposes: national, regional, schools
- Selecting pupils: streaming, types of school e.g. vocational
- Reporting performance: to parents, to pupils
- Informing national policies
- Monitoring equity
- Improving children's services systems
- Monitoring the Convention of the Rights of the Child

Some countries with high-performing education systems have distinguished between the monitoring and reporting mechanisms needed for evaluating the national education system as a whole, and for holding schools and other providers of services for children and young people to account. The monitoring of national trends requires routine collection of quantifiable data, based on standardised procedures that can be analysed in relation to different groups of children in different geographic areas. Indicators of attainment, participation in school and destinations on leaving school are important in this respect. School league tables as a reporting mechanism

were not seen as helpful in some countries – they were only used explicitly in Singapore; instead, schools were encouraged to judge themselves against other similar schools. Schools, on the other hand, need to be able to evaluate their own performance and need to review their performance against benchmarked data so that they can compare themselves with other schools in similar circumstances and report to governing bodies. Five countries (Belgium, Finland, Hungary, Ireland and Japan) reported data at national level but not at school-level as they were of the opinion that 'naming and shaming' schools in already poor socio-economic circumstance would not aid school improvement. Sweden used 'value added' indicators that took account of the profile of individual school's pupil population and used indicators to target resources to needs.

As well as noting that governments routinely use outcome indicators for monitoring and accountability, the researchers also found they were used for other purposes in most systems. Individual child indicators were used within schools for allocating pupils to teaching groups or to streams for particular activities and for managing admission to different types of schools (Singapore and the Netherlands). They were also used for formative and summative purposes in reporting progress and attainment to parents and pupils (Australia, Ireland, Japan, the Netherlands and Sweden).

Models of how indicators were used

It is clear that governments approach the deployment of data on child outcomes in different ways, balancing monitoring and accountability with different emphases. On the basis of the analytical maps, four models have been built that appear to characterise these different emphases; it is not the case these models exist in 'pure' forms in any of the study countries, but these types characterise different approaches.

Model 1: An accountability model In this model, outcomes are rigorously monitored at reporting levels (schools, regions and national) for the purposes of management and accountability, with a particular emphasis on schools. This approach requires national standards and benchmarks by which schools, states and local areas can compare their performance (Australia and the Netherlands).

Model 2: A school-community model This is a reporting model in which outcomes are monitored at national level and effort is focused at policy level on identifying and removing barriers to participation, but relatively relaxed about within system accountability. For example, in Sweden the indicator 'poor fluency in the national language' is used to direct resources to schools regarded as being in need of such additional funding. There may be no or very limited school inspection, with school self evaluation being important. External school inspection has an advisory role focused more on helping schools improve the curriculum, teaching and learning rather than evaluating the school's performance in terms of outcomes for children and young people. Child outcome indicators are used to direct resources to allow schools to compensate for inequalities, such as the failure of students to reach national attainment targets and to combat poor health by increasing health provision (Finland).

Model 3: A social capital model Here, improving child outcomes are part of plans to increase individual citizen's contributions to the economy of the country, and to establish a basis for strengthening social networks. Children are highly valued because declining birth rates and an aging population mean they will be the mainstay of the economy in the future. There is concern with young people's perceptions of their well-being and their take-up of opportunities for social participation, such as leisure activities, and their experiences of making a social contribution through volunteering and forming positive relationships with their peers. This spirit is embodied in the Japanese Zest for Living initiative that aims to improve the health of young people's minds and bodies and educational outcomes by

emphasising ethical living, public spirit and compassion for others as well as academic achievement (Japan and Singapore).

Model 4: A psycho-social model This model emphasises young's people's mental and physical health and well-being as prerequisites to improving learning outcomes. This draws on the relationship between health inequalities and access to education, and that between poor education outcomes and poor health outcomes. Young people who leave school earlier and who are unemployed perceive their health and their quality of life to be poorer than more advantaged groups who leave school later and have better jobs. The essence of this approach is glimpsed in Australia where each school is required to take account of issues, such as measures to combat bullying (Australia).

CHAPTER FOUR

Implications, or 'What does this mean?'

Strengths and limitations of the analytic maps

Limitations to the methods

In this study, the researchers departed from the systematic review convention by adapting a recently used search method pioneered by a previous review group (Bills et al., 2008) that used websites, rather than electronic databases to search for relevant material. In their research, they identified three different types of website containing country reports or comparative studies and developed a systematic search strategy for each, as well as undertook a more free ranging 'handsearch' of ministry websites. These search strategies helped ensure consistency in the website searches and enabled the research team members to use their time efficiently and effectively. In the case of government websites, the strategy provided a guide to searching that included an instruction to search more widely if the search produced limited results – similar to the 'handsearching' of journals used in the EPPI-Centre methodology. Materials found during the searches were screened for relevance in the usual way, using inclusion and exclusion criteria.

Their methods also differed from the EPPI-Centre methodology as they considered the range, volume and the relevance of material found for each country to the research questions so that they could assess whether or not they had a comprehensive evidence base from which to draw conclusions. Their judgements were based on the following:

- the number of sources of information for each country (range)
- the number of documents found for each country (volume)
- the amount of detail found in documents (relevance)

They did not appraise quality because they were not looking at research evidence. The materials examined were government documents, such as annual reports, policy documents or statistical profiles, and descriptive country or comparative reports from international organisations (OECD, INCA and Eurydice). From these documents, they were able to identify indicators used by governments and the uses to which they were put. The few academic studies available were descriptive reports of the approaches of different countries rather than evaluations or studies of relationships between different variables. These descriptive reports were useful as they contained information about whether and which child outcome indicators were used and the purposes for which they were used.

While their methods allowed them to triangulate their findings using different sources, their evidence base was restricted.

The systematic approach that they followed allowed for some deviation but in the main adhered to an agreed process. Language was a barrier and the researchers were aware from contacts that some of the information that they were seeking was available, but not in English. Clearly, the evidential base for this report substantially lay in policy documents and reports. Wherever possible, they tested their conclusions through direct contacts with informants in ministries and policy units in the study countries, but this was not always possible.

The focus of material included in the map

The searches produced a wealth of information about education indicators, but a smaller amount about health or well-being. This was not surprising as the search of government websites began with the ministry responsible for education. In most cases, the remit of the ministry responsible for education was narrowly focused on education; however in some countries, it was much wider and included some or all of the policy areas of culture, sport, science, technology, youth affairs, employment and community. The researchers' Irish contact helpfully directed them to information from their Ministry for Children. As they mostly dealt with education ministries, the majority of the government documents examined rarely contained measures of health and well-being; they mostly covered only educational outcomes.

It may be the case that, if the search were to have been widened to ministry sites responsible for health, social care or children, they might have found more outcome indicators in use. Where evidence of all three outcomes was found (for example, in Ireland and Japan, there was a greater likelihood that outcomes for children were being considered in a broader context, rather than as just in the domain of the education system.

The researchers expected to find evidence of the use of child outcome indicators in the annual reports of inspectorates of schools. They anticipated that these reports on education would hold the national system to account for outcomes for children. However, where annual reports on education systems were found, they tended to report the progress in implementing initiatives or compliance with government policies, such as the national curriculum and school self-evaluation, rather than national trends in outcomes for children and young people.

International studies with their focus on education provided information about the process of evaluation and monitoring used by different countries and gave us some details about the outcome data available. However, these international studies did not examine how services providing for children's health and well-being were monitored and evaluated which was a limitation to the evidence base.

What the analytic maps mean for decision makers

Some caveats have already been entered about the issues involved in learning from the practices of 'high-performing education systems', and the inherent difficulties in the concept on current measures. However, in this section, some possible implications of this work are identified.

In England, there is a rich collection of child outcome indicators for education, health and well-being. While the English dataset – especially in education – has been noted for its range and depth, combining indicators is more challenging. Currently assessment is something which is 'done to' pupils in whatever sphere of activity they are engaged. However, assessment could be augmented to include a greater emphasis on pupils' perceptions of their well-being and their experiences. Crucially, there may be a need for a periodic report that combines key education, health and well-being indicators to provide a comprehensive description of outcomes for children including trends that can be used by policymakers and planners.

The evidence of this study is that national standards can be monitored by analysing outcomes of standardised tests without the need to report at school-level, and this echoes recent policy work in the UK (Green et al., 2006). With appropriately benchmarked data, it is also possible to report the state of play and trends in schools in similar socio-economic circumstances - that is, in statistical neighbours. Such an approach would provide detailed information for the purposes of monitoring the performance of the education system as a whole, policymaking and prioritising the allocation of resources at lower overall cost. The evidence of some of the high-performing education systems explored is that sampling and rigorous national and sub-national reporting generates secure information about standards.

It also follows that child outcome indicators could be collected in alternative ways to current practices. Not all national testing needs to be annual or for the whole cohort; some subjects could be tested periodically and/or be carried out with a representative sample. For instance, within the primary and secondary phase, whole cohorts of pupils could be sampled and different subjects tested in different years. Online pupil perception surveys could be extended to a wider age range and given more importance and reported in a combined education, health and well-being annual report on outcomes for children.

An obvious use of indicators by high-performing education systems is for monitoring performance and socio-economic disparities between schools, and then using such data as a basis for developing policies for reducing social inequality. In England, due to the wide socio-economic differences in the population, equity is a major issue that the Government has found difficult to solve. The extent of the problem is illustrated by comparing performance and social equity in high-performing counties. The OECD/PISA (PISA, 2005) analysis of socio-economic disparities on student performance identified the UK and Ireland as having high average performance, but large socio-economic differences. The majority of high-performing countries in PISA in this study were identified as having high average performance and high social equity (Australia, Belgium, Hungary, Japan, Korea, the Netherlands, New Zealand, Sweden and Switzerland). This analysis suggests that reducing social inequality may itself be connected with higher performance. The effective use of indicators of equity by the Government could help focus resources and effort where it is needed.

Types and use of child outcome indicators

Of the types and range of education outcome indicators found across all high-performing education systems, the majority are collected in England at similar times in a young person's school and post-compulsory school career. Of those indicators not collected in England, it is worth noting 'competence in study skills', 'home and school environment' and 'pupils' psychosocial development' which were collected in the Netherlands. Young people's outcomes in these three areas are likely to contribute to their academic and personal development and may merit consideration as additional indicators.

An awareness of educational performance in comparison with other countries provides an international perspective that can help identify areas for development. Some countries made good use of their participation in international standardised assessment surveys and produced reports comparing their performance with other similar countries and exploring trends within their country. In general, it seemed that many of the systems examined were more conscious about their desire to compare their performance against international benchmarks, and less concerned to deploy data for intra-national comparison.

It is paramount that effective use is made of existing routinely collected data for health and well-being by those responsible for monitoring, evaluating and developing children's services. The successful use of

routinely collected data – and, concomitantly, the training of officers to make effective use of sometimes complex datasets – would seem to be sensible. General health indicators are available from the England Department of Public Health and indicators of children's health and life style may be routinely collected by other government bodies, such as indicators for 'housing and homelessness'. Young peoples' perception data is available through national online surveys and this could be extended to younger children. It would be worth reviewing the current data set for England against those in the list of health and well-being indicators given in the Technical Report, Appendix 4.1, Tables 4.1.1 and 4.1.2 and Appendix 4.2.

Monitoring education systems

These findings suggest that there are lessons to be learnt about which and how indicators of children's outcomes are collected, reported and used. As indicated above, some countries with high-performing education systems have distinguished between the monitoring and reporting mechanisms needed for *evaluating* the national education system as a whole and for holding schools and other providers of services for children and young people to account. In these countries, different approaches are used to monitor performance and to secure accountability for system development and outcomes. These approaches do not seem to increase the burdens on schools.

Monitoring equity

Reducing inequality requires indentifying pockets of deprivation and working to reduce it. In situations where the characteristics of the school population were changing on account of economic factors, such as immigration, or within country movements of population, equality and social cohesion were important considerations. To understand these issues, governments examined child outcome indicators: for example, 'competency in speaking the national language' in relation to specific groups of children and young people, such as children of recently arrived immigrants, second generation immigrants and indigenous people. These indicators were then used to target resources where they were most needed.

Monitoring the effectiveness of education, health and well-being, systems

There were few examples of the combined reporting of children's outcomes in education, health and well-being within one report. Only two examples were found of reports of national trends in a range of outcomes for children and young people: the annual statistical report *Japan's education at a glance 2006* (Japanese Ministry of Education, Culture, Sports, Science and Technology, 2006) (see Technical Report, Appendix 4.2, section 4.2.3 for a full list of contents) and the inaugural *State of the Nation's Children: Ireland 2006* (Ireland, Office of the Minister for Children, 2006)(see Technical Report, Appendix 4.2, section 4.2.2 for an extract from the summary of main findings).

The presentation of statistics in these ways gave the researchers the most complete descriptions of outcomes for children and young people in single countries. The evidence base for these reports drew on studies carried out periodically by various government departments coordinated in Japan by the ministry responsible for education, and in Ireland by the ministry responsible for children. The reports provided informative data sets for use by educators reviewing current provision and planning for the future.

A one-off report by the Australian Institute of Health and Welfare (*Australia's young people: their health and welfare 2003*) combined reporting of comprehensive health and well-being outcomes with some educational factors. The statistical analysis needed for this type of cross-cutting statistical report is a massive undertaking, requiring a coordinated initiative with access to data from across government departments and agencies.

In England, the government is data-rich and could produce a statistical analysis that brought together key indicators that would be useful for developing policy and strategic planning.

Methods of collecting child outcome indicators

The annual routine collection of outcome indicators for whole populations of groups of children is time-consuming for those involved in recording, collating and reporting. If the purpose of monitoring data is to provide information about the national system, other approaches may be more efficient and economic, such as the following:

- sampling, rather than whole population testing, as in Japan and New Zealand
- longitudinal cohort studies that sample groups of children as in the Netherlands
- periodic, rather than annual, sampling: Australia collected some data every three years, PISA standardised assessments are usually every three years.

The Government's concerns are noted that indicators used to measure the five 'every child matters' outcomes are mainly negative indicators. Some countries have used pupil perception data to collect positive information, such as 'participation rates in volunteer activities' or 'experience of helping to stop bullying or the bad behaviours of friends'. Perception data could be collected using on-line surveys, as in the Netherlands, and/or conducted at the same time as national tests as are student and teacher perception surveys in Japan.

Implications for future research

As mentioned earlier, there are many ways in which this approach breaks new ground. Methodologically, the adaptation of the conventional EPPI-Centre method, although not wholly unique, is relatively novel. Analytical maps can provide only a reasonably high level set of descriptions of practices, and beg many questions about the impact, effectiveness and the operation of the practices described. The concept of 'high-performing education systems', despite its ready acceptance in policy discourse, remains relatively unexamined.

For these reasons, the researchers conclude by outlining areas in which the Department may wish to consider undertaking more work. The first relates to understanding in greater detail the nature of high performance in education systems. While there appears to be evidence that some systems – notably those in Scandinavia – are able to sustain high levels of average performance, high levels of equity and high levels of children's well-being, other systems appear to be either actively or passively experiencing trade-offs between different aspects of performance. Considerable work is required to understand this, which will inevitably involve the relationship between children's outcomes, educational governance and the cultural settings in which children's outcomes are identified and managed. There is some evidence from the researchers' contacts to suggest that there might be considerable interest from other governments in addressing such questions.

The second area in which work might be done relates to the management of datasets and their use at various levels of the education system. The researchers have observed that education, health and well-being systems are not short of potential measures. The challenge is to use the data which is either already collected or which might be collected to inform action at various levels of the system: whether in terms of national policy-making, national administration, local administration or institutional leadership. While English schools have become expert users of data in the last decade and a half, in many cases the sophistication of the use of the data has far exceeded the reliability and validity of the data available – schools operate with very small sample sizes. Work might be done on addressing the scope to bring together more reliable and valid datasets and to equip policymakers and leaders with the skills needed

to use these constructively to inform policy and implementation; effectively, this would involve drawing on elements of what the researchers have called the 'school-community' model.

Linked to this, the researchers have been struck by the extent to which many of the systems have been seeking to benchmark and analyse their performance not in terms of its own internal strengths and weaknesses but against the findings of international surveys. Work might be done on linking the available English national datasets to international datasets which make this outward looking a comparatively routine activity.

References

Adonis A (2007) *Towards an 80% education system*. Speech at the Cass Business School, 25 June 2007.

Al-Yaman F, Sargeant H, Bryant M (2003) *Australia's young people: their health and well-being 2003*. Australian Institute of Health and Welfare. http://www.aihw.gov.au/publications/phe/ayp03/ayp03.pdf

Bachmann M, Husbands C, O'Brien M, Jones N, Shreeve A, Watson J (2007) *Children's Trust Pathfinders: innovative partnerships for improving the well-being of children and young people: National Evaluation of Children's Trust Pathfinders, Final Report March 2007*. London: DfES.

Barber M, Mourshed C (2007) *How the world's best performing education systems come out on top*. London: McKinsey and Company.

Bills L, Briggs M, Browne A, Gillespie H, Gordon J, Husbands C, Shreeve A, Still C, Swatton P (2006) Organisational structures, management processes and conceptual framework in initial teacher education: a systematic review. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Bills L, Briggs M, Browne A, Gillespie H, Gordon J, Husbands C, Phillips E, Still C, Swatton P (2008) International perspectives on quality in initial teacher education: an exploratory review of selected international documentation on statutory requirements and quality assurance. In: *Research Evidence in Education Library*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Department for Children, Schools and Families (2007) *Building brighter futures: the children's plan*. London: TSO.

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.

Goldin C, Katz L (2008) *The race between education and technology*. Cambridge, MA: Harvard University Press.

Gonzales P, Guzmán J C, Partelow L, Pahlke E, Jocelyn L, Kastberg D and Williams T (2004). Highlights From the Trends in International Mathematics and Science Study (TIMSS) 2003 (NCES 2005–005). U.S. Department of Education, National Center for Education Statistics. Washington, DC: US Government Printing Office.

Green S, Bell J, Oates T, Bramley T (2006) *Alternative approaches to national assessment*. Cambridge: Cambridge Assessment.

Grogger J, Karoly LA, Klerman J (2002) *Consequences of welfare reform: a research synthesis*. San Francisco: RAND.

HM Treasury (2003) *Every child matters*. London: The Stationery Office. www.everychildmatters.gov.uk/_files/EBE7EEAC90382663E0D5BBF24C99A7AC.pdf.

Ireland, Office of the Minister for Children (2006) *State of the nation's children: Ireland 2006*. Dublin: The Stationery Office. http://www.omc.gov.ie/documents/Publications/SONC_final.pdf.

Japanese Ministry of Education, Culture, Sports, Science and Technology (2006) *Japan's education at a glance 2006*. http://www.mext.go.jp/english/statist/07070310.htm

Jones C, Clark L, Grusec J, Hart R, Plickert G, Tepperman L (2002) *Poverty, social capital, parent and child outcomes in Canada*. Quebec: Human Resources Development.

Kendall S, Straw S, Jones M, Springate I, Grayson H (2008) *Narrowing the gap in outcomes for vulnerable groups: a review of the research evidence*. Slough: NFER.

PISA (2005) *School factors related to quality and equity: results from PISA2000*. Paris: OECD.

PISA (2007a) *PISA 2006 science competencies for tomorrow's world. Volume 1: analysis*. Paris: OECD.

PISA (2007b) *PISA 2006 science competencies for tomorrow's world. Volume 2: data*. Paris: OECD.

Plewis I, Smith G, Wright G, Cullis A (2001) *Linking child poverty and child outcomes: exploring data and research strategies*. Research Working Paper No 1. London: Department for Work and Pensions.

Pring R (2005) *Philosophy of education: aims, theory, common sense and research*. London: Continuum.

Statistics Commission (2004) *School education statistics: user perspectives*. London: Statistics Commission.

UNICEF (2007) *Child poverty in perspective: an overview of child well-being in rich countries*. Innocenti Report Card. Florence: UNICEF Innocenti Research Centre.

Wolf A (2002) *Does education matter? Myths about education and economic growth*. London: Penguin.

Appendix 1: Authorship of this report

This work is a report of a systematic review conducted by the School Accountability Review Group

The authors of this report are:

Chris Husbands (Institute of Education, University of London)
Ann Shreeve (Institute of Education, University of London)
Natalia R. Jones (Institute of Education, University of London)

They conducted the review with the benefit of active participation from the members of the review group.

For further information about this review, please contact:

Professor Chris Husbands
Faculty of Culture and Pedagogy
Institute of Education, University of London
20 Bedford Way,
London WC1H 0AL

Email: c.husbands@ioe.ac.uk

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
London WC1H 0AL

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

Acknowledgements

The review was sponsored and funded by the Department of Children, Schools and Families

Appendix 2: The standard EPPI-Centre systematic review process

What is a systematic review?

A systematic review is a piece of research following standard methods and stages (see figure 1). A review seeks to bring together and 'pool' the findings of primary research to answer a particular review question, taking steps to reduce hidden bias and 'error' at all stages of the review. The review process is designed to ensure that the product is accountable, replicable, updateable and sustainable. The systematic review approach can be used to answer any kind of review question. Clarity is needed about the question, why it is being asked and by whom, and how it will be answered. The review is carried out by a review team/group. EPPI-Centre staff provide training, support and quality assurance to the review team.

Stages and procedures in a standard EPPI-Centre Review

- Formulate review question and develop protocol
- Define studies to be included with inclusion criteria
- Search for studies – a systematic search strategy including multiple sources is used
- Screen studies for inclusion
  - o Inclusion criteria should be specified in the review protocol
  - o All identified studies should be screened against the inclusion criteria
  - o The results of screening (number of studies excluded under each criterion) should be reported
- Describe studies (keywording and/or in-depth data extraction)
  - o Bibliographic and review management data on individual studies
  - o Descriptive information on each study
  - o The results or findings of each study
  - o Information necessary to assess the quality of the individual studies

At this stage the review question may be further focused and additional inclusion criteria applied to select studies for an 'in-depth' review.

- Assess study quality (and relevance)
  - o A judgement is made by the review team about the quality and relevance of studies included in the review
  - o The criteria used to make such judgements should be transparent and systematically applied
- Synthesise findings
  - o The results of individual studies are brought together to answer the review question(s)
  - o A variety of approaches can be used to synthesise the results. The approach used should be appropriate to the review question and studies in the review
  - o The review team interpret the findings and draw conclusions implications from them

Quality assurance (QA) can check the execution of the methods of the review, just as in primary research, such as:

- Internal QA: individual reviewer competence; moderation; double coding
- External QA: audit/editorial process; moderation; double coding
- Peer referee of: protocol; draft report; published report feedback
- Editorial function for report: by review specialist; peer review; non–peer review

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=2424&language=en-US

First produced in 2008 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/

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.

The views expressed in this work are those of the authors and do not necessarily reflect the views of the funder. All errors and omissions remain those of the authors.

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