Clearinghouse – Research Series · 2010 Number 06

Input, Process, and Learning in primary and lower secondary schools

A systematic review carried out for The Nordic Indicator Workgroup (DNI)

Technical Report

By

Sven Erik Nordenbo
Anders Holm
Eyvind Elstad
Jaap Scheerens
Michael Søgaard Larsen
Michael Uljens
Per Fibæk Laursen
Trond Eiliv Hauge

DANISH CLEARINGHOUSE FOR EDUCATIONAL RESEARCH

THE DANISH SCHOOL OF EDUCATION
AARHUS UNIVERSITY

Copenhagen 2010

Input, Process, and Learning in primary and lower secondary schools - A systematic review

The Danish Clearinghouse for Educational Research
is a unit at the Danish School of Education, Aarhus University

Title	Input, Process, and Learning in primary and lower secondary schools - A systematic review carried out for The Nordic Indicator Workgroup (DNI)
Copyright	© 2010 by Danish Clearinghouse for Educational Research
Translation into English	Mike Robson
ISBN	978-87-7684-323-6
ISSN	1904-52-55
Review group	Professor Eyvind Elstad, University of Oslo Professor Trond Eiliv Hauge, University of Oslo (until 2009.04.28) Professor Anders Holm, University of Copenhagen Professor Per Fibæk Laursen, Aarhus University Professor Michael Uljens, Åbo Academy University Professor Jaap Scheerens, University of Twente
Peer reviewer	Professor Jan-Eric Gustafsson, University of Gothenburg
Advisory group at Danish Clearinghouse	Professor, Director Sven Erik Nordenbo Associate Professor, Deputy director Michael Søgaard Larsen Head of Communication Mette Thornval Evidence consultant Neriman Tiftikci *Scientific assistants* Anne Bang-Olsen Hanna Bjørnøy Sommersel Henriette Nobili Christiansen Majken Mosegaard Svendsen Rikke Eline Wendt Trine Kløveager Nielsen
Section	Technical report
Authors	Sven Erik Nordenbo (author and editor) Anders Holm Eyvind Elstad Jaap Scheerens Michael Søgaard Larsen Michael Uljens Per Fibæk Laursen Trond Eiliv Hauge (until 2009.04.28)
Dansk Clearinghouse – reference number	435201
Month and year ofpublication	June, 2010
This report shall be cited as	Nordenbo, S.E., Holm, A., Elstad, E., Scheerens, J., Søgaard Larsen, M., Uljens, M., Fibæk Laursen, P., & Hauge, T.E. (2010) Input, Process, and Learning in primary and lower secondary schools. A systematic review carried out for The Nordic Indicator Workgroup (DNI). In: The Evidence Base. Copenhagen: Danish Clearinghouse for Educational Research, DPU, Aarhus University
Contact address (postal address, phone, e-mail)	Danish Clearinghouse for Educational Research School of Education Aarhus University Tuborgvej 164 DK-2400 Copenhagen NV Phone: +45 8888 9980 clearinghouse@dpu.dk www.dpu.dk/clearinghouse

Foreword

This report presents a systematic review of empirical research published internationally between 1990 and 2008 on the relationship between factors in primary and lower secondary schools (input and process) and pupils' learning (output and outcome). The project was commissioned by the *Danish Evaluation Institute* (*Danmarks Evalueringsinstitut)* and was performed on behalf of the Nordic Indicator Workgroup (DNI). DNI is a workgroup nominated by the Nordic Evaluation Network, which consists of representatives of The Agency for the Evaluation and Quality Development of Primary and Lower Secondary Education in Denmark, the Danish Evaluation Institute, the Swedish National Agency for Education, the Norwegian Directorate for Education and Training, the Finnish National Board of Education and the Ministry of Education, Science and Culture in Iceland. The project was carried out in the period 1.10.2008-15.01.2010.

Danish Clearinghouse wishes to express its warmest thanks to the Review Group and the Peer Reviewer, which not only accepted our invitation to participate in the project, but also – despite large workloads outside the project – devoted additional time and effort at critical moments in order to meet the fixed and rather tight deadlines. We also wish to express our thanks to Professor Peter Allerup, Aarhus University, for developing a statistical test and to Professor Mads Jæger, Aarhus University, for taking part in the scientific discussions.

Danish Clearinghouse wishes to thank *the National Library of Education, Denmark* for exemplary assistance and for help in obtaining the many documents on which the report is based.

Finally, the Clearinghouse wishes to thank *the Nordic Indicator Workgroup (DNI)* for setting the task, and especially the excellent

working relationship with Special Advisor Signe Ploug Hansen, Danish Evaluation Institute, and Director of Education Gunnar Iselau, Swedish National Agency for Education, who acted as contact point to DNI.

This document was completed March 2010, and revised June 2010.

Sven Erik Nordenbo
Danish Clearinghouse for Educational Research

The results of this systematic research review are available in five formats:

Summary	Explains the purpose of the review and its principal conclusions
Data sheet	Describes the components of the technical report
Report	Describes the results without technical details
Technical report	Describes in detail the context, methods, studies and results
Database	Access to the database containing descriptions and classifications of the individual studies included in the review

All formats may be accessed through www.dpu.dk/clearinghouse

Summary

What do we want to know?

What empirical research has been carried out to examine the relationship between factors in primary and lower secondary schools (inputs and processes) and the learning achieved by primary and lower secondary school pupils (outputs and outcomes)? What are the results with weight of evidence of this empirical research?

Who wants to know and why?

The project was commissioned by the *Danish Evaluation Institute* (*Danmarks Evalueringsinstitut)* and was performed on behalf of the Nordic Indicator Workgroup (DNI). DNI is a workgroup nominated by the Nordic Evaluation Network, which consists of representatives of The Agency for the Evaluation and Quality Development of Primary and Lower Secondary Education in Denmark, the Danish Evaluation Institute, the Swedish National Agency for Education, the Norwegian Directorate for Education and Training, the Finnish National Board of Education and the Ministry of Education, Science and Culture in Iceland.

The task has been to establish which factors or constellation of factors in the school are the most important for producing desired results that might be relevant for the development of a reliable indicator instrument for supervision and development etc. within the primary and lower secondary school sector.

What did we find?

From 1990 to 2008, 109 studies were published on malleable school factors within school effectiveness research. Of these studies, 71 are of high or medium weight of evidence. Synthesising these studies establishes that 11 school factors (some with subcategories) are of im-

portance for high pupil achievement. The school factors and subcategories identified are the following: Human Resources (Management and Leadership); Educational Leadership (Management and Leadership); Opportunity to Learn (Curriculum/scheduling); Disciplinary Climate (School Culture and School Climate); Achievement/progress Orientation (School Culture and School Climate); Interrelational Climate (School Culture and School Climate); Social norms and values (School Culture and School Climate); Teacher behaviour (Teacher); Teacher as an Organisational Actor (Teacher); Pupil Composition of the School; and Parental Relationship.

What are the implications?

For practice: The school leader should realize that a number of aspects of his or her work are important for pupil learning: the more he or she is available for teachers the better; the more the principal's policies are concerned with teachers' growth the better; the more teachers and parents are involved in school decisions the better. The principal should demonstrate strong leadership, above all in the areas of curriculum and instruction, and should be able to involve other staff members in leadership activities and position. The principal's behaviour ought to be supportive and egalitarian and neither directive nor restrictive, and should be 'resource supportive', e.g. in deciding textbooks and contents of the teaching. The teacher's efficiency of organising the instruction process improves pupil learning; this is measured by the percentage of time teachers reported spending on the planning of their lessons for the following day, the making of a weekly teaching plan, keeping to the timetable, and the assigned time spent on lessons,. It also includes homework hours, which are total hours pupils spent on homework both in school and out of school per week. In a good school an orderly atmosphere prevails, and also an ordered environment, in which appropriate pupil behav-

iours are present. A good school for pupils is a school where pupils do not feel unsafe, since the proportion of pupils who feel unsafe has a significant negative effect on pupil achievement. A good school focuses on academic achievement and high expectations, high pupil engagement exists and negative peer pressure is absent. Teachers rate attentiveness and have established a 'learning climate'. In a good school, staff and pupils show affiliation and support/respect, there is a warm teacher/pupil relationship, teachers can obtain assistance, advice and encouragement and they are made to feel accepted by their colleagues. Pupils develop positive relationships with each other. Good schools employ various means of communication and interaction with the parents. Parents are invited to be active on School Boards, and are given the opportunity to participate in leadership decisions. Schools give parents tips about homework and encourage them to participate in focus groups and surveys to uncover children's and parents' needs. Parents' support of children and involvement in school matters and community partnership are important.

For policy: Policymakers can influence pupil learning through choice of the pupil composition of the school. Policymakers can promote pupil achievement by helping to identify strengths and weakness in school by developing indicator systems for malleable school factors and subcategories.

For research: Although research in the 'good school' to a certain extent is based on high quality data and sophisticated statistical models, taking into account that data is sampled as clusters (students within classes and classes within schools) and thus reporting the correct standard errors, it is also evident that no studies in this review seriously address causality in terms of using experimental or quasi-experimental data or statistical methods that allow for causal interpretation. It seems that there is a complete lack of interest in estab-

lishing causal directions in 'good school' research. Concerning the problems of the causal direction of school inputs, it is evident that it is crucial that future research takes causality more seriously. Taking causality seriously also means that new requirements must be made to data, requirements that are not always met by existing data sources. Thus the research community must also convince policy makers that a new causal agenda in school research is needed.

How did we arrive at these results?

The project has had four principal phases. First we searched all relevant sources for research that had been published during the period 1990-2008. Next we went through the studies that had been found in order to ensure that only those that were relevant were included. Then we extracted relevant data out of the studies using, among other things, a software programme developed by the EPPI-centre, University of London. Finally the research mapping was carried out on the relevant studies, and syntheses were formulated where possible.

For further information

The study is included in the Evidence Base set up by the Clearinghouse for Educational Research. Here a link can also be found to the basis for the research, the Concept Note, that governs the research process at the Danish Clearinghouse for Educational Research, see www.dpu.dk/clearinghouse.

Contents

Tables

Figures

Frames

1 Introduction

1.1 Background and problem area

This report has been written on the basis of a contract between the *Nordic Indicator Workgroup* (DNI) and *Danish Clearinghouse for Educational Research, DPU, Aarhus University*.

DNI is a workgroup nominated by the Nordic Evaluation Network, which consists of representatives of The Agency for the Evaluation and Quality Development of Primary and Lower Secondary Education in Denmark, the Danish Evaluation Institute, the Swedish National Agency for Education, the Norwegian Directorate for Education and Training, the Finnish National Board of Education and the Ministry of Education, Science and Culture in Iceland.

The research mapping and synthesis presented in this report consists of mapping and synthesis of research that addresses the relationship between the primary and lower secondary school's efforts and its pupils' learning.

1.2 Aim

The aim of this research assessment can be summarised in the question:

> What empirical research has been carried out to examine the relationship between factors in primary and lower secondary schools (inputs and processes) and the learning achieved by primary and lower secondary school pupils (outputs and outcomes)?
>
> What are the results with weight of evidence of this empirical research?

The questions can be addressed as follows:

- *By performing a systematic research mapping of the empirical research that has been carried out to examine the relationship between factors in primary and lower secondary schools (inputs and processes) and learning achieved by pupils (outputs and outcomes).*
- *By performing a systematic synthesis of research with sufficient weight of evidence identified in the systematic research mapping.*

1.3 Review group

To carry out the task, Clearinghouse established a review group with the following members:

Professor Eyvind Elstad, University of Oslo, Norway
Professor Trond Eiliv Hauge, University of Oslo, Norway (until 2009.04.28)
Professor Anders Holm, University of Copenhagen, Denmark
Professor Per Fibæk Laursen, Aarhus University, Denmark
Professor Jaap Scheerens, University of Twente, the Netherlands
Professor Michael Uljens, Aabo Akademi University, Finland

The review group participated with Danish Clearinghouse in the data extraction and coding of the research reports covered by this study. The final report was produced by Danish Clearinghouse for Educational Research and the review group in cooperation.

Clearinghouse has asked Professor Jan-Eric Gustafsson, University of Gothenburg, Sweden, to peer review an earlier version of the report, which was completed January 15th 2010. Jan-Eric Gustafsson accepted this commission. In working out the final version of the report Clearinghouse has learned both from the peer reviewer, from

comments of the review group members and from members of the *Nordic Indicator Workgroup* (DNI). Clearinghouse and the review group are solely responsible for the final version.

There have been no conflicts of interest for any member of the review group or the peer reviewer during the data extraction process and the preparation of the report. No review group member has participated in the coding of own research reports.

2 Methods used in the research mapping

2.1 Design and method

This research mapping has been carried out following a standardised procedure described in the *Concept Note* developed by *Danish Clearinghouse for Educational Research*
(see http://www.dpu.dk/site.aspx?p=9864).
The procedure is described in a protocol established at the start of the project. The procedure is characteristic in utilising transparent and explicit methods in a series of steps. This is explained further in this report and also (briefly) in the *Concept Note*.

A special software tool was used, developed especially for this type of study: the EPPI-Reviewer. This is explained in more detail on the producer's website: http://eppi.ioe.ac.uk.

Data extraction from relevant and suitably qualified documents was carried out following the methodology and systematic of the EPPI-Reviewer. This procedure was developed by the *EPPI-Centre* at the *Institute of Education, University of London*. In this particular research mapping the procedure was adapted to the conceptual universe of the research in question – see Chapter 3.

The research mapping was carried out on the basis of coding and evaluation of the research reports by a review group working together with the staff of *Danish Clearinghouse for Educational Research*. The studies were characterized and their thematic relationships analysed.

2.2 Conceptual delimitation

The starting point of the research mapping was the two review questions:

What empirical research has been carried out to examine the relationship between factors in primary and lower secondary schools (inputs and processes) and the learning achieved by pupils (outputs and outcomes)?

What are the results and conclusions of such research?

The research mapping was intended to uncover factors relevant for pupils' learning emerging from a broad interpretation of the concept of 'the good school' – including physical layout, ways of teaching, teacher competences, administration etc., thus bringing in all the data about inputs (the factors determined by the school), processes (the school's activities) and outputs (the pupils' results), that might be relevant for the development of a reliable instrument for supervision and development etc. within the primary and lower secondary school sector.

The task has been to establish which factors or constellation of factors in the school are the most important for producing the desired results. Since the way in which the various factors interact is also important for the combined effect, we have searched for studies that describe synchronous effects.

This implies that studies on a single feature of the school, for example 'teacher effectiveness' or 'the competence of school leaders' were not included. Individual factors were included only where they were viewed in relationship with other factors in the school, i.e. in a total perspective of the school. The approach adopted for this research mapping has been 'school effectiveness'.

In this approach the school is seen as an institution, and concepts are employed that make it possible to state which factors in the school lead to effects in the short term (output) and/or on the longer term (outcomes). In this research mapping exercise, 'the good school' is

therefore regarded as an empirical phenomenon. In other words, 'the good school' is a school that has proved that it lives up to certain desirable, explicit criteria, corresponding to those set up by research looking for *School Effectiveness*, a research tradition internationally anchored in the '*International Congress for School Effectiveness and Improvement*' (ICSEI).

The following concepts, taken from the 'ERIC Thesaurus', will be used:

School effectiveness

Degrees to which schools are successful in accomplishing their educational objectives or fulfilling their administrative, instructional, or service functions.

Effective schools research

Educational research focused on identifying unusually effective schools, studying the underlying attributes of their programs and personnel, and designing techniques to operationalise these attributes in less effective schools.

Research into effective schools is based on a theory that the results achieved by a school are based on (a) the individual abilities of the pupils, (b) the cultural, socio-economic and family background of the pupils and (c) what the pupil experiences at the school.

Effective schools research seeks information about factor (c), and must attempt to control and correct any influences arising from the other two factors. In effective schools research an analytical distinction is sometimes drawn between phenomena at the school level and at the classroom level (Creemers & Kyriakides, 2008). The classroom level is admittedly a part of the school, but is only of interest for the current study if it is seen in the context of the school as a whole. 'Good classrooms' can also be found in 'not very good schools', and

vice versa. In this study, the focus is 'school effectiveness', not 'teaching effectiveness'.

The concept of 'school effectiveness' only gives meaning in relation to certain *criteria* that an effective school must meet. The question then is to define these criteria. In research into school effectiveness, these criteria are formulated as the *desired effects* expressed as 'outputs' or 'outcomes'.

There is an indefinite number of possibilities. For the purposes of this study it has been decided that only effects on pupils have any interest. In the short term such effects might be e.g. the results achieved in specific school subjects, the acquisition of certain generally valued competences, or whether the pupils thrive in the school.

On a longer term, relevant effects might be the various functions or effects of the school seen from a societal viewpoint: economic effects, effects on the cohesiveness of local society, or effects on cultural life in the community. Such effects are *not* included in this analysis.

Initially it is unlikely to be the same basic factors in all schools that create such a diversity of effects. In the synthesis process it has been necessary to make additional conceptual distinctions in this area, cf. Chap. 4. In connection with this research mapping exercise, however, it is not necessary to introduce any other delimitation than stipulating that the effects must be relevant to the pupils.

Interest is also restricted to *schools* that in their nature are similar to the Nordic basic schools, i.e. schools internationally characterised as 'primary and lower secondary schools'. The study only considers normal schools, not special schools or vocational schools.

Most other industrialised countries have school systems that differ organisationally from the Nordic system. Most industrialised countries divide their school system into 'primary school' and 'secondary

school'. Since this research mapping covers research on schools similar to the Nordic basic school, it includes research focusing on 'primary school', and research focusing on 'lower secondary school'.

This study is only interested in schools in societies resembling the Nordic societies. This means in practice that studies on 3rd world schools are *not* considered relevant to this study.

'School' is generally recognised to be a non-constant phenomenon. Thus, in principle, any school research from any period in time cannot be relevant. However, it can be difficult to stipulate *one particular year* since which research can be considered to be particularly relevant to the present day. In the first half of the 1990's, however, the legal basis of the basic schools in a number of Nordic countries was changed considerably (Tjeldvoll, 1998). This might indicate that 1990 would be a good starting year for this research mapping exercise.

This cut-off year could also be defended from a viewpoint of research methodology, since around 1990 school effectiveness research began to utilise a new research design that made research results more reliable. At this time the research tradition began to employ new statistical methods that permitted simultaneous analysis of hierarchical data. This is interesting, because what the pupils experience in the school takes place both at classroom level and at a leadership and organisational level (Willms, 1994; Creemers , B. et al., 1992).

To this can be added that there are several thorough research reviews that cover research prior to 1990 in a competent manner (Scheerens, 1997; Teddlie & Reynolds, 2000; Townsend, 2007).

As an illustration of the relationship between the conceptual delimitations discussed here we can refer to Figure 2.1

The model indicates that there are at least three basic relationships contributing to what the pupil gets out of the school: (a) the individual abilities of the pupil, (b) the social background – in a broad sense - of the pupil and (c) the character of the school at which the pupil is taught. The present research mapping and synthesis only looks at the outputs and outcomes that can be ascribed to the contribution of the school itself. This is achieved by correcting as much as possible for factors related to (a) and (b).

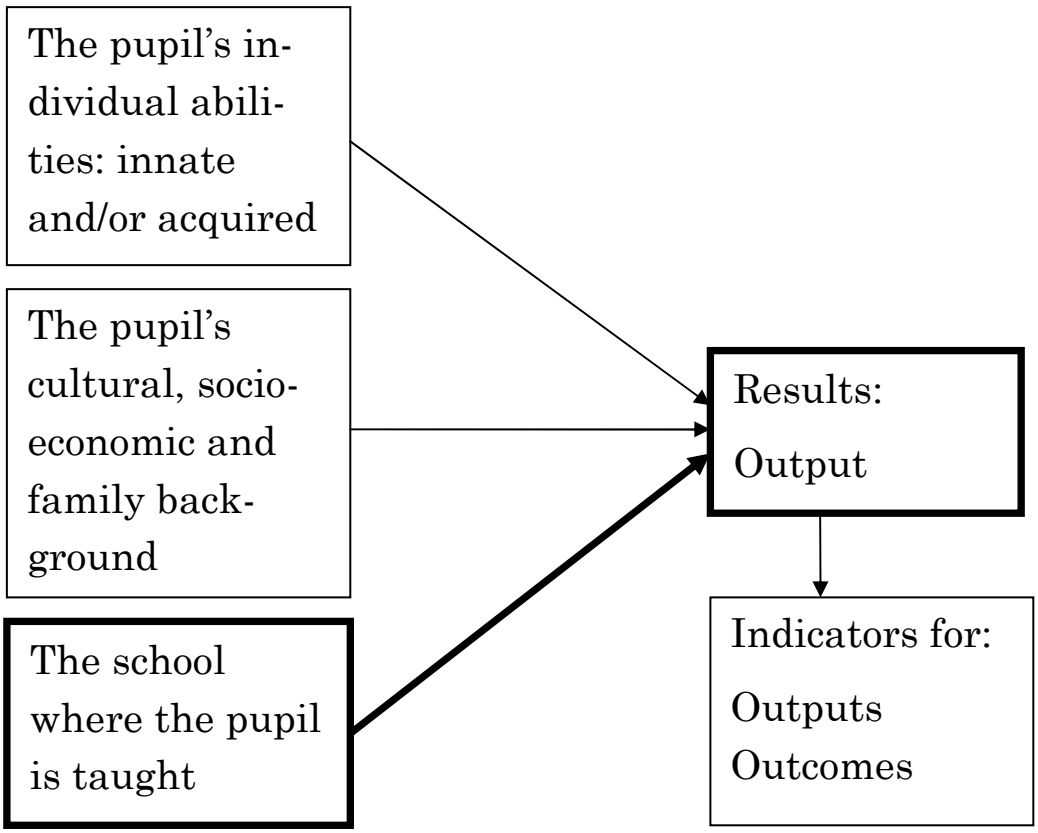

Figure 2.1: School effectiveness and indicators – conceptually simplified relationship

Similarly, as already mentioned, the establishment of an indicator system is a separate research task which can be undertaken after the conclusion of this research mapping and any subsequent research

synthesis built on the studies identified in this research mapping exercise.

2.3 Searches

Searches were carried out by the Clearinghouse. The review group had the opportunity to discuss and correct both the sources to be searched and the search profiles. Both the search sources and the search profiles were explicitly described in the research mapping protocol set up in the initial phase of the project.

From the start the review group as well as the members of the DNI Group were encouraged to suggest additional references. During the project, seven such suggestions were considered. Of these only one study fulfilled the inclusion criteria.

The core of the research mapping exercise has been 'the Good School', i.e. the characteristics of a school that creates the desired effects in its pupils. The special approach to school relationships adopted in school effectiveness research has also been used here.

The professional universe of this review covers didactics and educational research, including more psychologically oriented and more sociologically oriented directions. It was therefore desirable to achieve the same breadth of scope in the sources that were searched and in the search profiles that were employed. The linguistic universe was initially defined as Danish, Swedish, Norwegian, German, French and English. The search process did not specify any restrictions with regard to research methodologies; this aspect was taken into account in the screening process – see Section 2.4. Sources and hits are shown in Table 2.1. All searches were uploaded in the software EPPI-Reviewer.

Source	Date of search	Number of hits
BEI (dialog)	21/11/2008	150
AEI (Dialog)	24/11/2008	500
Psychinfo(CSA)	24/11/2008	260
ERIC(CSA)	21/11/2008	1293
Evidensbasen	27/11/2008	21
Sociological abstracts(CSA)	25/11/2008	98
Fis Bildung	26/11/2008	801
CBCA Education (Proquest)	26/11/2008	107
Dansk Pædagogisk Base(DPB)	27/11/2008	29
forskningsdatabasen.dk	03/12/2008	10
Libris (Sweden)	27/11/2008	17
Skolporten.com	27/11/2008	2
Norbok (Norway)	01/12/2008	12
Bibsys Forskdok publikasjoner (Norway)	01/12/2008	52
Jykdok	01/12/2008	6
Swetswise	01/12/2008	122
Google Scholar	03/12/2008	153
References from included studies	Continuous during review process	11
References from review group/DNI Group	Continuous during review process	7

Table 2.1: Searches performed

2.3.1 Search profiles

The searches covered material published during 1990-2008, as presented below. All search profiles were formed in accordance with the theme of the research mapping, paying particular attention to the subject data systems and professional content of the sources that were searched. All searches were done in November-December 2008.

2.3.1.1 Searches performed

BEI (Dialog)

('HIGH SCHOOLS' OR 'COMMUNITY SCHOOLS' OR 'ELEMENTARY SCHOOLS' OR "INDEPENDENT SCHOOLS' OR "MAINTAINED SCHOOLS' OR "MIDDLE SCHOOLS' OR "PRIMARY SECONDARY EDUCATION' OR "SECONDARY EDUCATION' OR 'SECONDARY SCHOOLS' OR 'ELEMENTARY SCHOOLS' OR 'PRIMARY EDUCATION' OR 'JUNIOR SCHOOLS' OR 'PRIMARY SCHOOLS') AND

('SCHOOL EFFECTIVENESS') AND:

Year of Publication=('1990' OR…..'2008')

AEI (Dialog)

AEI Subject Headings=('SECONDARY EDUCATION' OR 'ELEMENTARY SCHOOLS' OR 'JUNIOR PRIMARY SCHOOLS' OR 'PRIMARY EDUCATION' OR 'PRIMARY GRADES' OR 'PRIMARY SECONDARY EDUCATION' OR "CENTRAL SCHOOLS' OR "LOWER PRIMARY YEARS' OR "MIDDLE PRIMARY YEARS' OR "PRIMARY SCHOOLS' OR "UPPER PRIMARY YEARS' OR "YEAR 1' OR "YEAR 2' OR "YEAR 3' OR "YEAR 4' OR "YEAR 5' OR "YEAR 6' OR "YEAR 7' OR "YEAR 8' OR "YEAR 9' OR "YEAR 10' OR "HIGH SCHOOLS' OR "SECONDARY SCHOOLS' OR LOWER SECONDARY YEARS' OR ELEMENTARY SCHOOLS') AND

YEAR OF PUBLICATION=("2008' OR "2007' OR "2006' ….. "1990') AND

AEI subjects headings=("SCHOOL EFFECTIVENESS' OR "EFFECTIVE SCHOOLS PROJECTS' OR "EFFECTIVE SCHOOL RESEARCH')

Psychinfo (CSA)

(DE=('elementary schools' or 'high schools' or 'junior high schools' or 'middle schools')) and ("effective* school*' or "school* effective*')

Limited to: Publication Year: 1990 -2008

ERIC (CSA)

((DE='effective schools research') or (DE='school effectiveness')) AND (PT=(142 reports: evaluative) or PT=(143 reports: research))

Limited to:

Publication year 1990-2008

And

Limited to:

Education level:

Elementary education or elementary secondary education or grade1 or grade 2 or grade 3 or grade 4 or grade 5 or grade 6 or grade 7 or grade 8 or grade 9 or grade 10 or high schools or intermediate grades or junior high schools or middle schools or primary education or secondary education

Evidensbasen

Dk=37.3? and (ti=school? Eller ti=skol?)

Sociological abstracts (CSA)

Sociological abstracts searched 2008- 11-25

(DE=('schools' or 'elementary schools' or 'private schools' or 'public schools' or 'secondary schools')) and((DE='effectiveness') or('effective* school*' or 'school* effective*'))

FIS-Bildung

(Titelsuche: schul* ODER school) UND

(Slagwörter suche: Effizienz ODER effektivitaet) UND

(Jahr:>=1990)

CBCA education (Proquest)

Effective* W/2 school*

Limited to 1990-2008

Limited to scholarly journals

Dansk pædagogisk base

DK=37.3? and (skoleeffektivitet eller effektiv? eller 'god skole') and år=1990 til 2008

Forskningsdatabasen.dk

'god? skole?'=skoleeffektivitet='effektiv? skole?' FR:1990 TO:2008

Libris (Svensk bogfortegnelse)

(skol* SAME effektiv* OR skol* SAME bra) AND tree:em AND(Prod:NB NOT (styp:n OR styp:p)) AND (ÅR:1990 OR ÅR:1991 OR ÅR:1992 OR ÅR:1993 OR ÅR:1994 OR ÅR:1995 OR ÅR:1996 OR ÅR:1997 OR ÅR:1998 OR ÅR:1999 OR ÅR:2000 OR ÅR:2001 OR ÅR:2002 OR ÅR:2003 OR ÅR:2004 OR ÅR:2005 OR ÅR:2006 OR ÅR:2007 OR ÅR:2008)

Skolporten.com

Under 'Forskning & utvickling'

Under 'Avhandlingar'

Browsing of all titles

Norbok

(DEWEY SØK: 3?0 OR 37? OR 370.193?) AND

(ORDSØK: bra OR god? OR effektiv?) AND

(ORDSØK: skol?) AND

Publication Year: 1990 - 2008

BIBSYS Forskdok

(tittel, ordsøk = effektiv? or tittel, ordsøk = bra or tittel, ordsøk = god?) and tittel, ordsøk = skol? and årstall = 1990-2008

Jykdok

('school? effectiv?')[in Kaikki sanat/All fields] OR ('effectiv? school?')[in Kaikki sanat/All fields] OR (skol? AND effektiv)[in Kaikki sanat/All fields]

With search limits:

Place of publication: Finland AND

Year of publication: 1990-2008

Swetswise

(Within all fields: effective* schools* OR

Within all fields: school* effective*) And

Publication Year: 2008 And

Within subject category: Education

This base was searched only to obtain references that were not yet available in the other bibliographic sources listed above.

Google Scholar

Limited to: the social sciences, art and humanities

Limited to: published in 2008

alleititel: (school OR schools) (good OR excellent) OR (effective OR effectiveness)

This search was also performed for the same reason as for Swetswise.

2.4 Screening

The searches were performed in such a way as to ensure that all relevant material would be found. However, not all that is found may be relevant to the study. All 3651 [3682] hits were therefore screened, and sorted according to their relevance.

The screening gave no weighting to research quality or the quality of the way in which the study was carried out and reported. Attention was given solely to whether the material belonged in the conceptual universe described above in Section 2.2.

The screening process also looked at whether the reference reported primary research. Popular presentations, secondary research reporting and discussions of scientific methodology etc. were not included.

Reasons for inclusion/exclusion	Reason described	Number
EXCLUDE wrong scope	**Not dealing with the relation between factors in schools analyzed explicitly as contributing to school effectiveness and positive effects on pupils**	**2219**
EXCLUDE Wrong paper	**Not a paper with data from empirical research: editorials, commentaries, book reviews, policy documents, resources, guides, manuals, bibliographies, opinion papers, theoretical papers, philosophical papers, research methodology papers**	**706**
EXCLUDE Wrong research	**Not offering data from original research i.e. only summarizing research done by others. (Systematic reviews can be included)**	**156**
EXCLUDE Wrong research design	**When none of these three criteria are part of the study design: 1. Control is present for differences in pupils' socioeconomic background 2. Control is present for differences in pupils' scholastic aptitude 3. A pre(-post) is present. When one criterion is found the study must be included.**	**52**
EXCLUDE Wrong institution	**Not an ordinary general primary or lower secondary school. For example special schools or vocational schools or educational institutions which function at other levels.**	**117**
EXCLUDE Wrong social context of schooling	**The document only deals with schooling in developing countries.**	**117**
MARKER Insufficient information at present	**New information is necessary in order to exclude/include**	
MARKER Overview	**A document which provides historic or conceptual overview of the review theme**	**167**
INCLUDE Inclusion	**Original empirical research on 'effective schools' which deals with ordinary primary and lower secondary schools in industrialized nations published after 1990 with a proper research design (pupils' socioeconomic background or scholastic aptitude are controlled for or with a pre (-post) test) OR Systematic reviews on 'effective schools'**	**148**

Table 2.2: Overview of complete screening

Prior to the screening process all duplicates were eliminated. As a natural consequence of the search process, duplicates must be expected to occur. 165 duplicates were removed. After this, the screening was carried out as a two-phased process:

2.4.1 Phase 1: Screening of references

All references obtained were loaded into EPPI-Reviewer and were screened for inclusion using title and abstract. The results of the screening process can be seen in Table 2.2.

After removal of duplicates, all the hits uploaded to EPPI-Reviewer were sorted into 11 categories. All references for which the information was deemed inadequate were regularly subjected to additional searches in order to supplement with abstract or other additional information. This lack of information applied in particular to Nordic references.

This phase included everything that could not be excluded *with confidence*. Both 'certain' and 'uncertain' references were thus included at this stage. Only references with a high degree of certainty were excluded.

Exclusion was hierarchical, such that exclusion took place firstly on the grounds of 'wrong scope', then of 'wrong paper', then of 'wrong research' … etc. Since the exclusion criterion 'wrong research design' was deemed impossible to apply with certainty in the screening of references, this category was only introduced in the next phase of the screening process.

After the first screening phase there remained 353 references.

2.4.2 Phase 2: Full text screening

In Phase 2 the books, articles or reports that were the subject of all the remaining references were obtained and they were then screened on the basis of the full text.

The screening was carried out using the same criteria as in Phase 1 with the addition of the exclusion criterion 'wrong research design'. This criterion was included so as to ensure that the included studies did in fact ascribe actual positive effects to the school on the basis of some form of control.

It is important to emphasise in connection with the screening process that reports from evaluations or innovative school experiments were not excluded solely on the grounds that they report evaluations or school experiments.

It is important to remember as a general point that research quality or reporting quality was not used as a basis for inclusion/exclusion.

2.5 Coding and data extraction

The *EPPI-Centre* at the *Institute of Education, London University*, was established in 1996. It has created a generalised coding and data extraction system for educational research. This is known as the *EPPI-Centre data extraction and coding tool for education studies V2.0*. This system has been used in a shortened and edited form for all coding and data extraction in this study. It is presented as Appendix 1, p. 157, and in Chapter 3. The coding and data extraction system is an integrated part of the EPPI-reviewer.

The EPPI-reviewer was used to make a coding and data extraction of all the documents included in the study. A prerequisite for creating an overview or synthesis covering all the documents is that they are described using the same system. The principle of *tertio comparationis* is employed here. That is to say, a comparison between two

elements is made possible by introducing and comparing them with a third (common) element.

Coding and data extraction consists of answering questions about the studies in such a way that relevant data is drawn out for use in the comparison. The system is built up in sections which are subdivided into questions which in turn are subdivided into multiple choice answers. At all points it is possible to insert notes and explanatory remarks linked to the selected multiple choice answer. In terms of content, the system covers the purpose of the study, its focus with respect to policy and practice, the factors investigated in the school, the focus on pupil performance, sampling considerations, results and conclusions, design and method, quality of research and reporting. The original EPPI questions have been modified considerably, as indicated in Chap. 6: Appendix 1, in the light of the actual theme of this review.

Coding and data extraction was performed by the members of the review group in such a way that individual members were responsible for specific studies. The studies were also distributed to the scientific assistants at the Clearinghouse, who also were given responsibility for specific studies. The peer review principle was then applied systematically, and every study was examined by at least two people.

Special focus was given to ensuring the quality of the evaluation of the weight of evidence, which forms part of the coding and data extraction.

In this connection a procedure was employed to permit establishment of an 'agreed version': if there were differing opinions as to the evaluation of the four questions in the section concerning weight of evidence (cf. Chap. 6: Appendix 1, Section N, Question 11-14), a dialogue took place between the member of the review group and the staff member of the Clearinghouse, in which explicit arguments for

the differences were exchanged with a view to establishing agreement. If agreement could not be reached in this way, a third party was assigned the task of establishing an 'agreed version' on the basis of the presented arguments.

In this review differences were originally noted in connection with 105 out of 444 individual evaluations of weight of evidence (24 %). The disagreements applied to 57 of a total number of 116 studies (51 %). In connection with this review it was not necessary to employ the services of a third party in any single case.

An example of a complete coding and data extraction for one document is presented in Chapter 6.

The work of coding and data extraction provided the basis on which the research mapping could be carried out. The research mapping was performed using the analysis and reporting facilities available in the EPPI-Reviewer.

2.6 Summary of the review process

Figure 2.2 presents in graphic form the process *from search to research mapping*. The figure also indicates that a research synthesis can potentially be performed starting from the research mapping that has been carried out. Grey boxes indicate sub-processes for which Clearinghouse is mainly responsible, and white boxes indicate sub-processes for which the review group and Clearinghouse are jointly responsible.

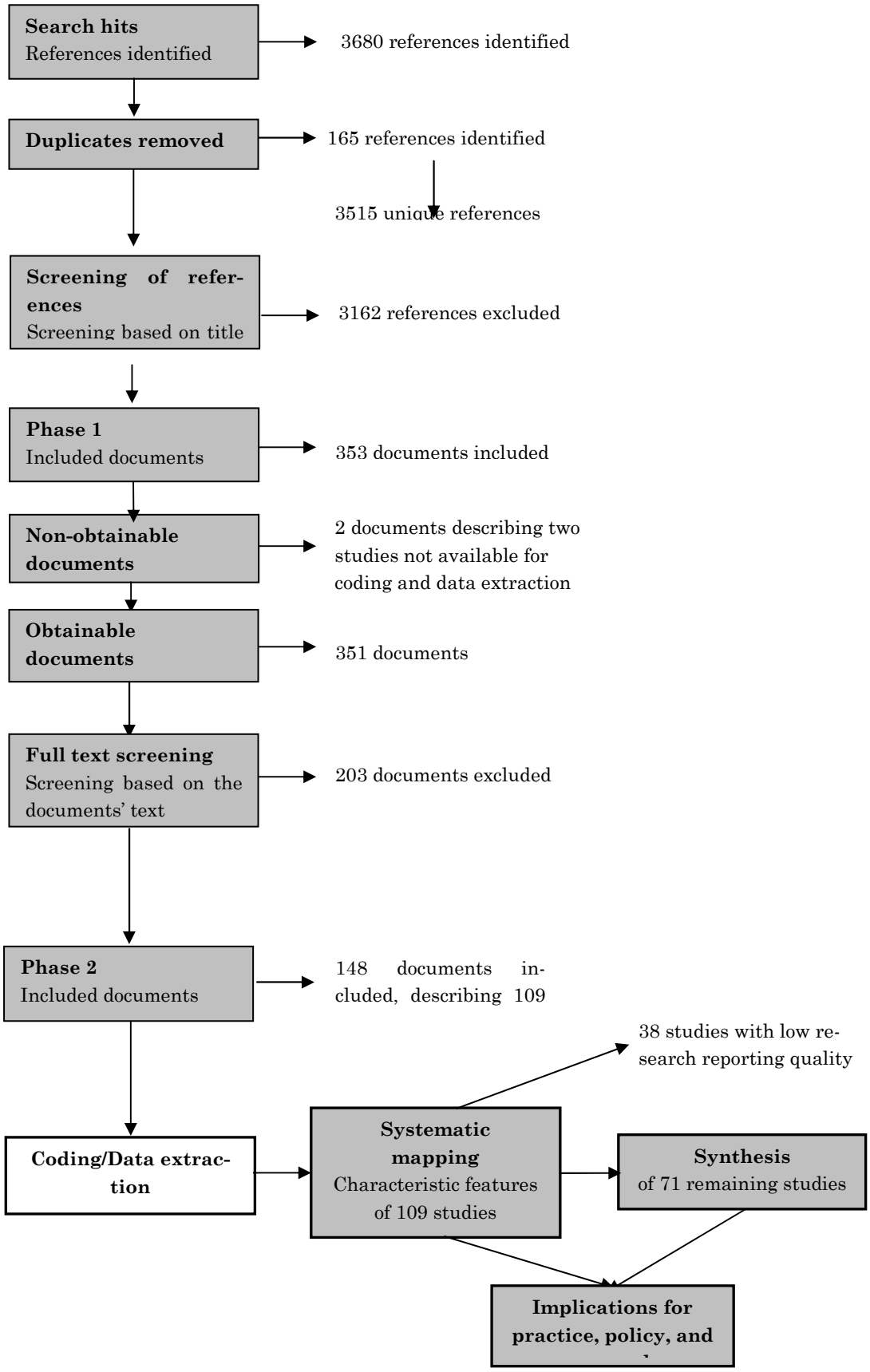

Figure 2.2: Filtering of references from search results to mapping and synthesis

3 Research mapping and research assessment

This chapter gives a general description of all the 109 studies included in the survey.

The studies are described cross-sectionally and are evaluated in the light of the research assessment, so as to create a combined picture of current research, its character and quality.

First we examine the context of the studies: Where and in what types of schools were they carried out? Next, we look at the content: What factors in the schools have been studied? Which subject areas were covered? Which pupils and what effects on pupils were looked at? Subsequently we look at the aims of the studies and their design and methodology. The chapter concludes with an analysis of the quality of the studies.

3.1 The context of the studies

School effectiveness research is an international research effort, and this manifests itself in the material in two different ways. Firstly, the 109 studies draw their data from a total of 38 different industrialised countries. Secondly, some of the investigations were in fact comparative educational studies that used data from a number of countries in one and the same research process.

Table 3.1 shows the distribution of the studies amongst the various countries. It will be seen that 64 % of all the studies involve data from the USA. UK, Holland, Australia and Belgium account for 12 %, 11 %, 9 % and 7 % of the studies respectively.

Countries	Number of studies
USA	70
United Kingdom	13
Netherlands	12
Australia	10
Belgium	8
France, Germany, Ireland, Canada, Norway	5 (from each country)
Hong Kong, Switzerland, Spain	4 (from each country)
Korea, Cyprus, Denmark, Greece, Hungary, New Zealand, Sweden, Portugal, Singapore	3 (from each country)
Austria, Czech Republic, Finland, Italy, Japan, Iceland, Thailand, Slovenia	2 (from each country)
Luxembourg, Taiwan, Slovak Republic, Russian Federation, Romania, Lithuania, Poland, Latvia	1 (from each country)

Table 3.1: Countries in which the studies took place
(N=109 studies; several categories permitted per study)

It will also be seen from the table that some studies have data from Nordic countries: Norway (5 studies), Denmark and Sweden (3 studies each) and Finland and Iceland (2 studies each). The total number of studies with Nordic data is six (Dronkers & Robert, 2008; Grøgaard, Helland & Lauglo, 2008: Martin et. al., 2000; Postlethwaite & Ross, 1992; Reynolds et al., 2002; Ringsmose & Mehlbye, 2004).

Owing to the comparative studies there are data from more than 111 countries mentioned in the 109 studies.

The reports of the studies are almost all in English (95 % of all studies), as seen in Table 3.2. Even though the searches covered a much larger linguistic universe, only a few non-English studies were found and included.

Language	Number of studies
English	104
German	3
Danish	1
Norwegian	1

Table 3.2: Language in research reports
(N=109 studies)

As already described in Chapter 2, studies that investigated schools of types not corresponding to the Nordic 'basic school' were excluded. However, this still permitted several possible types of school in the studies. The distribution amongst school types is shown in Table 3.3.

This table shows that 42 % of the studies were made in schools covering the first six school years only ('primary school'). 37 % of the studies concerned schools covering 7th to 12th grade ('secondary school'). 28 % of the studies were made in schools corresponding exactly to the Nordic type ('Primary and lower secondary'). 17 % of the studies took place in schools comprising grades 7 to 9 or 10 ('lower secondary school').

A number of studies examine both 'primary school' and 'secondary school'. This means that the number of school types indicated in the table is greater than the number of studies.

School	Number of studies
Lower secondary school	19
Primary and lower secondary school	30
Primary school	46
Secondary school	40

Table 3.3: Educational setting of the studies
(N=109 studies; several categories permitted per study)

3.2 School and pupil factors studied

The inclusion criteria for this systematic review included the requirement that a given study had to examine at least two different school factors in order to be included. The range of actually studied school factors is shown in Table 3.4.

Here the full breadth of the studies becomes obvious: there are studies covering every one of the previously defined categories of phenomena and factors in the school. The most frequently investigated factors are the socio-economic composition of the pupils at the school, school culture, teacher and leadership. These are covered by 63 %, 61 %, 58 % and 50 % of the studies respectively. Factors such as physical school environment, class size, support systems, staff development, school size and teacher teams are less frequently included in the studies, being covered by 8 %, 11 %, 15 %, 16 %, 18 % and 20 % of the studies respectively. [1]

[1] The system of school factors/phenomena was changed and improved later during the process of research synthesis. The new system for sorting school factors/phenomena is described in Chapters 4 and 7.

School factor/phenomena	Number of studies
Class size	12
Curriculum/scheduling	42
Ethnic composition of the pupils in the schools	31
Leadership	55
Management	37
Other	43
Physical environment	9
School culture	67
School size	20
Socio-economic composition of the pupils in the schools	69
Staff development	17
Support systems	16
Teacher	63
Teacher teams	22

Table 3.4: Phenomena/factor in school addressed in the studies (N=109 studies)

Even though the studies address school effectiveness as such, this may often be combined with other (school) subjects, either by viewing the school's effectiveness in relation to a subject success criterion such as performance in mathematics, or in the form of a study of special professional aspects of the school's activities, such as how reading is taught. This is shown in Table 3.5. It is seen here that 61 % of all studies have an inbuilt mathematical aspect, while 51 % look at literacy in the mother tongue. 25 % of the studies make no reference to specific factors in the curriculum.

Curriculum area	Number of studies
Cross-curricular	2
Environment	2
General	2
Geography	2
Hidden	1
History	4
Literacy - first languages	56
Literacy - further languages	7
Literature	8
Maths	66
N/A (not on a specific curriculum area)	27
Phys. Ed	1
Science	20
Other	7

Table 3.5: Curriculum area of the studies

(N=109 studies; several categories permitted per study)

The studies can also be sorted by asking whether they examine the effect of the school on the pupils in general, or its effect on specific groups of pupils. This is shown in Table 3.6. 47 % of the studies examine the effects on pupils in general, while 38 % and 23 % examine effects on pupils with low socio-economic status and pupils from ethnic groups respectively. On the other hand, there are very few studies looking at gender differences, differences in competence, and handicaps. Several studies include a number of different pupil groups. As a result, the number of studies listed under the various pupil groups is greater than the total number of studies.

Group of pupils	Number of studies
No specific group	51
Pupils with high competence	5
Pupils with low competence or handicaps	3
Yes, girls	5
Yes, boys	6
Yes, other specific groups	12
Yes, pupils from ethnic groups	25
Yes, pupils with low SES	41

Table 3.6: Pupil result focus: Specific group of pupils
(N=109 studies; several categories permitted per study)

Table 3.7 shows the distribution of studies sorted by the effect on the pupils, taken in a narrow academic context. Only 3 % of the studies made no reference at all to this aspect. 96 % include this focus and interpret it as pupil performance, usually measured by achievement or examination performance. 5 % look at the academic effect in other ways, for example as a successful transition to the next stage in the educational system. 4 of the latter studies also include performance measurements. As a result, the combined number of studies listed in the various categories is greater than the total number of studies.

Focus on academic effects	Number of studies
Without such focus	3
Yes, achievement or per-formance	105
Other academic effects	5

Table 3.7: Pupil result focus: Academic effects
(N=109 studies; several categories permitted per study)

A number of the studies include an evaluation of effects on pupils apart from the academic effect. This includes topics such as the pupils' well-being. Table 3.8 shows the distribution: 72 % of the studies did not include such effects, whilst 30 % of the studies did. A number of studies examined a number of non-academic effects on the pupils. For this reason the combined number of studies listed by effects examined is slightly greater than the total number of studies.

Focus on non-academic effects	Number of studies
No	78
Other	11
Yes, physical	4
Yes, psychical	18

Table 3.8: Pupil result focus: non-academic effects
(N=109 studies; several categories permitted per study)

3.3 Purpose, design and methodology of the studies

The following section gives a short description of the studies seen from a research viewpoint.

The aims of the various studies are listed in Table 3.9. Here it is seen that several studies have more than one purpose. For this reason, the combined number of studies listed by purpose is greater than the total number of studies. The table also shows that explorations of relationships and description are the most frequent purposes, covering 57 % and 51 % of the studies respectively. The purpose 'what works' appears in only 9 % of the studies. Methods development, here understood as research methodological development, is a purpose in 9 % of the studies.

Purpose	Number of studies
Description	**56**
Exploration of relation-ships	**62**
What works?	**10**
Methods development	**10**

Table 3.9: Purpose of the study

(N=109 studies; several categories permitted per study)

When choosing the appropriate design for a study, this question is usually linked to the purpose of the study. The designs that were actually used in the studies are listed in Table 3.10. Here again, a number of studies can be assigned to more than one category, and therefore the combined number of studies listed by design is greater than the total number of studies.

The primary impression is one of considerable breadth in choice of design. The most frequently used designs are secondary data analysis (48 %), cross-sectional studies (41 %) and studies of views (35 %). There are also a considerable number of studies using a case study design (27 %) and cohort design (28 %). Studies using an experimental or adapted experimental design are rare.

The data collection procedures in the studies were also diverse, as shown in Table 3.11. Many of the studies employed several data collection methodologies. Thus the combined number of studies listed by data collection method is greater than the total number of studies. The most frequently employed methods for data collection were: self-completion questionnaires (68 %), use of secondary, already existing data (42 %), and curriculum-based assessment or measurement (40 %). One-to-one interviews (36 %) and observation (30 %) were also employed in a considerable number of studies.

Design	Number of studies
Case study	29
Case-control study	8
Cohort study	30
Comparative study	7
Cross-sectional study	45
Document study	2
Ethnography	19
Experiment with non-random allocation to groups	5
Experiment with random allocation to groups	1
Methodological study	3
One group post-test only	1
One group pre-post test	1
Secondary data analysis	52
Views study	38

Table 3.10: Design in studies

(N=109 studies; several categories permitted per study)

Data collection method	Number of studies
Curriculum-based assessment	44
Examinations	8
Focus group interview	18
Not stated/ unclear	1
Observation	33
One-to-one interview (face to face or by phone)	39
Other documentation	16
Please specify any other important features of data collection	9
Practical test	1
Psychological test (e.g. I.Q test)	7
School/ college records (e.g. attendance records etc)	19

Data collection method	Number of studies
Secondary data such as publicly available statistics	46
Self-completion questionnaire	74
Self-completion report or diary	5

Table 3.11: Methods applied in data collection in the studies (N=109 studies)

3.4 Quality of studies

A quality assessment of the research is a necessary step in the process of establishing an overview of what the research actually shows. Only studies carried out and reported to a sufficiently high standard can be viewed with confidence. For this reason, all the studies included in this mapping have been assessed in relation to a broad range of questions concerning their quality; cf. Chapter 6: Appendix 1. For each individual study an assessment is made of the evidence that the study can provide. In Section 2.5, a description was given as to how peer review was employed in the assessment process, with at least two different persons responsible for each assessment.

Table 3.12 displays how a number of relevant factors were evaluated concerning the adequacy of the description of the study that was available in the report. Here we see that the fewest problems were encountered with regard to the description of the context and aims (11 % and 19 % respectively). The greatest number of problems was encountered in the descriptions of avoidance of selective reporting bias (39 %). In addition, 36 % of the studies would not be replicable on the basis of the description in the report. Since the material includes a considerable number of qualitative studies this situation is not especially remarkable.

Table 3.13 to Table 3.20 indicates the distributions of answers to a number of core assessments of the quality of the individual studies. These assessments, together with the assessments in Table 3.12,

serve as the basis for the weight of evidence assigned to the individual studies.

Wording of question	Yes	No
Is the context of the study adequately described?	97	12
Are the aims of the study clearly reported?	88	21
Is there an adequate description of the sample used in the study and how the sample was identified and recruited?	79	30
Is there an adequate description of the methods used in the study to collect data?	85	24
Is there an adequate description of the methods of data analysis?	80	29
Is the study replicable from this report?	70	39
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?)	66	43

Table 3.12: Quality of studies – reporting

(N=109 studies; several categories permitted per study)

Only a minority of the studies indicate problems of a research ethical nature concerning the involvement of participants or relatives of participants. Table 3.13 shows that this was only a problem in 8 (7 %) of the studies.

Answer	Number of studies
No, but involvement would be desirable	2
No, involvement is not relevant	71
Yes, however users/relatives are not appropriately involved	6
Yes, users/relatives are appropriately involved	30

Table 3.13: Were users / relatives of users involved in the design or conduct of the study?

(N=109 studies)

Moving to the choice of research design in the individual studies, it is our opinion, cf. Table 3.14, that this was only completely satisfactory in 53 % of the studies. In the remaining studies, there were consid-

ered to be major or minor problems in the design employed. Here it must also be recalled that studies that do not control for obvious alternative reasons than the school for the success of the pupil have not been included. This question is addressed in section 2.2.

Answer	Number of studies
No	51
Yes, completely	58

Table 3.14: Was the choice of research design appropriate for addressing the research question(s) posed?

(N=109 studies)

An evaluation of the attempts made by the studies to establish reliability and repeatability of data collection is presented in Table 3.15. Here, 80 % of the studies have made a good attempt or at least some form of attempt, while 20 % have made no attempt to ensure reliability and repeatability.

Answer	Number of studies
No, none	22
Yes, good	46
Yes, some attempt	41

Table 3.15: Have sufficient attempts been made to establish the repeatability or reliability of data collection methods or tools?

(N=109 studies)

The attempts made in the studies to ensure the validity of the data collection procedures are analysed in Table 3.16. 77 % have made a good attempt or some form of attempt, while 23 % of the studies have made no attempt to ensure the validity of their data collection procedures.

Answer	Number of studies
No, none	25
Yes, good	44
Yes, some attempt	40

Table 3.16: Have sufficient attempts been made to establish the validity or trustworthiness of data collection tools and methods?

(N=109 studies)

The reliability and repeatability of the data analysis has been adequately established in 61 % of the studies. Major or minor problems of data analysis were noted in 39 % of the studies, cf. Table 3.17.

Answer	Number of studies
No	43
Yes	66

Table 3.17: Have sufficient attempts been made to establish the repeatability or reliability of data analysis?

(N=109 studies)

Table 3.18 shows that 72 % of the studies have made good or some attempts to ensure the validity of their data analysis. 28 % of the studies made no apparent attempt to do this.

Answer	Number of studies
No, none	30
Yes, good	42
Yes, some attempt	37

Table 3.18: Have sufficient attempts been made to establish the validity or trustworthiness of data analysis?

(N=109 studies)

In each study a choice has been made of research design and methodology. Table 3.19 indicates whether the chosen design and methods have been capable of ruling out other explanations than the one arrived at in the study itself. 25 % of the studies were found to be designed in such a way that they could rule out alternative explana-

tions to a great extent. 49 % of the studies were found to be designed so that they could rule out alternative explanations to a limited extent. 26 % of the studies were not capable of ruling out alternative explanations at all.

Answer	Number of studies
A little	53
A lot	27
Not at all	29

Table 3.19: To what extent are the research design and methods employed able to rule out any other sources of error/bias which would lead to alternative explanations for the findings of the study?

(N=109 studies)

Table 3.20 examines whether the authors of this review arrived at different findings and conclusions from the authors of the studies in question. This was found to be the case in 34 % of the studies, while for 66 % of the studies the reviewers were in agreement with the authors.

Answer	Number of studies
Not applicable (no difference in conclusions)	72
Yes	37

Table 3.20: In light of the above, do the reviewers differ from the authors over the findings or conclusions of the study?

(N=109 studies)

The combined assessment of the contributions of the individual studies to the weight of evidence is shown in Table 3.21.

Question	Number of studies		
	High	medium	Low
Weight of evidence A: Taking account of all quality assessment issues, can the study findings be trusted in answering the study question(s)?	38	40	31
Weight of evidence B: Appropriateness of research design and analysis for addressing the question, or sub-questions, of this specific systematic review.	29	48	32
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	46	63	-
Weight of evidence D: Overall weight of evidence	17	54	38

Table 3.21: Weight of evidence of the studies (N=109 studies)

Weight of Evidence A indicates whether the individual study has been carried out in good agreement with its own declared aims, design, methods and results. It is a combined result based on how the study has been scored in all the assessments presented in Table 3.12 to Table 3.20. The distribution turns out to be fairly even: 35 % of the studies have a high weight of evidence, 37 % have a medium weight of evidence and 28 % have a low weight of evidence.

Weight of evidence B indicates whether the design employed by the individual study has been appropriate for providing an answer to the review question on which this mapping and synthesis is based. Here the studies are distributed with 27 % in the high category, 44 % in the medium category and 29 % in the low category.

Every study has its own focus and its own way of viewing phenomena and their context. *Weight of evidence C* addresses the relevance of each study's focus with respect to the review question of this research assessment. Here 42 % of the studies are found to have a high weight of evidence and 58 % a medium weight of evidence. Low weight of evidence is not included as an optional response category,

since studies with such low weight of evidence were removed in the screening process, cf. Section 2.4.

The position of the studies with respect to *weight of evidence D*, the combined weight of evidence, decides whether the studies should be included in a research synthesis covering the results emerging from the research within this field. Studies with a high weight of evidence D (16 %) and medium weight of evidence D (49 %) qualify for inclusion in the synthesis. Studies with low weight of evidence D (35 %) should not be included in the synthesis.[2]

[2] The six studies with data from Nordic countries, cf. p. 36, have the following overall evidence weight: *Medium* (Dronkers & Robert, 2008; Grøgaard, Helland & Lauglo, 2008: Martin et. al., 2000; Postlethwaite & Ross, 1992; Ringsmose & Mehlbye, 2004); *Low* (Reynolds et al., 2002). Only the first five studies are included in the synthesis. The school factors addressed in the Nordic studies are recorded in Chap. 7 under each factor.

4 Syntheses of primary research

4.1 Introductory remarks

The previous chapter identified the primary studies included in the systematic synthesis process. We have already seen (cf. Table 3.10) that only one randomised controlled experiment about the review questions relating to this study was carried out in the years 1990-2008. This excludes the possibility of conducting a systematic synthesis in the form of meta-analyses, but some sort of result could still be obtained, however, from the quantitative data in the primary studies; cf. Chap. 8: Appendix 3, and Section 4.3.1.

As an addition to these analyses we apply in the following a procedure called *Narrative Synthesis in Systematic Reviews* (see Popay et al., 2006). The aim of this analysis is both to get closer to fundamental features of the 'good school' and to uncover some of the mechanisms that cause and explain the outcomes from 'good schools'.

According to this procedure, the narrative synthesis process consists of four elements, which analytically are presented in a given order, but which in the practical process of synthesis might well contain iterative movements between the various elements.

The four elements can briefly be described as follows:

The first element consists of developing a theoretical model of how the effect(s) that are the object of study come about, why they do so and for whom. There is at times talk of establishing a 'theory of change' (see Weiss, 1998, 55), which in Wholey's (1987, 78) description suggests 'the chain of causal assumptions that link programme resources, activities, intermediate outcomes and ultimate goals'. The theory can be used to interpret the review's findings and can be use-

ful in an assessment of how broad the applicability of these findings is.

The second element consists of developing a preliminary synthesis. In this phase it is necessary to organise the studies included in such a way that their direction — and if possible their strength — can be established. At the same time a pattern is sought that also relates to factors that in various ways might prove to have an influence on the effect. In this phase the task is to establish possible syntheses, while it is reserved for a later stage to determine how robust they are.

The third element is devoted to a survey of the factors that are common to the studies and can explain variations in the direction and strength of the effect studied. Also included here is a treatment of the question of why a phenomenon has or does not have an effect, and of whether particular factors play a part here that can explain how the effect in a given context is strengthened or weakened.

The fourth element is an assessment of the robustness of the synthesis. This is a complex notion which, somewhat simplified, can be said to consist of three aspects.

In the first place the robustness of a synthesis depends on *the methodological quality* of the primary studies. The trustworthiness of a synthesis will depend both on this quality and on the quantity of the evidential basis it is constructed upon. If primary studies of poor quality are uncritically included in the systematic review, the trustworthiness of the synthesis will be affected.

In the second place, the trustworthiness of the syntheses will also be affected by the *methods* used in the synthesis. Which precautions are taken to minimise bias by, for example, giving a similar weighting to primary studies of uniform technical quality?

Finally one aspect deals with the degree to which the screener and the reviewer have *sufficient information* to be able to be certain in including a primary study in the synthesis. This can present a serious problem, in particular as regards the investigation of effects connected to complex factors, since it is not always clear from the primary study what the conditions are that the various effects are linked to.

At the conclusion of the synthesis process, these aspects should be brought together to produce an overall assessment of the strength of the evidence, which allows conclusions to be drawn on the basis of a narrative synthesis.

4.2 A theoretical model

On page 23, the purpose of this systematic review is formulated as follows**:**

> *What empirical research has been carried out to examine the relationship between factors in primary and lower secondary schools (inputs and processes) and the learning achieved by primary and lower secondary school pupils (outputs and outcomes)?*
>
> *What are the results with weight of evidence of this empirical research?*

The research mapping dealt with in Chapter 3 entailed a list of school factors deduced from the data extraction from primary studies with high or medium weight. *The school factors and subcategories developed are thus based exclusively on the factors and categories addressed in the included primary studies.* This list of factors is reproduced in Table 4.1. The list has to be seen in connection with the more comprehensive overview over the individual school factors and the possible subcategories reproduced in Chap. 7, Appendix 2. Clear-

inghouse has developed the list and the overview, and has discussed both with the Review Group.

According to the procedure of a narrative synthesis, the first task consists of developing a theoretical model of how the effect(s) that are the object of study come about, why they do so and for whom. That is, to develop a model that connects school factors to specified groups of pupils by identifying the activities that the primary studies have demonstrated to cause the desired achievements.

The list of school factors identified in the data extraction process normally forms part of a theoretical context that describes relationships between pupils, teachers, teaching-learning, schools, etc. In the European tradition this theoretical field has been named '*Didaktik*' from the German. A 'didactic theory' is a theory that among other things outlines the internal relationships between these school factors and points to the pedagogical activities going on between pupils and teacher, the teacher activities, the processes in the classroom, and processes in the school and its environment. To see these school factors in their proper educational context we therefore relate them in the following to features of a *critical school didaktik* theory (CSD), originally developed in Uljens (1997).

School Factor	Subcategories
1.School size	-
2.Class size	-
3.Management and Leadership	3.1 Human resources
	3.2 Rational goal leadership
	3.3 Educational leadership
	3.4 Administrational leadership
	3.5 Other
4. Curriculum/scheduling	4.1 Opportunity to learn
	4.2 Alignment
	4.3 Learning goals
	4.4 Other
5. School culture and climate	5.1 Disciplinary climate
	5.2 Achievement/progress orientation
	5.3 Interrelational climate
	5.4 Social norms and values
	5.5 Other
6. Teacher	6.1 Teacher behaviour
	6.2 Teacher beliefs
	6.3 Teacher self-efficacy beliefs
	6.4 Teacher subject knowledge
	6.5 Teacher as an organisational actor
7. Support teams	-
8. Physical environment	-
9. Pupil composition of the school	-
10. Parental Relationship	-
11.Other	-

Table 4.1: School Factors and Subcategories

The elements of the theory are discussed with reference to the model presented in Figure 4.1.

Figure 4.1 indicates that the model consists of four commonly accepted main components in understanding institutionalized schooling. These components refer to (1) the planning, (2) realisation and (3) evaluation of the pedagogical process, as well as to (4) the multiple contexts providing the framework for the pedagogical activity in schools. The model thus identifies the constitutive elements of the teachers' pedagogical work (planning, teaching and evaluation). Con-

cerning the context, major distinctions are made between (1) the classrooms (learning situation), (2) the school as an organisational context and (3) the local society and culture as framing the pedagogical work on the first two levels.

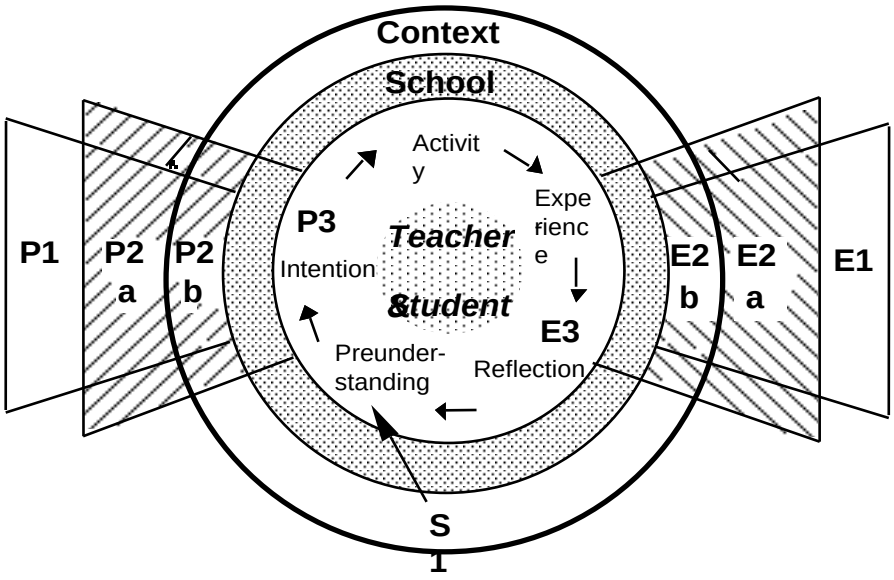

Figure 4.1: Levels and forms of pedagogical activity according to Critical Theory of School Didaktik

S1 = Pupils' pre-understanding, intentions and experiences (life-history) in approaching the school

P1 = Planning (intentions) at a (formal) collective level

P2 (a) = Teachers' planning before a pedagogical sequence in relation to planning at a collective level

P2 (b) = Teachers' planning before a pedagogical sequence in relation to the individual, local culture and the school as context

P3 = Teachers' and pupils' ongoing intentional planning

E3 = Teachers' and pupils' ongoing evaluative reflection of their teaching- and learning experiences

E2 (b) = Teachers' evaluation of process and results after a pedagogical sequence in relation to the individual, local culture and the school as context

E2 (a) = Teachers' evaluation after a pedagogical sequence in relation to curriculum and evaluation at a collective level

E1 = Evaluation at a formal, collective level

School = The classroom and local school as context

Context = Non–formal cultural context of education

- The inner circle describes the dynamic developmental process of teachers' and pupils' individual and shared intentions, activities, experiences, reflection and their situated teaching- and learning–experiences
- The outer circle indicates the various contexts framing the Teaching-Studying-Learning process in the school

The left wing consists of two parts, P1 and P2. P1 is planning at the collective level (e.g. national curriculum). P2 refers to the teacher's planning activities before the actual process. Finally, we must differentiate between the teacher's preparatory planning (P2) and the teacher's ongoing planning (P3), which goes right into the heart of the model.

A similar differentiation to the one concerned with planning (P) can be made with respect to evaluative activities (E). There is the ongoing evaluation during the teaching process (E3), as well as the teacher's evaluation after some kind of pedagogical sequence (E2). Evaluation after a sequence covers both evaluation of the pupils' results in relation to the goals as well as the teacher's evaluation of his/her own activities such as the choice of relevant content, form of representation etc. Finally, E1 refers to evaluation at the collective level (e.g. national evaluation, IEA, PISA, etc.).

Finally, the model elaborates the pedagogical process itself in many important details, which are not dealt with here.

It is important to notice that the present model works on three principally different but complementary levels; the collective, the individual and the interactional levels. Questions related to choice and treatments of subject matter are dealt with on all three levels, but in different ways.

To conclude, it should be observed that this theory of didactics accepts and defends an interdependence concerning decisions in the practical pedagogical situation with respect to content, method, media, goals and context/pupil. The approach is not method-centred on

content only, nor mainly centred at the aims or media of education. The model emphasizes that every pedagogical situation must be analyzed with respect to all these dimensions.

As emphasized several times, the primary studies in this systematic review were selected from the school effectiveness research tradition. Therefore, a second model is introduced, i.e. an integrated model of school effectiveness, originally developed in Scheerens (2000). By doing this, we are pointing out the similarities between the didactic theory and the integrated model of school effectiveness, and – in the syntheses below – are thus relating the results from school effectiveness research to the pedagogical activities going on in classrooms and schools.

Whereas Uljens's model in a general way places the identified school factors in a school didactic context, Scheerens's (2000, chap. 2) model is created as an integrated model with basis in research linked to the paradigm of school effectiveness. The structure of the model emerges from the integrated model of school effectiveness (ibid.), cf. Figure 4.2.

As Scheerens says in connection with this model, 'the fundamental design of school-effectiveness research [is] the association of hypothetical effectiveness-enhancing conditions and measures of output, usually calculated in terms of pupil achievement. The basic model can be taken from systems theory, where the school is seen as a black box, within which processes or 'throughput' takes place to transform this basic design. The inclusion of an environmental or contextual dimension completes this model. The major task of school effectiveness research is to reveal the impact of relevant input characteristics on output and to 'break open' the black box in order to show which process or throughput factors 'work', as well as the impact of contextual conditions. Within the school it is helpful to distinguish between

school and classroom levels and the corresponding school organisational and instructional processes' (Scheerens, 2000, 35).

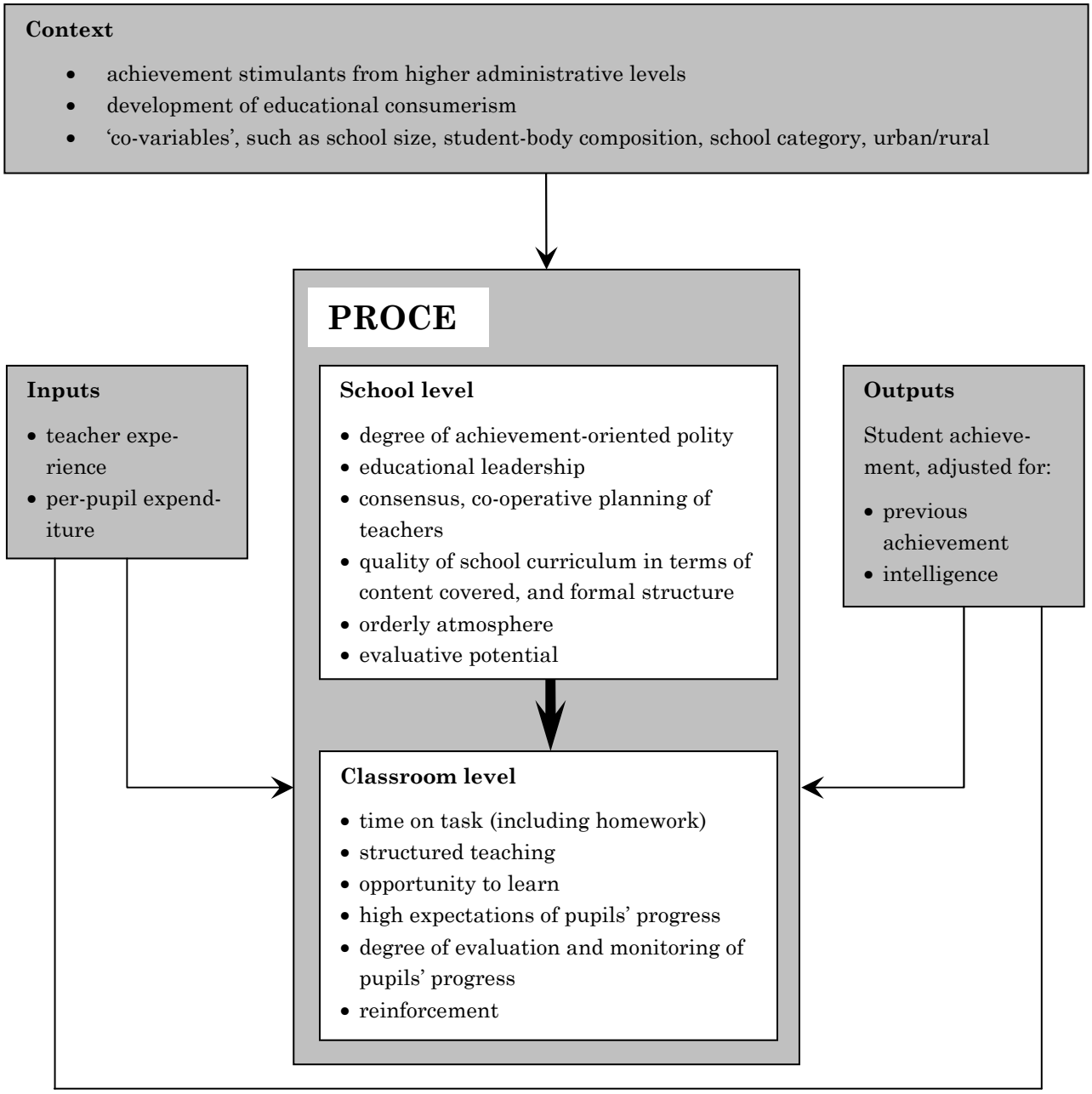

Figure 4.2: An integrated model of school effectiveness

(*Source*: J. Scheerens, 2000)

In opposition to the basic systems model of school functioning, the Integrated Model of School Effectiveness expounds the content of the

five boxes in the figure: context, inputs, outputs, school level and classroom level, cf. Scheerens (2000, fig. 3, p. 54). This will not be discussed any further here.

If we compare Uljens's critical theory of School *Didaktik* with Scheerens's integrated model of School Effectiveness it appears that the models have a lot in common. The main difference is that Uljens accentuates the teachers' pedagogical activities and the teacher-pupil interaction in and out of the classroom more than does Scheerens's systems-theoretical approach. It appears likewise that the two models refer to school factors of the same character, as mentioned in Table 4.1. Uljens's model places the ten school factors and subcategories in a didactic context, whereas Scheerens's model calls attention to fundamental results of school effectiveness research that comprise the ten school factors and subcategories mentioned in Table 4.1.

4.3 Syntheses based on the theoretical model

Having advanced a model for the synthesis we turn now to the next task: Developing a preliminary synthesis, cf. p. 66. In this phase it is necessary to organise the included studies in such a way that their direction — and if possible their strength — can be established. At the same time a pattern is sought that also relates to factors that in various ways might prove to have an influence on the effect. In this phase the task is to establish possible syntheses, while it is reserved for a later stage to determine how robust they are.

The primary studies included in the synthesis are divided into two main groups: One group comprises 63 primary studies that apply quantitative methods. The second group comprises 16 primary studies that apply qualitative methods. The reason for this division is the following: several studies apply both quantitative and qualitative methods. Studies applying qualitative methods in part of the study do not necessarily have the same weight of evidence here as they

were given in the data extraction process. This assessment has the consequence that five studies, which in the research mapping were assigned high or medium weight of evidence, have been excluded in this group with the assessment: low weight of evidence. This applies to the following studies (Perez et al., 2007; Teddlie, 1993; Stringfield, 1993; Witte & Walsh, 1990; Young, 2001).

Having carried out the synthesis for the two groups separately, the final task is to consider the possibility of synthesising the results from the two groups in one analysis.

4.3.1 Group 1: Synthesis of quantitative studies

This group includes all studies that were recorded in the research mapping as being quantitative, as having high or medium weight of evidence, and as having a measure of significance. Group 1 comprises 16 studies with high weight of evidence and 39 studies with medium weight of evidence.

To supplement the narrative syntheses we apply an additional procedure in this section. As will be seen from the tables (i.e. Table 4.12-Table 4.21 & Table 4.26), for every school factor/subcategories it is recorded how many primary studies have been proven to have significance relative to the school factor/subcategory in question, or are insignificant or intractable in that respect. In the tables in the section it is also recorded whether these circumstances are found in primary studies with high or medium weight of evidence. However, in the syntheses that were performed, this last distinction was disregarded. Then the number of significant primary studies for each school factor/subcategory is related to the sum of all studies with high or medium weight of evidence that have addressed the school factor/subcategory in question.

The power calculation table in Table 9.1, Chap. 9, Appendix 4, is applied in the assessment of whether a synthesis is feasible or not. [1] The calculation is done as calculations of power in the light of a binomial distribution with the parameter n=10,11,12,13,...,100, where a hypothesis H_0: ss=0.05 is tested in comparison with an alternative H_1: ss=0.10 0.15 0.20. The figure 0.05 is chosen since this is the general level of significance in the studies analysed.

The presentation is organised as follows. First, all the studies having a similar idea of the 'good school' are collected together (Pupil Achievement). These studies are then distributed into Pupil Groups in keeping with the categories applied in the data extraction system. Finally, for each of the created groups, school factors and school activities are recorded according to Table 4.1. This is done by enumerating the number of significant (positive and negative) and insignificant studies for each school factor and subcategory, respectively. Studies without measurements of significance are not included.

4.3.1.1 What is a 'good school'?

16 studies having high weight of evidence in this group focus on Pupil Achievement as academic achievement. Table 4.2 shows on which subjects the estimation of Pupil Achievements as academic achievement is based.

[1] The table was developed by Peter Allerup, Professor of Statistics at the Department of Education, Aarhus University.

School subjects / Studies	Reading	Mathematics	Language	Writing	Citizenship	Science
Campbell	x	x				
Meelissen		x				
Opdenakker, 2007						
Rogers		x	x			
Rumberger & Palardy	x	x				
Tarter	x	x				
Van Damme		x	x			
Yu	x	x		x	x	x
Taylor	x					
Reezigt		x	x			
Van der Werf, 1997		x				
Van der Werf & Weide, 1996	x					
Teddlie						
Young, 1992						x
Ross et al., 2006		x		x		
Woessmann		x				x
Sum	**5**	**11**	**3**	**2**	**1**	**3**

Table 4.2: Which subjects are covered when measuring Pupil Achievements as academic achievement? (high weight of evidence studies)

(N = 16)

Four of the studies having high weight of evidence in this group focus additionally on Pupil Achievement as non-academic achievement of various types, see Table 4.3.

Studies	Area
Meelissen	- self confidence in maths - liking maths - stereotyped views
Opdenakker, 1997	- the learning efforts of students
Rumberger & Palardy	- drop out - transfer - attrition
Van Damme	- environment - work - self - peers

Table 4.3: Which topics are covered when measuring Pupil Achievements as non-academic achievement? (high weight of evidence studies)

Table 4.3 shows in which area and/or non-academic aspects the estimation of Pupil Achievements as non-academic achievement is based.

39 studies having medium weight of evidence focus on Pupil Achievement as academic achievement. Table 4.4 shows on which subjects the estimation of Pupil Achievements as academic achievement is based.

Studies \ School subjects	Mathematics	Reading	Science	Subject(s) not specified	Language	Writing	Social Science	History	Citizenship	Physical	Education
Bain	X	X									
Bondi		X									
Bottoms, 2006	X	X	X								
Choi	X										
Coates	X	X	X			X	X				
Dumay	X										
Florida	X	X									
Foley	X	X									
Franklin	X	X	X		X		X				
Fullarton	X										
Griffith, 2002				X							
Griffith, 2003				X							
Grisay	X	X									
Grøgaard		X			X	X					
Heck	X	X									
Hofman, 1996	X	X									
Hofman, 2002	X										
Hoy	X	X									
Kennedy	X	X			X						
Kyriakides										X	
Lamb	X										
Mandeville	X										
Martin	X		X								
Meijnen	X	X									
Opdenakker, 2000	X	X									
Papanastasiou	X										
Perez	X	X									
Ringsmose	X	X			X						
Sammons	X	X	X		X			X			
Senkbeil	X		X								
Smyth				X							
Sweetland	X	X									
Thomas, 1995	X	X		X							
Traufler	X	X	X								
Waxman				X							
Webster	X										
Witte	X	X									
Young, 2001	X		X								
Zigarelli	X	X	X					X	X		
Sum	32	23	9	5	5	2	2	2	1	1	

Table 4.4: Which subjects are covered when measuring Pupil Achievements as academic achievement? (medium weight of evidence studies)

Seven of the studies having medium weight of evidence in this group focus additionally on Pupil Achievement as non-academic achievement of various types.

Studies	Area
Bain	- student self concept - motivation
Foley	- dropout rates
Kyriakides	- psychomotor skills
Opdenakker, 2000	- well-being-indicators
Silins	- participation - engagement
Smyth	- absenteeism - potential drop out - stress - academic self-image - locus of control - body image
Witte	- dropout rates

Table 4.5: Which topics are covered when measuring Pupil Achievements as non-academic achievement? (medium weight of evidence studies)

(N=7)

Table 4.5 shows in which area and/or non-academic aspects the estimation of Pupil Achievement as non-academic achievement is based.

In consequence, the 55 studies having high or medium weight of evidence employ two different definitions of the 'good school', firstly, a school with 'high academic achievements' and secondly a school with 'high non-academic achievements'. However, since the two definitions relate to the same studies it would seem reasonable to suppose that the two definitions should not be seen as mutually exclusive, but rather as supplementing each other. We can therefore say that in 44 studies Pupil Achievements are defined solely as 'high academic

achievements', and in 11 studies Pupil Achievements are defined as 'both high academic and non-academic achievements'.

In the following presentation we discuss the 44 and the 11 studies in separate groups. This is because we cannot know in advance whether those Pupil Activities that lead to high academic achievements are the same as those that lead to high non-academic achievements. This is what the synthesis is intended to show.

As shown in Table 4.2, out of 16 high weight of evidence studies that look at Pupil Achievement as high academic achievement, 11 studies cover Maths, 5 cover Reading, 3 Language and 2 Writing, 3 Science, and 1 the subject 'Citizenship'. Two studies give no clear indication of subjects covered. It is a tradition to consider the 3R's (Reading, wRiting, and aRithmetic) as important indicators of Pupil Achievement. The studies included here reflect this viewpoint.

Table 4.3 shows that amongst the 4 high weight of evidence studies looking at Pupil Achievement as non-academic achievement it can be difficult to find any common criteria.

As shown in Table 4.5, out of 39 studies with medium weight of evidence that look at Pupil Achievement as high academic achievement, 32 studies cover Maths, 23 cover Reading, 9 cover Science, 5 give no clear indication of subjects covered, 5 cover Language and 2 Writing, 2 Social Science, 2 History, and 1 the subjects 'Citizenship' and Physical Education, respectively. As mentioned above, it is a tradition to consider the 3R's (Reading, wRiting, and aRithmetic) as important indicators of Pupil Achievement. These studies also reflect this viewpoint.

Table 4.5 shows that amongst the seven medium weight of evidence studies looking at Pupil Achievement as non-academic achievement it can be difficult to find any common criteria.

4.3.1.2 A good school – for whom?

The included studies specify for which Pupil Groups the 'good school' is good.

Taking first the 16 studies with high weight of evidence that define a 'good school' as a school with high academic achievement, Table 4.6 indicates for which Pupil Groups the 'good school' is good.

13 studies with high weight of evidence do not specify Pupil Groups, but three studies look exclusively at Gender, Low SES, and Ethnic groups, respectively.

Looking next at the four high weight of evidence studies that define a 'good school' as a school with high non-academic achievements, Table 4.7 shows for which Pupil Groups the 'good school' is good.

For Whom? / Studies	Pupils with no specified SES and/or gender and ethnicity	Gender	Low-SES	Ethnic groups
Campbell	X			
Meelissen		X		
Opdenakker, 2007	X			
Rogers	X			
Rumberger & Palardy	X			
Tarter	X			
Van Damme	X			
Yu	X			
Taylor			X	
Reezigt	X			
Van der Werf, 1997	X			
Van der Werf & Weide, 1996				X
Teddlie	X			
Young, 1992	X			
Ross et al., 2006	X			
Woessmann	X			
Sum	13	1	1	1

Table 4.6: The 'good school' for whom? (Academic Achievements; high weight of evidence studies)

(N = 16)

For Whom? / Studies	Pupils with no specified SES and/or gender and/or ethnicity	Gender
Meelissen		X
Opdenakker, 2007	X	
Rumberger & Palardy	X	
Van Damme	X	
Sum	3	1

Table 4.7: The 'good school' for whom? (Non-academic Achievements; high weight of evidence studies)

(N = 4)

4.3.1.3 The good school – how?

This section discusses the results of the studies with respect to which School Factors and Activities are selected from the possibilities considered as being important for high Pupil Achievements.

Studies \ For Whom?	Pupil with no specified SES and/or gender and ethnicity	Low SES	Middle SES	High SES	Gender	Ethnic groups
Bain		x				
Bondi	x					
Bottoms, 2006		x				
Choi	x					
Coates	x					
Dumay	x					
Florida		x				
Foley	x					
Franklin		x	x	x		
Fullarton	x					
Griffith, 2002		x				x
Griffith, 2003	x					
Grisay		x				x
Grøgaard	x					
Heck	x					
Hofman, 1996	x					
Hofman et al., 2002	x					
Hoy	x					
Kennedy	x					
Kyriakides	x					
Lamb, 2002	x					
Mandeville		x			x	x
Martin	x					
Meijnen et al.		x	x	x	x	x
Opdenakker, 2000	x					
Papanastasiou	x					
Perez	x					
Ringsmose	x					
Sammons		x			x	x
Senkbeil	x	x				
Smyth	x					
Sweetland	x					
Thomas, 1995	x					
Traufler		x				x
Waxman	X					
Webster	x					
Witte	x					
Young, 2001	x					
Zigarelli	x					
Sum	29	11	2	2	3	6

Table 4.8: The 'good school' for whom? (Academic Achievements; medium weight of evidence studies)

(N=39)

The presentation is structured as follows: starting with the Pupil Groups shown in Table 4.6 and Table 4.8 we look first at the studies that define the 'good school' on the basis of high academic achievement in each of the School Factors and related Activities mentioned in the study in question.

For whom? / Studies	Pupils with no specified SES and/or gender and ethnicity	Low-SES	Middle SES	High SES	Gender	Ethnic groups
Bain		X				
Foley	X					
Kyriakides	X					
Opdenakker, 2000	X					
Silins	X					
Smyth	X					
Witte	X					
Sum	**6**	**1**	**0**	**0**	**0**	**0**

Table 4.9: The 'good school' for whom? (Non-academic Achievements; medium weight of evidence studies)

(N=7)

Following this we take the Pupil Groups from Table 4.7 and Table 4.9 as starting point for looking at those studies that define the 'good school' on the basis of high non-academic achievement, here too taking each of the School Factors and related Activities mentioned in the studies.

4.3.1.3.1 The 'good school': academic achievement, for pupils with no specified SES

The studies grouped together here all look into School Factors and Activities of importance for creation of a 'good school' for Pupil

Groups with no specified SES and/or gender and/or ethnicity, where the 'good school' is defined as a school with 'high academic achievements'. In the following we look at the activities underlying these School Factors, cf. Table 4.10 and Table 4.11.

School Factors / Studies	1. School size	2. Class size	3. Management and Leadership	4. Curriculum/scheduling	5. School culture and climate	6. Teacher	7. Support teams	8. Physical environment	9. Pupil composition of the school	10. Parental Relationship
Campbell				x	X					
Opdenakker, 2007	x			x	X	x			X	
Rogers		x		x	X	x			X	x
Rumberger				x	X	x				
Tarter			X		X	x			X	
Van Damme					X	x			X	
Yu						x				
Reezigt			X	x	X	x				
Van der Werf, 1997	x			x		x				x
Teddlie			X		X	x				x
Young, 1992						x			X	
Ross, 2006b			X							x
Woessmann		x	X	x		x				
Sum	**2**	**2**	**5**	**7**	**8**	**11**	**0**	**0**	**5**	**4**

Table 4.10: Academic achievement (high weight of evidence) – no specific group of pupils

(N = 13)

School factor / Studies	1. School size	2. Class size	3. Management and leadership	4. Curriculum/ scheduling	5. School culture and climate	6. Teacher	7. Support teams	8. Physical environment	9. Pupil composition of the school	10. Parental Relationship
Bondi	x								X	
Choi					x	x	x		X	
Coates	x	x		X		x				
Dumay				X	x					
Foley	x		X		x	x	x			x
Fullarton	x					x				
Griffith, 2003					x	x				x
Grøgaard		x	X	X	x	x	x	x	X	
Heck	x		X			x				
Hofman, 1996			X		x					x
Hofman, 2002			X	X	x	x				x
Hoy			X		x	x				
Kennedy			X		x	x			X	x
Kyriakides			X	X	x	x	x			
Lamb	x	x		X	x	x				
Martin	x	x		X	x	x			X	
Opdenakker, 2000					x	x				
Papanastasiou					x	x				
Perez	x	x	X			x	x			
Ringsmose		x	X	X	x	x	x	x	X	x
Senkbeil			X		x	x			X	x
Smyth	x	x		X	x	x				
Sweetland			X	X		x				x
Thomas, 1995		x				x	x		X	
Waxman			X	X	x	x				
Webster			X		x	x				
Witte			X		x					x
Young, 2001					x				X	
Zigarelli			X	X	x	x				x
Sum	**9**	**8**	**16**	**12**	**22**	**24**	**7**	**2**	**9**	**10**

Table 4.11: Academic achievement (medium weight of evidence) – no specific group of pupils

(N=29)

1. School size

The school factor 'School size' is concerned with the number of pupils in the school. No subcategories are made regarding this factor.

Table 4.10 and Table 4.11 show that two high weight of evidence studies and nine medium weight of evidence studies deal with the school factor School Size, cf. also Appendix 2, Section 7.1.

Table 4.12 indicates how positive and negative significance, insignificance and intractability are distributed over the 11 studies.

School factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	Insignificant	Intractable		Positive	Negative	Insignificant	Intractable	
1. School size	2	1		1		9	2	1	6		18 %

Table 4.12: School size (academic achievement; no specific group)

(N=11) The final column 'Relative % weight' indicates the percentage of studies with 'high' weight of evidence in relation to the total number of studies included for the elucidation of a given aspect

It appears from the figures in Table 4.12 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the school factor 'School Size' (n=11; m=4) is significant compared to the alternative frequency of 0.10 with $m \geq 2$ with the power 0.91, and compared with the alternative frequencies 0.15 and 0.20 with $m \geq 3$ with the power 0.93 and 0.84 respectively.

Because school size varies greatly from country to country (in TIMSS, average school size for eighth-grade students ranged from about 180 students in Norway to over 1200 in Singapore), TIMSS defined the size of a school in relation to the average school size in each country. Large schools were those with student enrolment greater than the average for the country (Martin, 2000).

According to Bondi (1991) and Fullarton (2004) neither urban-rural location nor School Size can account for between-school variations of pupil achievements. Martin (2000), however, states that in the TIMSS study the school factor School Size and Location did discriminate between the high- and low-achieving schools in some countries, although the variable did not work consistently across all countries.

As the concept of 'school size' is defined in relation to the average school size in a country,It is seen *that the concept 'school size' is applied inconsistently in the studies.Therefore, no conclusion is warranted concerning this factor*.

2. Class size

Class size concerns the number of pupils in the class. Besides the number of pupils, this factor also comprises dimensions such as student-teacher ratios, teacher aid and teaching assistance. No subcategories are made regarding this school factor.

Table 4.10 and Table 4.11 show that two high weight of evidence studies and eight medium weight of evidence studies deal with the school factor Class Size, cf. also Appendix 2, 7.2.

Table 4.13 indicates how positive and negative significance, insignificance and intractability are distributed over the 10 studies.

It appears from the figures of Table 4.13 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the school factor 'Class Size' (n=10; m=4) is significant compared to the alternative frequencies of 0.10 and 0.15 with $m \geq 2$ with the powers 0.93 and 0.82 respectively, and compared with the alternative frequency of 0.20 with $m \geq 3$ with the power 0.88.

It has been found, for instance, that classes having smaller numbers of pupil outperformed classes having a greater number of pupil. Similarly, classes containing pupils with learning disabilities, or pu-

pils who have failed or are repeating a grade, may be more successful if the class size is reduced (Rogers).

School factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	Insignificant	Intractable		Positive	Negative	Insignificant	Intractable	
2. Class size	2	1		1		8	2	1	5		20 %

Table 4.13: Class size (academic achievement; no specific group) (N=10)

However, another study (Woessmann) produced a more ambivalent result. There was no proof that smaller classes performed better than larger ones, since in fact it seemed that larger classes performed better. A possible explanation for this could be that low-performing pupils are placed into smaller classes, which results in larger classes performing better. Woessmann's conclusion is thus that no class size effect could be shown.

As it is seen that the studies have not controlled for unbiased sampling to school classes *no conclusion is warranted concerning 'class size'*.

3. Management and Leadership

The scope of the school factor 'management and leadership' is defined as follows:

The concepts of management and leadership are often used interchangeably in the study of schools. Leadership could, however, be

seen as the broader concept relative to the two narrower concepts: management and educational leadership.

Management concerns the local school level as the decision-making authority. It is related to decisions concerning curricula, instructional technologies, and other school initiatives. Three areas of decision-making can be school based: budget (e.g. decisions regarding personnel, equipment, materials, and staff development), personnel (e.g. recruitment), and curriculum (e.g. decisions regarding the curriculum and instructional strategies at the school level within a framework of district or state goals).

Educational leadership is traditionally associated with people in positions such as principals and superintendents. Accordingly, principals and superintendents are the parties most responsible for crafting the essential educational agreements upon which schools either succeed or fail.

All studies with a bearing on this factor/phenomenon have been classified on the two following dimensions: [2] External orientation of leadership, internal orientation of leadership

Content of leadership: Human resources, rational goal leadership (Quinn & Rohrbauch, 1983), educational leadership, administrational leadership, etc.

Several studies employed more than one measure to assess leadership/management. In such cases each measure has been classified according to this system.

Table 4.10 and Table 4.11 show that five high weight of evidence studies and 16 medium weight of evidence studies deal with the sub-

[2] An attempt was also made to classify the studies on the following leadership/management dimensions: Transactional, transformational, distributive, not applicable. As this led to 3 out of 4 of the studies being classified as 'not applicable', these dimensions have been left out of the analysis.

categories of the complex school factor Management and Leadership, cf. also Appendix 2, Section 7.3.

School factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	Insignifi-cant	Intractable		Positive	Negative	Insignifi-cant	Intractable	
3.1 Human resources	3	1		2		9	3		6		25 %
3.2 Rational goal leadership						2	1		1		0
3.3 Educational leadership	3		2	1		10	4		6		23 %
3.4 Administrational leader-ship	2	2				7	1		6		22 %
3.5 Other						5	3		2		0

Table 4.14: Management and Leadership (academic achievement; no specific group) (N=21)

Table 4.14 indicates how positive and negative significance, insignificance and intractability are distributed over the subcategories.

In the light of the primary studies included in this synthesis it appears from the figures of Table 4.14 that a number of studies show relationships with significance. This indicates that the complex school factor Management and Leadership is of importance for creation of a 'good school' for Pupil Groups with no specified SES.

If we look closer into the subcategories, two are of particular interest:

It appears from the figures of Table 4.14 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the subcategory 'Human Resources' (n=12; m=4) is significant compared to the alternative frequency of 0.10 with $m \geq 2$ with the power 0.89, compared with the alternative frequency of 0.15 with $m \geq 3$ with the

power 0.91, and compared with the alternative frequency of 0.20 with $m \geq 4$ with the power 0.88.

The subcategory Human Resources (3.1) covers three main aspects: the principal's years of experience, hours spent working and his or her availability for the teachers. It also covers the principal's policy concerning the teachers' growth, and influence on hiring and firing staff. Finally, it looks into the influence of members of the school organisation such as teachers and the principal, but especially the parents, on the decisions of the school board.

It also appears from the figures of Table 4.14 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the subcategory 'Educational leadership' (n=13; m=6) is significant compared to the alternative frequency of 0.10 with $m \geq 2$ significant with the power 0.87, compared with the alternative frequency of 0.15 with $m \geq 3$ significant with the power 0.88, and compared with the alternative frequency of 0.20 with $m \geq 4$ significant with the power 0.90.

The subcategory Educational leadership (3.3) covers the situation that the principal demonstrates strong leadership, especially in the areas of curriculum and instruction, and is able to involve other staff members in leadership activities and positions, that the principal's behaviour is supportive and egalitarian and neither directive nor restrictive, and that the principal is 'resource supportive', e.g. in deciding textbooks and contents of the teaching.

Several studies emphasise the ways in which this leadership ought to be expressed.

Tarter (2004) says that principals and teachers ought to work as colleagues while retaining their distinctive roles.

Ross & Gray (2006b) claims that where leadership has produced collective teacher efficacy, commitment to school mission, commitment to professional community, and commitment to community partnerships, the pupils demonstrate high academic achievement. However, it has been found contrariwise that professionalization policy (observation of lessons by teachers, principal sitting in on lessons) had a significant negative effect (Reezigt et al., 1999). To this it may be added that Educational Leadership was found to be negatively related to effectiveness, meaning that less effective schools manifested more educational leadership (Van der Werf, 1997). This result could be interpreted to say that less effective schools cause more educational leadership.

Finally, there is an ambiguous result concerning Human Resources: Pupils in schools with autonomy in determining teacher salaries performed better. However, in Hong Kong the opposite was the case (Ross & Gray, 2006b).

We cannot conclude anything about the subcategories 'Rational Goal Leadership' (3.2) and 'Administrative Leadership' (3.4) with n=2 and n=9 respectively.

4. Curriculum/scheduling

The scope of the school factor 'curriculum/scheduling' is defined as follows:

Curriculum is often defined as covering only those topics actually taught by teachers. However, the definition of curriculum can range from virtually everything that takes place in a classroom to the topics that are defined as instructional requirements in the legal regulation of an educational system. Curriculum can further be subdivided into three components: the intended, the implemented, and the at-

tained. Typical examples could be opportunity to learn, homework, coordination and alignment of the curriculum, and learning goals.

All studies with a bearing on this factor/phenomenon have been classified on the following subcategories:

Opportunity to learn: This subcategory consists of the curriculum actually offered to the students. (Homework is placed in 'Opportunity to learn').

Alignment: 'Alignment' is about coordination, i.e., bringing purpose and means together. (Differentiation on an organisational level such as 'single gender classroom' is placed in this subcategory).

Learning goals

Other (School resources such as books are categorised in the 'Other' category).

Table 4.10 and Table 4.11 show that seven high weight of evidence studies and 12 medium weight of evidence studies deal with the subcategories of the complex school factor Curriculum/scheduling, cf. also Appendix 2, Section 7.4.

Table 4.15 indicates how positive and negative significance, insignificance and intractability are distributed over the subcategories.

School factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	~	Intractable		Positive	Negative	~	Intractable	
4.1 Opportunity to learn	6	6				11	7	1	3		35 %
4.2 Alignment						2		1		1	0
4.3 Learning goals	1	1				1	1				50 %
4.4 Other						2	1	1			0

Table 4.15: Curriculum/scheduling (academic achievement; no specific group) (N = 19)

It appears from the figures of Table 4.15 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the subcategory 'Opportunity to Learn' (n=17; m=13) is significant compared to the alternative frequency of 0.10 with $m \geq 3$ significant with the power 0.92, compared with the alternative frequency of 0.15 with $m \geq 4$ significant with the power 0.90, and compared with the alternative frequency of 0.20 with $m \geq 5$ significant with the power 0.89.

The subcategory Opportunity to Learn (4.1) covers number of teaching hours, including homework hours. It could be seen as the teacher's efficiency of organising the instruction process, measured by the percentage of time teachers reported spending on the planning of their lessons for the following day, the making of a weekly teaching plan, keeping to the timetable, and by the assigned time spent on

lessons. It also includes homework hours, which are the total hours pupils spent on homework both in school and out of school per week.

Reezigt et al. (1999), Campbell et al. (2000) and Woesmann (2003) have demonstrated that frequency of homework has a significant positive effect on pupil achievement.

However, Campbell claims that there is a lack of significant effect of the 'time on task' variable both at pupil and classroom level – the study concludes that further research is needed to investigate this phenomenon. In contrast, Van der Werf (1997) claims that efficient allocation of opportunity and time to learn within arithmetic lessons seems to be especially important in explaining the differences in pupils' arithmetic achievement across schools.

One curious result must be mentioned: Classes in which mathematics was taught in the morning outperformed classes in which mathematics was taught in the afternoon (Rogers et al., 2006).

We cannot conclude anything about the subcategories 'Alignment' (4.2) and 'Learning goals' (4.3) with n=2 and n=2 respectively.

5. School Culture and School Climate

'School culture and climate' is understood in terms of the feel, atmosphere, tone, ideology, or milieu of a school. The concepts of school climate and school culture are often used interchangeably in the study of schools. Some authors, however, make a distinction between the two.

While much of the school climate literature focuses on the structural dimensions of schools, culture looks beyond structural elements, both the formal and informal specifics, to the meanings those specifics hold for the participants and how they make use of them.

When school climate and school culture are seen as synonyms, the indicators of school culture/climate can range from perceptions and normative views to behavioural characteristics and factual circumstances (e.g. shared visions, goals and values, monitoring progress, achievement orientation, internal relationships, evaluative potential, feedback reinforcement and behavioural rules).

All studies with a bearing on this factor/phenomenon have been classified into the following subcategories:

Disciplinary climate

Achievement/progress orientation (This subcategory includes an evaluative culture; it also includes the students' attitude towards the school and school work as well as the students' self-concept regarding the school work)

Interrelational climate

Social norms and values (Pupil involvement is assigned to this subcategory)

Other.

Table 4.10 and Table 4.11 show that eight high weight of evidence studies and 22 medium weight of evidence studies deal with the subcategories of the complex school factor School Culture and School Climate, cf. also Appendix 2, Section 7.5.

Table 4.16 indicates how positive and negative significance, insignificance and intractability are distributed over the subcategories.

In the light of the primary studies included in this synthesis it appears from the figures of Table 4.16 that a considerable number of studies show relationships with significance.

If we look closer into the subcategories, three are of particular interest:

It appears from the figures of Table 4.16 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the subcategory 'Disciplinary climate' (n=11; m=4) is significant compared to the alternative frequency of 0.10 with m ≥ 2 significant with the power 0.91, and compared with the alternative frequencies 0.15 and 0.20 with m ≥ 3 significant with the power 0.93 and 0.84 respectively.

The subcategory Disciplinary Climate (5.1) covers a school where an orderly atmosphere prevails. Orderly atmosphere had a significant positive effect (Reezigt et al., 1999), as did an ordered environment in which appropriate pupil behaviours are present (Ross et al., 2006b). Rogers et al. (2006) showed that the adjusted school mean of schools with more severe disciplinary problems was 0.071 standard deviations lower than the adjusted school mean of schools with less severe problems. A 'good school' for pupils with no specified SES is a school where pupils do not feel unsafe, since the proportion of pupils who feel unsafe has a significant negative effect on pupil achievement (Rumberger et al., 2005).

It appears from the figures of Table 4.16 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the subcategory 'Achievement/Progress Orientation' (n=21; m=16) is significant compared to the alternative frequency of 0.10 significant with m ≥ 3 with the power 0.85, compared with the alternative frequency of 0.15 with m ≥ 4 significant with the power 0.80, and compared with the alternative frequency of 0.20 with m ≥ 6 significant with the power 0.89.

The subcategory Achievement/Progress Orientation (5.2) covers a school where the focus is on academic achievement and high expectations (cf. also Ross et al., 2006b; Campbell et al., 2000; Teddlie & Stringfield, 1993; Rumberger et al., 2005).

School Factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	Insignifi.	Intractable		Positive	Negative	Insignifi.	Intractable	
5.1 Disciplinary climate	1			1		10	3	1	4	2	9 %
5.2 Achievement/progress orientation	5	4			1	16	12		3	1	24 %
5.3 Interrelational climate	4	2	1	1		8	4		4		33 %
5.4 Social norms and values	1			1		9	5		4		10 %
5.5 Other						2	1			1	0

Table 4.16: School Culture and School Climate (academic achievement; no specific group)

(N = 30)

It may be added that schools in which academic achievement was more frequently recognized at the school level outperformed schools in which academic achievement was less frequently recognized (Rogers). Where high pupil engagement (Ross et al., 2006b) and teacher rated attentiveness (Campbell) were also present, this also led to high Pupil Achievement. One aspect of this is that 'learning climate' has a significant positive effect on pupil achievement. The term 'learning climate' refers to a very broad variable (Van Damme).

It appears also from the figures of Table 4.16 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05, the

subcategory 'Interrelational Climate' (n=12; m=7) is significant compared to the alternative frequency of 0.10 with $m \geq 2$ significant with the power 0.89, compared with the alternative frequency of 0.15 with $m \geq 3$ significant with the power 0.91, and compared with the alternative frequency of 0.20 with $m \geq 4$ significant with the power 0.88.

The subcategory Interrelational Climate (5.3) covers affiliation, support/respect for staff and pupils, and warmth in teacher/pupil relationships. Teachers can obtain assistance, advice and encouragement, and are made to feel accepted by their colleagues. Pupils develop positive relationships. Teddlie & Stringfield (1993) found that an absence of negative peer pressure correlates positively with achievement.

Finally, it appears from the figures of Table 4.16 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05, the subcategory 'Social norms and values' (n=10; m=5) is significant compared to the alternative frequencies of 0.10 and 0.15 with $m \geq 2$ with the power 0.93 and 0.82 respectively, and compared with the alternative frequency of 0.20 with $m \geq 3$ with the power 0.88.

The subcategory Social norms and values (5.4) covers teachers' professional values like interest in their work and professional development, and an interest in new educational plans and experimentation, classroom openness and individualisation. Teachers appreciate taking a full participation in school activities, feeling ownership of what happens in the school and accept that a work pressure dominates the school environment. Students enjoy class work, and they are involved and think that they are doing a good job in classes. Physical features of rooms, equipment, and buildings are maintained and kept orderly.

6. Teacher

Teacher is understood in terms of the teacher as an individual and/or the teacher as part of an organisation.

All studies with a bearing on the 'teacher as an individual teacher' have been classified into the following subcategories:

Teacher behaviour: covers the ways teachers ensure that pupils behave in an appropriate manner both towards each other/the teacher, and in relation to the learning that is to take place in the school. It is about getting the teaching right (e.g. by differentiation/using a variety of teaching strategies). Examples of teacher behaviour are:

Classroom management: teacher's organisation and structuring of teaching

Behaviour management: Correction of student misbehaviour e.g. rewards truly praiseworthy behaviour.

Classroom climate: Contribution from the teacher to the classroom climate e.g. high expectations, teacher enthusiasm, avoids criticism.

Teacher beliefs: represents teacher's theories about how pupils function, i.e. their beliefs about what constitutes 'good teaching'.

Subject knowledge: is about the teacher's content knowledge of his/her subject

Teacher self-efficacy beliefs. This is covered by two concepts:

Teachers' self-concept (a person's perception of him-/herself, formed through interaction with the environment)

Teachers' self-efficacy (a teacher's judgment of his/her capabilities to bring about desired outcomes of the student engagement and learning)

The scope of 'teacher as an organisational actor' is determined as follows:

The aspect could contain teacher groups/teams, the teachers' job satisfaction, teachers' gender, teacher corps stability, teachers' formal competence (certified/uncertified teacher/teaching assistant)

Table 4.10 and Table 4.11 show that 11 high weight of evidence studies and 23 medium weight of evidence studies focus on the subcategories of the complex school factor Teacher, cf. also Appendix 2, Section 7.6.

Table 4.17 indicates how positive and negative significance, insignificance and intractability are distributed over the subcategories.

In the light of the primary studies included in this synthesis it appears from the figures of Table 4.17 that a number of studies show relationships with significance.

If we look closer into the subcategories, two are of particular interest:

It appears from the figures of Table 4.16 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the subcategory 'Teacher behaviour' (n=24; m=14) is significant compared to the alternative frequency of 0.10 with $m \geq 4$ significant with the power 0.91, compared with the alternative frequency of 0.15 with $m \geq 5$ significant with the power 0.86, and compared with the alternative frequency of 0.20 with $m \geq 6$ significant with the power 0.81.

The subcategory Teacher Behaviour (6.1) covers a number of aspects. It is therefore not surprising that there are no clear conclusions with respect to this school factor. We consider first the teacher's behaviour in the classroom.

Ross et al. (2006b) state that direct instruction, use of higher level questioning, sustained writing, and high usage of academic focus all strengthen pupil achievement.

On the other hand, Rogers states that mathematics classes in which the pupils spent a greater proportion of time working either alone or in small groups outperformed classes in which a greater proportion of time was spent on direct teaching to the whole class or on small group instruction.

School Factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	[illegible]	Intractable		Positive	Negative	[illegible]	Intractable	
6.1 Teacher behaviour	7	3	2		2	17	8	1	5	3	29 %
6.2 Teacher beliefs	1	1				1	1				50 %
6.3 Teacher self-efficacy beliefs	1				1	2			2		33 %
6.4 Teacher subject knowledge						2	2				0
6.5 Teacher as an organisational actor	8	4		1	3	15	6	1	4	4	35 %

Table 4.17: Teacher (academic achievement; no specific group) (N = 34)

Van der Werf (1997) claims that in highly effective schools, teachers give more whole-class instruction and spend their lesson time more efficiently. On the other hand she also claims that teachers in effective schools spend more time on learning activities and evaluation of

learning tasks, provide more arithmetic content (opportunity to learn), and — unexpectedly — that teachers in effective schools give fewer arithmetic lessons than in non-effective schools.

Finally Reezigt makes the claim that in language subjects, feedback had a significant positive effect. In contrast, in mathematics, feedback had – surprisingly – a significant negative effect.

It appears also from the figures of Table 4.16 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the subcategory 'Teacher as an organisational actor' (n=23; m=11) is significant compared to the alternative frequency of 0.10 with $m \geq 3$ significant with the power 0.81, compared with the alternative frequency of 0.15 with $m \geq 5$ significant with the power 0.88, and compared with the alternative frequency of 0.20 with $m \geq 6$ significant with the power 0.84.

The subcategory 'Teacher as an Organisational Actor' (6.5) covers a number of aspects connected with the teacher's role in the school organisation, e.g. formal competence, average years of experience, average years of education, staff job satisfaction, teacher engagement, teacher cooperation, gender composition (number of female teachers in the school), and number of teachers with in-service training.

Two additional factors require brief comment: teacher experience, and the gender of the teacher.

Rogers claims that teaching experience at a variety of grade levels was negatively related to class performance, while Woesmann found that the impact of teacher experience in years varies between countries, and was positive up to a certain level.

Rogers states that female teachers achieved slightly better results than male teachers, while Young found that at school level there was

a significant negative effect related to the percentage of female teachers.

We cannot conclude anything about the subcategories 'Teacher beliefs' (6.2), 'Teacher self-efficacy beliefs' (6.3) or 'Teacher subject knowledge' (6.4) with n=2, n=3 and n=2 respectively.

7. Support teams

The scope of 'support teams' is defined as follows:

'Support teams' is concerned with non-instructional services or extra-curricular activities with the goal of addressing students´ needs, e.g., school dentist, nurse, advisors, and leisure-time activities.

No subcategories have been defined for this factor.

Table 4.10 and Table 4.11 show that seven medium weight of evidence studies deal with the school factor Support Teams, cf. also Appendix 2, Section 7.7.

School factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	Insignificant	Intractable		Positive	Negative	Insignificant	Intractable	
7. Support teams	0					7	1		4	2	0

Table 4.18: Support teams (academic achievement; no specific group) (N = 7)

Table 4.18 indicates how positive and negative significance, insignificance and intractability are distributed over the seven studies.

In the light of the primary studies included in this synthesis it appears from the figures of Table 4.18 that it cannot be established whether a variation in the school factor 'Support Team' (n=7; m=1) is of importance for the 'good school', where the 'good school' is defined as a school with 'high academic achievements'.

8. Physical environment

Studies grouped within the school factor 'Physical environment' all deal with the physical characteristics of the school. Examples of such characteristics are facilities such as furnishing, materials and supplies, equipment and information technology, characteristics of the school building and various aspects of the school layout such as athletic fields and playgrounds.

No subcategories have been defined for this school factor.

School factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	~gant	ble		Positive	Negative	~gant	ble	
8. Physical environment	0					2	1			1	0

Table 4.19: Physical environment (academic achievement; no specific group) (N = 2)

Table 4.10 and Table 4.19 show that two medium weight of evidence studies focus on the school factor Physical Environment, cf. also Appendix 2, Section 7.8.

Table 4.19 indicates how positive and negative significance, insignificance and intractability are distributed over the two studies.

In the light of the primary studies included in this synthesis it appears from the figures of Table 4.19 (n=2; m=1) that it cannot be established that a variation in the school factor Physical Environment is of importance for the 'good school', where the 'good school' is defined as a school with 'high academic achievements'.

9. Pupil Composition of the School

Studies that are grouped in the category 'Pupil composition of the school' all deal with the effects of percentages of different groups of pupils in the school (e.g. social economic status, special educational needs and ethnicity). This factor is more accurately defined as *"the aggregate characteristics of a student group on a student's learning over and above the effects on learning associated with that student's individual characteristic"* (Wilkinson, 2002 in Dumay & Dupriez, 2007). Hence, this factor is not to be confused with the inclusion criteria which every study has met in order to be included in this review: "control is present for differences in pupils' socio-economic background" or "control is present for differences in pupils' scholastic aptitude". *(NB! It has sometimes been difficult to determine whether the studies addressing this factor are dealing with pupil composition of the school as a control variable or as a malleable school factor)*

No subcategories have been defined for this factor.

Table 4.10 and Table 4.11 show that five high weight of evidence studies and nine medium weight of evidence studies deal with the school factor Pupil Composition of the School, cf. also Appendix 2, Section 7.9.

Table 4.20 indicates how positive and negative significance, insignificance and intractability are distributed over the 14 studies.

It appears from the figures of Table 4.20 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the school factor 'Pupil composition of the school' (n=14; m=11) is significant compared to the alternative frequency of 0.10 with $m \geq 2$ with the power 0.84, compared to the alternative frequency of 0.15 with $m \geq 3$ with the power 0.85, and compared with the alternative frequency 0.20 with $m \geq 4$ significant with the power 0.87.

School Factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative				Positive	Negative			
9. Pupil composition of the school	5	4		1		9	6	1	2		36 %

Table 4.20: Pupil Composition of the School (academic achievement; no specific group)

(N = 14)

Some studies comment on this factor:

Opdenakker states that a student body originating from high SES homes and with good academic abilities has a positive effect on outcomes. Tarter finds that the socio-economic level of the school had a significant positive effect on pupil achievement. Van Damme finds that the proportion of girls in the class had a significant positive effect on pupil achievement, and that the level of the initial cognitive ability has a significant positive effect on pupil achievement. Young found that the student characteristics of greatest significance were pupil gender, attitude towards science, ethnicity, verbal ability, quantitative ability and socio-educational level.

10. Parental Relationship

The school factor 'Parental Relationship' is used to group studies that deal with the emphasis on parental involvement in school policy and contact with parents. Illustrative examples of the school's role in encouraging parental involvement include practices such as conducting workshops for families, and communicating to parents about their children's education.

No subcategories have been defined for this school factor.

Table 4.10 and Table 4.11 show that four high weight of evidence studies and 10 medium weight of evidence studies deal with the school factor Parental Relationship, cf. also Appendix 2, Section 7.10.

Table 4.21 indicates how positive and negative significance, insignificance and intractability are distributed over the 14 studies.

School factor	High studies	Significance				Medium studies	Significance				% Relative weight
		Positive	Negative	Insignifi-	Intracta-		Positive	Negative	Insignifi-	Intracta-	
10. Parental Relationship	4	2		1	1	10	4	1	5		29 %

Table 4.21: Parental Relationship (academic achievement; no specific group) (N=14)

It appears from the figures of Table 4.21 and the power calculation table in Chap. 9 Appendix 4 that with significance p=0.05 the school factor 'Parental Relationship' (n=14; m=7) is significant compared to the alternative frequency of 0.10 with $m \geq 2$ with the power 0.84, compared to the alternative frequency of 0.15 with $m \geq 3$ with the

power 0.85, and compared with the alternative frequency 0.20 with m ≥ 4 with the power 0.87.

It seems thus to be the case that the 'good school' for pupils with no specified SES is a school where parents support child and school, are concerned about grades and education, and are committed to community partnership (Rogers; Teddlie; Ross & Gray, 2006b).

4.3.1.3.2 The 'good school': non-academic achievement, for pupils with no specified SES

The studies grouped together here all look into School Factors and Activities of importance for creation of a 'good school' for Pupil Groups with no specified SES and/or gender, where the 'good school' is defined as a school with 'high non-academic achievements', cf. Table 4.22 and Table 4.23.

School factors / Studies	1.School size	2. Class size	3. Management and Leadership	4. Curriculum/scheduling	5. School culture and climate	6. Teacher	9.Pupil composition of the school
Opdenakker, 2007	x		x	x	x	x	x
Rumberger			x	x	x	x	
Van Damme					x		
Sum	1	0	2	2	3	2	1

Table 4.22: No specific group of pupils (Non-academic achievement; high weight of evidence)

(N = 3)

Many of these studies have at the same time also looked into School Factors and Activities of importance for creation of a 'good school' for Pupil Groups with no specified SES and/or gender, where the 'good school' is defined as a school with 'high academic achievements', i.e. the studies already analysed in the former Section 4.3.1.3.1. As none

of the school factors in this group has an n ≥ 10, we cannot conclude anything about the school factors in this group.

School factor / Studies	1. School size	2. Class size	3. Management and leadership	4. Curriculum/ scheduling	5. School culture and climate	6. Teacher	7. Support teams	8. Physical environment	[illegible] sition of the school	10. Parental Relationship
Foley	x		x		x	x	x			x
Kyriakides			x	X	x	x	x			
Opdenakker, 2000					x	x				
Silins	x		x		x	x			x	
Smyth	x	x		X	x	x			x	
Witte			x		x					x
Sum	3	1	4	2	6	5	2	0	2	2

Table 4.23: No specific group of pupils (Non-academic achievement; medium weight of evidence)

(N = 6)

4.3.1.3.3 The 'good school': academic achievement, for pupils with low SES

The studies grouped together here all look into School Factors and Activities of importance for creation of a 'good school' for Pupil Groups with low SES, where the 'good school' is defined as a school with 'high academic achievements'. It appears that the activities underlying these School Factors, cf. Table 4.24 and Table 4.25, only one complex school factor 'Teacher' has an n ≥ 10. Therefore, this is the only school factor treated in this group.

School factors / Studies	1. School size	2. Class size	3. Management and leadership	4. Curriculum/ scheduling	5. School culture and climate	6. Teacher	7. Support teams	8. Physical environment	9. Pupil composition of the school	10. Parental Relationship
Taylor				x	x	x				x

Table 4.24: Academic achievement (high weight of evidence) for Pupil Groups with low SES

(N = 1)

School factors / Studies	1. School size	2. Class size	3. Management and leadership	4. Curriculum/ scheduling	5. School culture and climate	6. Teacher	7. Support teams	8. Physical environment	9. Pupil composition of the school	10. Parental Relationship
Bain		x				x				
Bottoms				X	x	x	x			
Florida	x	x	x	X	x	x	x		x	x
Franklin	x	x				x				
Griffith, 2002					x	x				
Grisay	x	x	x	X	x	x	x	x	x	
Mandeville	x			X		x			x	
Meijnen				X		x				
Sammons			x	X	x	x				x
Senkbeil			x		x	x			x	x
Traufler			x	X	x	x				x
Sum	4	4	5	7	7	11	3	1	4	4

Table 4.25: Academic achievement (medium weight of evidence) for Pupil Groups with low SES

(N=11)

6. Teacher

For the definition of the category ‘teacher’ and its subcategories, see p. 102.

School Factor	High studies	Significance				Medium studies	Significance				Relative % weight
		Positive	Negative	Insignifi-	Intractable		Positive	Negative	Insignifi-	Intractable	
6.1 Teacher behaviour	1				1	9	2		5	2	10 %
6.2 Teacher beliefs											
6.3 Teacher self-efficacy beliefs						1			1		
6.4 Teacher subject know-ledge											
6.5 Teacher as an organ-isational actor						8	3		4	1	

Table 4.26: Teacher (academic achievement; Pupil Groups with low SES) (N = 12)

Table 4.24 and Table 4.25 show that one high weight of evidence study and 11 medium weight of evidence studies deal with the sub-categories of the complex school factor 'Teacher'.

Table 4.26 indicates how positive and negative significance, insignificance and intractability are distributed over the subcategories.

In the light of the primary studies included in this synthesis it appears from the figures of Table 4.26 that some studies show relationships with significance.

4.3.1.3.4 The 'good school': non-academic achievement, for pupils with low SES

The studies grouped together here all look into School Factors and Activities of importance for creation of a 'good school' for Pupil Groups with low SES, where the 'good school' is defined as a school with 'high non-academic achievements'. Table 4.7 and Table 4.9 show that only one medium weight of evidence study falls in this category. An analysis is therefore not possible.

4.3.2 Group 2: Synthesis of qualitative studies

The presentation is organised as follows. First, all the studies having a similar idea of the 'good school' (Pupil Achievement) are collected together. Then, these studies are distributed into Pupil Groups in keeping with the categories applied in the data extraction system. Finally, for each of the created groups, school factors and school activities are recorded.

4.3.2.1 What is a 'good school'?

All 15 qualitative studies identify the 'good school' as a school with high academic achievement. They thus apply an unambiguous definition of 'good school'. Table 4.27 shows on which subjects the estimation of Pupil Achievement is based.

10 studies look at pupil achievements in 1st Language, 7 in Maths, and 3 in Science, 2nd Language, and History respectively. 4 studies have no specified subjects.

Subjects / Studies	1st Language	Mathematics	Science	2nd Language	History	Subject not specified
Lindsay	x					
Mosenthal	x					
Florida	x	x				
Kitchen		x				
Ringsmose						x
Picucci	x	x				
Willis	x	x				
Pressley, 2004						x
Taylor	x					
Towns						x
Pressley, 2007	x					
Texas	x	x				
Stringfield						x
Sammons	x	x	X	x	x	
Foley	x	x				
Sum	10	7	1	1	1	4

Table 4.27: Which subjects measure Pupil Achievements (Qualitative studies; academic achievements)?

(N=15)

4.3.2.2 A good school — for whom?

The included studies specify for which Pupil Groups the 'good school' is good, cf. Table 4.28.

10 studies have examined the 'good school' for pupils with low SES, 6 studies focused on bilingual or ethnic minorities, 3 studies on pupils with middle SES and one study on pupils with high SES. One study looks at gender, i.e. the 'good school' for boys. 4 studies have no specified SES and/or gender.

For Whom? / Studies	Boys	Low SES	Middle SES	High SES	Language & ethnic minorities	specified SES and/or gender
Lindsay	x	x			x	
Mosenthal						x
Florida		x				
Kitchen		x			x	
Ringsmose		x	X	x		
Picucci		x				
Willis		x			x	
Pressley, 2004		x			x	
Taylor		x	X			
Towns		x			x	
Pressley , 2007			X			
Texas		x			x	
Stringfield						x
Sammons						x
Sum	1	10	3	1	6	3

Table 4.28: The 'good school' for whom? (Qualitative studies; academic achievements)

(N=15)

4.3.2.3 The good school — how?

Table 4.29 reports a survey of the school factors identified by the 15 studies as important for high Pupil Achievement.

All school factors are seen as important for high Pupil Achievement, but some are stressed more than others. 14 studies point to the factor Management and Leadership, 14 studies to School Culture and Climate, 11 to Curriculum/Scheduling, and 11 to the Teacher. 6 studies

point to Parental Relationship, 5 to Physical Environment and 4 to Support Teams.

It has to be noted that the four most significant school factors are identical with the four factors pointed out in the analyses of the statistical significance of the four categories discussed in Section 8.3. This means that the quantitative and qualitative studies are in agreement on this issue.

4 studies point to no specific factor or combination of factors. As seen in the table, however, these last four studies mention some factors of significance for the 'good school'.

School Factors / Studies	3. Management and leadership	4.Curriculum/Scheduling	5.School culture and climate	6.Teacher	7. Support teams	8. Physical environment	10. Parental Relationship	Factor(s) not specified
Lindsay	x	x	x	X			x	x
Mosenthal	x	x	x	X				x
Florida	x							x
Kitchen	x	x	x	X	x			
Ringsmose	x	x	x		x			
Picucci	x		x	X	x			
Willis		x	x	X		x		
Pressley, 2004	x	x	x	X		x		
Taylor	x	x	x	X			x	
Towns	x	x	x	X		x	x	
Pressley, 2007	x	x	x	X		x	x	x
Texas	x	x	x	X		x	x	
Stringfield	x		x					
Sammons	x		x	X			x	
Foley	x	x	x		x			
Sum	14	11	14	11	4	5	6	4

Table 4.29: School factors of importance for high Pupil Achievement without regard for Pupil Groups

(N = 15)

After this survey we turn to the final task of specifying School Factors and Activities of importance for creation of a 'good school' for the Pupil Groups specified in Section 4.3.2.2.

The presentation will be divided into three groups. However, one of the groups consists of two aspects treated together: low SES and Bilingual or Ethnic Minorities. They are covered by the same studies and in most cases the two aspects are not given a separate treatment

in the included primary studies. The first collected group therefore consists of 10 studies. The second group comprises Pupil Groups with middle SES and consists of 3 studies. The last group discussed here comprises 3 studies with no specified SES and/or gender. One study treats pupils with high SES and another study treats boys. They are accordingly not synthesised separately, but are included in the groups to which they also belong.

4.3.2.3.1 The 'good school' for pupils with low SES (including bilingual and ethnic minorities)

School Factors / Studies	3. Management and leadership	4. Curriculum/Scheduling	5. School culture and climate	6. Teacher	7. Support teams	8. Physical environment	10. Parental Relationship	Factor(s) not specified
Lindsay	x	x	X	X			x	x
Florida	x							x
Kitchen	x	x	X	X	x			
Ringsmose	x	x	X		x			
Picucci	x		X	X	x			
Willis		x	X	X		x		
Pressley, 2004	x	x	X	X		x		
Taylor	x	x	X	X			x	
Towns	x	x	X	X		x	x	
Texas	x	x	X	X		x	x	
Sum	9	8	9	8	3	4	4	2

Table 4.30: School Factors of importance for Pupil Groups with low SES, including bilingual and ethnic minorities

(N=10)

The studies grouped together here all look into School Factors and Activities of importance for creation of a 'good school' for Pupil Groups with low SES and/or belonging to bilingual or ethnic minori-

ties (in the following named 'with low SES'). From Table 4.28 and Table 4.29 we can extract the following Table 4.30.

In the following we look at the activities behind these School Factors.

3. Management and Leadership

The 'good school' for pupils with low SES is a school where:

- The leadership has focus on staff support.

This support could be an everyday occurrence such as making teaching resources available (Kitchen), giving high priority to staff development, and allowing teachers to experiment with the teaching, when this is founded on principles/ideas (Texas).

The 'good school' for pupils with low SES is a school where:

- Educational leadership is strong.

The concept of 'strong educational leadership' stands for good acquaintance with staff, an active encouragement of teachers, explicit educational goals communicated to staff, and visible leadership (Kitchen, 2006; Picucci, 2002; Pressley, 1994; Taylor, 2000; Texas, 2000). Other studies underline democratic leadership, where leadership is in dialogue with teachers and appreciates their contributions (Florida, 1994; Picucci, 2002; Pressly, 2004).

The 'good school' for pupils with low SES is a school where:

- Teachers are free to experiment with teaching and curriculum.

The 'good school' for pupils with low SES is a school where:

- The leaders and managers have formal competences (are certified) and are experienced teachers.

4. Curriculum/Scheduling

In analysing the school factor Curriculum/Scheduling four themes appear: assessment, opportunity to learn, experimental approach and native language teaching.

The 'good school' for pupils with low SES is a school where:

- Assessment and accountability procedures focus on low achievers.

A constituent part is remediation by the use of after-school tutoring and enrichment programmes (Texas, 2000). Kitchen (2006) stresses that teachers prepare students to be successful on standardised tests, but in fact they teach beyond the test. Pressley (2004) claims that standardised test-preparation and focus on test-taking skills are important elements in the school practice of 'good schools'.

The 'good school' for pupils with low SES is a school where:

- Pupils are given extended opportunities to learn.

This can for instance consist of extended class periods for teaching mathematics, pupils' participation in summer schools, attending classes on Saturdays (Kitchen), or receiving club and tutorial support (Lindsay, 2006). According to Pressley (2004), pupils spend long days in intensive schooling followed by homework, cf. also Texas (2000) and Lindsay (2006) on recommendation of homework.

The 'good school' for pupils with low SES is a school where:

- Teachers are free to experiment with teaching and curriculum.

This item refers to the effect of Curriculum/Scheduling, when children are exposed to experimental approaches.

The 'good school' for pupils with low SES is a school where:

- The curriculum contains structured schedules with time for instruction in Native Language and 2nd Language.

The ‘good school’ for bilingual or ethnic minorities includes situations where manipulative and hands-on activities are used extensively, for instance in teaching Maths and Science, and when state-adopted materials and other resources are available in the classroom in both Native Language and 2nd Language versions (Spanish and English) (Texas).

5. School Culture and Climate

The 'good school' for pupils with low SES is a school where:

- School Culture is dominated by a shared vision of common goals.

This vision is very powerful if it comprises the value of ‘success for every child’ combined with the hope that stakeholders go beyond ordinary expectations to ensure student success. Towns (2001) defines this as ‘going the extra mile’, i.e. looking beyond the goal and not being satisfied with merely ‘getting by’. The shared vision includes expectations with regard to standards and styles of teaching and disciplinary policies communicated throughout the entire school (Lindsay, 2006).

Aspects of this shared vision of common goals occur in two other forms:

The 'good school' for pupils with low SES is a school where:

- The school is perceived as a community.

A version of this is that the school is experienced as a ‘family’, broad enough to accommodate all members of the community (Lindsay, 2006; Willis, 1996). This ‘family’ is recognised by relational trust between the members of the school community (Picucci, 2002). The positive atmosphere in the school expresses itself via praise, and a

caring atmosphere free from threat and physical harm (Texas, 2000; Lindsay, 2006).

The 'good school' for pupils with low SES is a school where:

- Focus is on academic achievement and high expectations.

One aspect of this is a principled philosophy of not accepting 'barriers on the road', but being creative and innovative in ensuring pupil learning (Town, 2001). For under-achieving boys it appears that effective schools with a 'global approach' of focusing on every pupil in the classroom at the same time helps the under-achieving boys in particular (Lindsay, 2006). Frequent monitoring of pupil progress is an ingredient of the effective school (Texas, 2000), and Pressley (2004) emphasizes evaluation of pupils and mentoring of college admission as important for creating academic success. Lindsay (2006) mentions performance monitoring, i.e. targeting 'at risk' students.

6. Teacher

The 'good school' for pupils with low SES is a school where:

- The instruction supports high academic achievement with an emphasis on understanding, rather than low-level learning.

Teachers promote high academic achievements if they believe that poverty does not prevent high academic achievement and they focus on the children (and teaching) through e.g. scaffolding, encouragement of self-regulation and feedback (Pressley, 2004). A study of boys with an ethnic minority background emphasizes the effects of the presence of successful male teachers from the relevant communities (Lindsay, 2006).

The 'good school' for pupils with low SES is a school where:

- Teachers have high staff morale i.e. strong internal support and a sense of ownership.

The 'good school' for pupils with low SES is a school where:

- Teachers focus on the individual pupil.

Overall, relationship building, teachers understanding and caring about their pupils, and teachers meeting pupils' emotional needs and making them feel good about themselves, are seen as essential parts of effective schools. They provide targeted interventions and extra services to ensure that no child becomes invisible (Taylor, 2000).

- The 'good school' for pupils with low SES is a school where:

Teachers inform their planning through research and professional development.

The teachers that align professional development with identified needs, seek outside technical assistance, and communicate and collaborate among staff, contribute to the 'good school' (Picucci, 2002; Taylor, 2000). The teachers having a strong belief in self-efficacy contribute to the 'good school'.

8. Physical environment

The 'good school' for pupils with low SES is a school where:

- Schools have good buildings and reduced gang activity.

The 'good school' for pupils with low SES is a school where:

- Schools have good facilities, such as designed buildings, labs, library, books, and computers.

10. Parental Relationship

The 'good school' for pupils with low SES is a school where:

- Parents support the basic mission of the school and are involved in the school community.

Studies show that good home-school relationships support 'good schools'.

4.3.2.3.2 The 'good school' for pupils with middle SES

The studies grouped together here all look into School Factors and Activities of importance for creation of a 'good school' for Pupil Groups with middle SES. From Table 4.28 and Table 4.29 we can extract the following Table 4.31.

In the following we look at the activities behind these School Factors.

3. Management and Leadership

The 'good school' for pupils with middle SES is a school where:

- Leadership is visible.

A visible leader is present in the staff room and takes part in school arrangements (Ringsmose). It is also important that leaders have clear ideas of what is going on in individual classrooms (Pressley 2007).

The 'good school' for pupils with middle SES is a school where:

- The leadership supports the teachers' professional development.

Taylor underlines that new educational programmes take a long time to implement. Presley states that we have to offer teachers as much professional development as possible.

School Factors / Studies	3. Management and leadership	4. Curriculum/Scheduling	5. School culture and climate	6. Teacher	7. Support teams	8. Physical environment	10. Parental Relationship	Factor(s) not specified
Ringsmose	x	x	x		x			
Taylor	x	x	x	X			x	
Pressley, 2007	x	x	x	X		x	x	
Sum	3	3	3	2	1	1	2	0

Table 4.31: School Factors of importance for Pupil Groups with middle SES (N = 3)

4. Curriculum/Scheduling

The 'good school' for pupils with middle SES is a school where:

- High priority is given to teaching in school subjects.

The school has to focus on academic achievements (Pressley). Pupils are given a lot of homework and are offered special courses if necessary. In addition, screening is conducted early in the school year to make sure that low achievers get the necessary support (Ringsmose). The daily school work is characterised by routine.

5. School culture and climate

The 'good school' for pupils with middle SES is a school where:

- Focus is on academic achievement and high expectations.

Teachers test pupils continuously by common classroom-based assessment (Taylor). Attention is on pupil progression. School Climate is marked by discipline where 'law and order' is appreciated (Ringsmose). Pupils are called on to behave pro-socially and to create a positive learning environment (Pressley).

6. Teacher

The school factor 'Teacher for pupils with middle SES' also contributes to the ‘good school’, but the studies synthesised here do not refer to the same activities.

Taylor points to teacher co-operation and communication about pupils’ test scores. Pressley (2007) claims that teachers in ’good schools’ profit by professional development and awareness of the context of high stakes assessments that pupil have to pass.

It has to be noted that no contradictions exist between the recommended activities.

10. Parental Relationship

The 'good school' for pupils with middle SES is a school where:

- Parents have great power in influencing their children’s growth.

Schools employ different means of communication and interaction with the parents. Parents are invited to be active on School Boards, and are given the opportunity to participate in leadership decisions. Schools give parents tips about homework and encourage them to participate in focus groups and surveys to uncover children’s and parents’ needs (Taylor).

4.3.2.3.3 The ’good school’ for pupils with no specified SES and/or gender

The studies grouped together here all look into School Factors and Activities of importance for the creation of a ‘good school’ for Pupil Groups with no specified SES and/or gender indicated (in the following: no specified SES). From Table 4.28 and Table 4.29 we can extract the following Table 4.32.

In the following we look at the activities behind these School Factors.

3. Management and leadership

The 'good school' for pupils with no specified SES is a school where:

Leadership communicates with and consults staff.

There is also significant systematic knowledge sharing (Stringfield, 2008).

School Factors / Studies	3. Management and leadership	4. Curriculum/Scheduling	ture and climate	6.Teacher	7. Support teams	8. Physical environment	10. Parental Relationship	Factor(s) not specified
Mosenthal	X	x	x	X				x
Stringfield	X		x					
Sammons	X		x	X			x	
Sum	3	1	3	2	0	0	1	1

Table 4.32: Factors of importance for Pupil Groups with no specified SES and/or gender

(N= 3)

The 'good school' for pupils with no specified SES is a school where:

- Leadership allots a significant degree of autonomy for teachers to make decisions about how to shape their educational programs.

The 'good school' for pupils with no specified SES is a school where:

- Leadership supports teachers' professional development.

The 'good school' for pupils with no specified SES is a school where:

- Leadership is dynamic, is an 'enabler', a source of ideas (Sammons, 1997).

5. School Culture and Climate

The 'good school' for pupils with no specified SES is a school where:

- Focus is on academic achievement and high expectations.

The ‘good schools’ have a shared vision that all children can succeed (Mosenthal, 2001). Additionally, the ’good schools’ have developed assessment policies that engender a sense of ownership (Sammons, 1997).

The 'good school' for pupils with no specified SES is a school where:

- Staff is focused and working hard with genuine mutual respect.

In this way shared successes and failures create learning communities, i.e. ‘turning schools into knowledge generators’ (Stringfield, 2009).

The 'good school' for pupils with no specified SES is a school where:

- A praise culture exists, where pupils' attendance and punctuality is recorded.

Furthermore, at these school the pupils participate in creating the school regulations (Sammons, 1997).

6. Teacher

The 'good school' for pupils with no specified SES is a school where:

- The quality of teaching is recognised as being of fundamental importance.

Teachers have high expectations and are expert teachers. Their teaching is lively and well-managed, and time is spent on instruction and practice (Sammons, 1997; Mosenthal 2001).

4.4 Direction and strength of the effects examined

The third element of the synthesis consists of going through the factors running across the studies to find the ones that can explain differences in the direction and strength of the studied effect. In this connection the question is also addressed as to why a phenomenon

does or does not have an effect, and whether there are special circumstances that play a part and can explain why an effect is strengthened or weakened in a given context.

In the theoretical model (see Section 4.2) the hypothesis has already been formulated that the direction of the effect goes from school factors towards pupil achievement. At the same time it is noted that the analytical systems model, recognizing context, input, process and output variables, serves quite well as a general frame of reference to determine the position of process-indicators.

4.4.1 Direction and strength of the influence

It is not easy to decide the direction of the influence. And it cannot be excluded that causality, at least partially, may run from pupil achievement to school inputs. Schools and administrations may respond to low pupil achievement by changing school inputs or teachers. The mechanism will act as a suppressor effect, and will lower the observed correlation between school input and pupil achievement compared to the true school input effects. Conversely, teachers and leaders might self-select into schools on the basis of pupil achievement, so that the best teachers and school managers select themselves into better schools, thus generating a spurious correlation between teacher and school management quality.

For example, as mentioned on page 94, Educational Leadership as a subcategory of the school factor Management and Leadership was found to be negatively related to effectiveness, meaning that less effective schools manifested more educational leadership. This result could be interpreted to say that less effective schools cause more educational leadership, not the other way round. This turns the tables concerning what is presupposed in this systematic review.

There is a very good reason why it is difficult to decide the direction of the influence. All the primary quantitative studies investigate cor-

relations, not causality. The argument for a certain causal effect is thus based on theoretical preconceptions and not on causal empirical research.

Furthermore, it has not been possible in this synthesis to measure the strength of the influences exerted by the school factors and their subcategories. As shown in Section 8.1, the data available make it impossible to carry out a meta-analysis on even some of the factors or subcategories included in the primary studies in the review. The reason is the heterogeneity of the studies. It has thus in a number of cases been impossible to identify the content of the variables and measures applied. This also has the consequence that the data has not allowed us to calculate an effect size for the school factors and subcategories. Instead, the synthesis of the quantitative studies is complemented by a counting of significances. In consequence we cannot rank the relative importance of the various a school factors and subcategories.

4.4.2 The significance of context

Research into effective schools is based on a theory that the results achieved by a school are derived from (a) the individual abilities of the pupils, (b) the cultural, socio-economic and family background of the pupils and (c) what the pupil experiences at the school. Effective schools research seeks information about factor (c), and attempts to control and correct any influences arising from the other two factors, cf. p. 29 .

In the research mapping (Section 3.1 and 3.2) possible contexts have been noted that could be considered to have significance for an assessment of the direction and strength of the effect. Table 3.1, Table 3.2, Table 3.3, Table 3.5, and Table 3.6 respectively show the studies' distribution by country, language used in the research reports, the educational setting of the studies, the curriculum area of the studies,

and the specific group of pupils that have been examined. In this context it is worth mentioning that it has not been possible to show that studies that include data from Nordic countries, give conclusions that differ from studies based on data from non-Nordic countries.[3] Added to these are the contexts that have been brought to light by the reading of the mapped studies. As mentioned in Table 3.12, most studies (97 out of 109) give an adequate description of the context.

It must also be noted that the very concept of the 'good school' from the outset should be considered to be both politically controversial and dependent on cultural context. However, as shown above (page 80), the studies employ only two different definitions of the 'good school': firstly a school with 'high academic achievements' and secondly a school with 'high non-academic achievements'. However, since the two definitions in most cases relate to the same studies, it would seem reasonable to suppose that the two definitions are not mutually exclusive, but rather supplement each other, i.e. the 'good school' is in general understood in terms of 'high pupil achievements'.

Can something be said about the significance of context for the present systematic review? One of the primary studies included in the systematic review (see Rumberger & Palardy fig. 1, p. 11) provides a helpful figure about the relationships between context, school processes and pupil experience, and school outputs and pupil outcomes, by creating a conceptual framework for analysing school performance, cf. Figure 4.3 .

[3] A similar, resently published Swedish study reaches the same conclusion: "In broad outline, the same factors appear as significant when we compare results from Swedish and international studies" (Skolverket, 2009, 33).

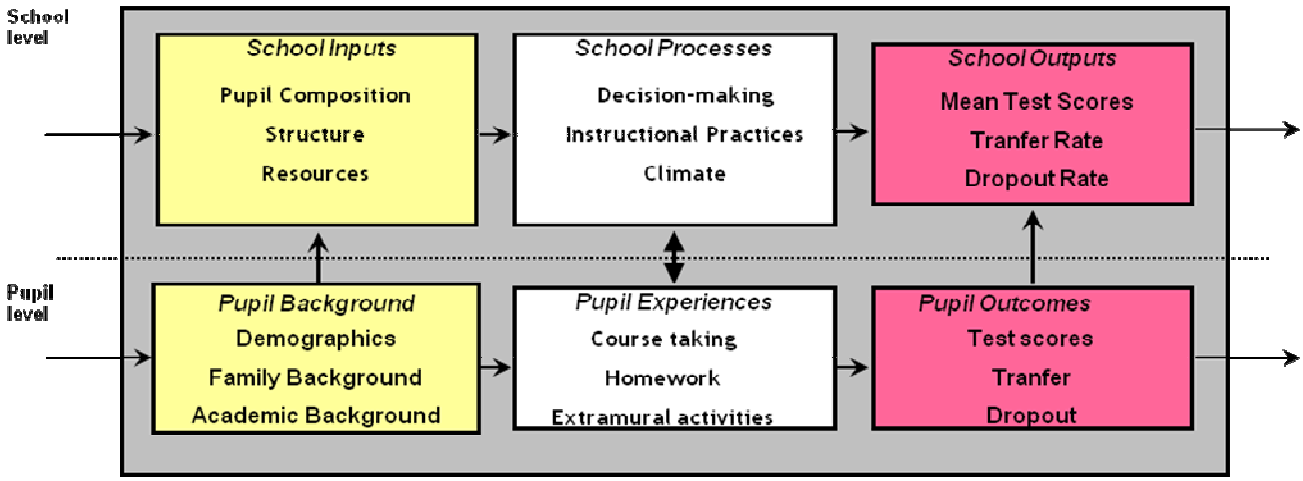

Figure 4.3: Conceptual Framework for Analysing School Performance (Source: Rumberger & Palardy, Am. Educ. Research J. 2005, fig. 1, p.11)

If we apply this conceptual framework to the present systematic review, we get Figure 4.4. In the figures, the yellow boxes represent Pupil Background and School Inputs, the content of which shapes the empirical environment for both School Processes and Pupil Experiences, represented by white boxes, resulting in school outputs and pupil outcomes, represented by red boxes.

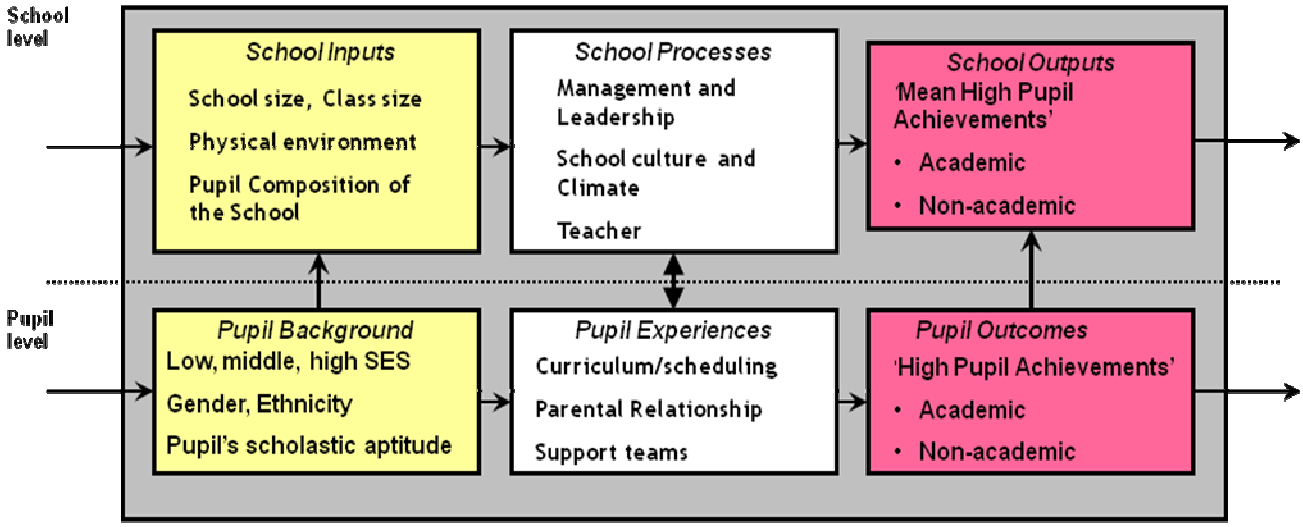

Figure 4.4: Conceptual framework for the present systematic review

It has been stated in the systematic review that factors relating to background, such as the teacher's gender, age, socio-economic back-

ground and ethnicity, by and large play a part in the mapped studies. Some studies especially bring up teacher's gender and ethnicity as important for specific groups of pupils. In this context we call attention to the fact that one of the reasons for the exclusion of a study was the following screening criterion: 'When none of these three criteria are part of the study design: 1. Control is present for differences in pupils' socioeconomic background; 2. Control is present for differences in pupils' scholastic aptitude; 3. A pre(-post) is present. When one criterion is found the study must be included', cf. Table 2.2, p. 41. As it appears from Figure 4.4, we have in fact included the following contextual factors, if possible, in the systematic review:

- Low, middle, and high SES
- Gender
- Ethnicity
- Pupil's scholastic aptitude

We have applied this information in the studies focusing on the pupils' socioeconomic background to carry out the syntheses for pupils belonging to separate SES groups, cf. the analyses in Sections 4.3.1.2 and 4.3.2.2. It has not been possible to show in this systematic review that these contexts call for special educational treatment for specified Pupil Groups.

4.5 The robustness of the narrative syntheses

In the fourth element of the synthesis process, an attempt must be made to assess robustness. This involves three aspects, as mentioned above (Section 4.1): the *methodological quality of* the primary studies, the *methods* employed in the synthesis, and the amount of *information* about the primary studies that has led to their inclusion in the systematic review. These three aspects will be examined individually in the following. It must be pointed out, however, that in

Section 3.4 above we have already made an overall quality assessment of the studies reviewed, with special emphasis on the reporting quality and contribution of evidence.

4.5.1 Methodological quality of the primary studies

The first aspect concerns the methodological quality of the primary studies. Table 3.10 shows the research designs used by the studies in question. In this systematic review, studies are assigned 'high', 'medium' or 'low' weight of evidence, see Table 3.21. The overall weight of evidence of the individual study — as already described — is based on a combined evaluation of the credibility of the research and of the reporting, the relevance of the study's purpose, and the extent to which the chosen research design and analysis are appropriate for supplying an answer to the review question. It should be added that the overall weight of evidence of a study may well be higher or lower than the weight of evidence of the individual aspects of the assessment. After removing studies with 'low' overall weight of evidence, the summary of research designs used in the studies included in the syntheses is as indicated in Table 4.33.

Attribute	Number
Case study	11
Case-control study	5
Cohort study	20
Comparative study	4
Cross-sectional study	33
Ethnography	6
Random experiment with random allocation to groups	1
Experiment with non-random allocation to groups	5
Methodological study	2
One group pre-post test	2
Secondary data analysis	38
Views study	32

Table 4.33: Distribution of research designs of the studies used in the syntheses (N = 63; multiple categories per study permitted)

Of the 109 studies included in this systematic review, 71 (65 %) remain after exclusion of the studies with low overall weight of evidence. After further exclusion of the studies with no measurement of significance there remain 63 (58 % of all included studies; 89 % of all included studies with high or medium overall weight of evidence). Of these 63, 17 are considered to have an overall high weight of evidence, and 46 are considered to have an overall medium weight of evidence. The preliminary syntheses have wherever possible been based primarily on studies with overall high weight of evidence, but since some of the relationships of interest have only been addressed

in a small number of studies, it has not always been appropriate in such cases to discriminate on the basis of overall weight of evidence.

The assessments already presented in Sections 3.3 and 3.4 are relevant in connection with this first aspect of the methodological quality of the studies. In addition there is the following consideration: on the basis of Petticrew & Roberts (2003; 2006, p.30), Rieper & Foss-Hansen (2007, p.79 fig.7.1) constructed an evidence typology concerning the relationship between research question and research design. This typology indicates that for review questions concerning effects studies, randomised controlled trials (RCTs) command the highest weight of evidence, followed by cohort studies and quasi-experimental studies. In the 63 studies examined here, there are 1 RCT, 5 experiments with non-random allocation to groups, 20 cohort studies and 2 studies using a quasi-experimental design.

A group of studies included in the qualitative part of the synthesis demands a special comment. Many of these primary studies are outlier studies in the form of case studies. By making a detailed study of schools with exceptional or unexpectedly high pupil achievements it is assumed that either important school inputs or school processes (or both in combination) can be uncovered as the cause(s) of these results. However, measures must be taken to ensure that researchers do not just 'see' factors which in fact reproduce private or public assumptions, and do not include factors as explanations for observed effects that on closer analysis turn out to be concomitant, not causal. We have therefore treated the results from these studies with caution and evaluated the results from the quantitative part of the synthesis as being more reliable.

4.5.2 Method in synthesis creation and weight of evidence

If we now turn to the method employed in creating the seven groups of syntheses, and the weight of evidence that the various syntheses

have been assigned, we arrive at the following results – see Table 4.34.

Synthesis		High	Medium	Total	Relative % weight
1	Quantitative Academic No specified	13	29	42	31 %
2	Quantitative Non-academic No specified	3	6	9	33 %
3	Quantitative Academic Low SES	1	11	12	8 %
4	Quantitative Non-academic Low SES		1	1	0
5	Qualitative Academic No specified		3	3	0
6	Qualitative Academic Low SES	1	9	10	10 %
7	Qualitative Academic Middle SES	1	2	3	33 %

Table 4.34: Distribution of 'high' and 'medium' weight of evidence in the various syntheses

(N=63 studies; multiple categories per study permitted. The right-hand column: "relative weighting" shows the percentage of studies with high weight of evidence within the total number of studies covering the aspect in question)

When developing the syntheses, the studies with 'high' weight of evidence were the first to be considered. Subsequently studies with 'medium' weight of evidence were brought in where this was possible. As can be seen from the figures, synthesis 1 of quantitative studies about a 'good school' for Pupil Groups with no specified SES, where

the ‘good school’ is defined as a school with ‘high academic achievement’, presents the greatest weight of evidence, followed by synthesis 3 of quantitative studies about a ‘good school’ for Pupil Groups with low SES, where the ‘good school’ is defined as a school with ‘high academic achievement’. Notice that there has been no attempt to create syntheses of groups 3, 4 and 5.

It was noted on p. 65 that the data do not permit us to carry out a meta-analysis on school factors or subcategories. The reasons for this have been given in Section 8.1. Instead, another procedure is applied based on the power calculation table in Table 9.1, Chap. 9: Appendix 4. By assessing whether a synthesis is possible or not on the background of how positive and negative significance, insignificance and intractability are distributed over the individual school factors or subcategories, a more firm reason is given for the syntheses that have been conducted.

This method has certain problems, however. The main one is the problem of publication bias. Whereas primary studies that report ‘positive’ results are more likely to find a publisher, studies that report insignificant results, where ‘positive’ was hoped for, are more likely to remain unpublished. A consequence for the method applied to the quantitative studies in this systematic review could be that the number of primary studies reporting insignificant results is underestimated for each individual school factor/subcategory.

However, if we look at the results reported, this risk can with good reason be considered to be smaller than immediately expected. If we look at the data in Table 4.21, for instance, where n = 14 and m = 7 for the school factor ‘Parental Relationship’ and keep m = 7, more than 56 insignificant primary studies would have to be reported to make the school factor insignificant, i.e. 49 more than in fact were

found in the research mapping according to the power calculation table in Chapter 9.

4.5.3 Information about the primary studies

If we finally turn to the third aspect concerning the robustness of the study, this has to do with the degree of *information* about the primary studies that has led to their inclusion in the systematic review. In chapters 2 and 7, Appendix 2, detailed descriptions are given of this review's conceptual delimitations, search profiles and techniques, inclusion and exclusion principles when screening, and methods for extracting data from the studies that were selected.

An important consequence of the inclusion criteria chosen is that no factor reporting the influence of economic resources or individual factor studied is included. The reason for this choice is that we have concentrated the main interest in mapping and synthesising primary studies which only look at the outputs and outcomes that can be ascribed as the results of various factors interacting with the contribution of the school itself.

4.5.4 Concluding evaluation

The examination of the statistical significance of the different schools conducted in chap. 8 concludes that there is an indication that the School Culture and Climate factor is more significant on average than other factors, and that the Management and Leadership factor is less significant than other factors. The figures do not allow inclusion of the school factors School size, Class size, Support teams, Physical environment, Pupil Composition of the School, and Parental Relationship in this analysis.

However, syntheses covering the quantitative research of the last 20 years or so within the area of interest do seem to indicate that the

more general features and tendencies in the research can be identified.

On the one hand it is striking that the research gives a fairly consistent picture, presented in the following Section 4.6. On the other hand, it should not be overlooked that a mapping of the last two decades of research in a given area can also be interpreted as a reflection of the prevailing professional opinions and expectations of the researchers and commissioners of research within the area in question. Viewed in this light, the research that has been reviewed gives a picture of what researchers and those commissioning the research considered it worthwhile to study, and which frames of reference and answers they considered fruitful.

4.6 Concluding remarks – school factors across syntheses

The synthesis carried out in Sections 4.3.1.3 and 4.3.2.3 for quantitative studies and qualitative studies respectively, has been divided into seven individual syntheses. This was done to make sure that Pupil Groups that perhaps require different educational treatments, and studies with different definitions of a 'good school' understood as either 'high academic achievement' or 'high non-academic achievement' were kept apart in the synthesis.

Each of the seven syntheses discusses the ten school factors and their subcategories if the underlying primary studies make it possible. In this section we try to sum up the results for each of the ten school factors and their possible subcategories across the seven syntheses, cf. Table 4.34. Results from Section 4.3.1 (Group 1: Synthesis of the quantitative studies) are the main guide. Results from Section 4.3.2 (Group 2: Synthesis of the qualitative studies) are considered as supplementary.

1. School size

This factor is addressed in three high weight of evidence studies and 16 medium weight of evidence studies.

In the light of the primary studies included in this synthesis a sufficient number show relationships with significance for the creation of a 'good school' for Pupil Groups with no specified SES.

School size varies greatly from country to country. As the concept of 'school size' in the studies is defined in relation to the average school size in a country,it is concluded *that the concept 'school size' is applied inconsistently.Therefore, no conclusion is warranted concerning this factor*.

2. Class size

This factor is addressed in two high weight of evidence studies and 13 medium weight of evidence studies.

In the light of the primary studies included in this synthesis a sufficient number show relationships with significance for the creation of a 'good school' for Pupil Groups with no specified SES.

As it, however, has been noted that the studies included have not controlled for unbiased sampling to school classes *no conclusion is warranted concerning 'class size'*.

3. Management and Leadership

This factor is addressed in eight high weight of evidence studies and 33 medium weight of evidence studies. It is divided into four subcategories.

Looking across the studies, a considerable number show relationships with significance. They indicate that the subcategories 'Human Resources' and 'Educational Leadership' of the complex school factor Management and Leadership are of importance for the creation of a

'good school' for Pupil Groups with no specified SES. The examination of the statistical significance of the different school factors conducted in chap. 8 concluded that the Management and Leadership factor is less significant than the other complex school factors, i.e. 4. Curriculum/scheduling, 5. School Culture and School Climate, and 6. Teacher.

The subcategory 3.1 *Human Resources* shows relationships with significance for the 'good school' understood as 'high academic achievement' for Pupil Groups with no specified SES. It covers three main aspects, i.e. the principals' years of experience, hours spent working and his or her availability for the teachers. It also covers the principal's policy concerning the teachers' growth, and influence on hiring and firing staff. Finally, it looks into the influence of members of the school organisation, like teachers and the principal, but especially the parents, on the decisions of the school board.

The subcategory 3.3 *Educational Leadership* shows relationships with significance for the 'good school' as 'high academic achievement' for Pupil Groups with no specified SES. This category covers situations where the principal demonstrates strong leadership, above all in the areas of curriculum and instruction, and is able to involve other staff members in leadership activities and position, where the principal's behaviour is supportive and egalitarian and neither directive nor restrictive, and where the principal is 'resource supportive', e.g. in deciding textbooks and contents of the teaching.

4. Curriculum/scheduling

This factor is addressed in 11 high weight of evidence studies and 28 medium weight of evidence studies. It is divided into three subcategories.

The subcategory 4.1 *Opportunity to Learn* of the complex school factor Curriculum/scheduling shows in particular relationships with significance for the 'good school' understood as 'high academic achievements' for Pupil Groups no specified SES.

It covers number of teaching hours, including homework hours. It could be seen as the teacher's efficiency of organising the instruction process, measured by the percentage of time teachers reported spending on the planning of their lessons for the following day, the making of weekly teaching plan, keeping to the timetable, and by the assigned time spent on lessons. It also includes homework hours, which are the total hours that pupils spent on homework both in school and out of school per week.

5. School Culture and School Climate

This factor is addressed in 13 high weight of evidence studies and 42 medium weight of evidence studies. It is divided into four subcategories.

Looking across the studies, a considerable number show relationships with significance. This indicates that the complex school factor School Culture and School Climate is of importance for the creation of a 'good school' for Pupil Groups no specified SES. The examination of the statistical significance of the different school factors conducted in chap. 8 concluded that the School Culture and Climate factor is more significant on average than any of the other complex factors.

The subcategory 5.1 *Disciplinary Climate* shows relationships with significance for the 'good school' for Pupil Groups with no specified SES. It covers a school where an orderly atmosphere prevails, and also an ordered environment in which appropriate pupil behaviours are present. A 'good school' for pupils with no specified SES is a school where pupils do not feel unsafe, since the proportion of pupils

who feel unsafe has a significant negative effect on pupil achievement.

The subcategory 5.2 *Achievement/progress Orientation* shows relationships with significance for the 'good school' for Pupil Groups with no specified SES. It appears to be the most important subcategory for creation of a 'good school'. It covers a school which focuses on academic achievement and high expectations, where high pupil engagement exists and these is absence of negative peer pressure; and where teachers rate attentiveness and have established a 'learning climate'.

The subcategory 5.3 *Interrelational Climate* shows relationships with significance for the 'good school' for Pupil Groups with no specified SES. It covers affiliation, support/respect for staff and pupils, and warmth in teacher/pupil relationships: teachers can obtain assistance, advice and encouragement, and are made to feel accepted by their colleagues. Pupils develop positive relationships.

The subcategory 5.4 *Social norms and values* shows relationships with significance for the 'good school' for Pupil Groups with no specified SES. It covers teachers' professional values like interest in their work and professional development, and an interest in new educational plans and experimentation, classroom openness and individualisation. Teachers appreciate taking a full participation in school activities, feeling ownership of what happens in the school and accept that a work pressure dominates the school environment. Students enjoy class work, and they are involved and think that they are doing a good job in classes. Physical features of rooms, equipment, and buildings are maintained and kept orderly.

6. Teacher

This factor is in total treated in 15 high weight of evidence studies and 48 medium weight of evidence studies. It is divided into five subcategories.

Looking across the studies, a considerable number show relationships with significance. This indicates that the complex school factor Teacher is of importance for creation of a 'good school' for Pupil Groups with low and no specified SES.

The subcategory 6.1 *Teacher behaviour* shows in particular relationships with significance for creation of a 'good school' for Pupil Groups with low and no specified SES. It covers a number of aspects that relates to the teachers' planning and teaching methods.

The subcategory 6.5 *Teacher as an Organisational Actor* shows relationships with significance for creation of a 'good school' for Pupil Groups no specified SES. It covers a number of aspects connected to teacher's role of the school organisation, e.g. formal competence, average years of experience, average years of education, staff job satisfaction, teacher engagement, teacher cooperation, composition of sex (number of female teachers in the school), and number of teachers with in-service training.

7. Support Teams

This factor is in total treated in 12 medium weight of evidence studies.

In the light of the primary studies included in this synthesis it appears that it cannot be established that a variation in the school factor Support Teams is of importance for the 'good school'.

8. Physical environment

This factor is in total treated in 7 medium weight of evidence studies.

In the light of the primary studies included in this synthesis it appears that it cannot be established that a variation in the school factor Physical Environment is of importance for the 'good school'.

9. Pupil Composition of the School

This factor is in total treated in 6 high weight of evidence studies and 15 medium weight of evidence studies.

Looking across the studies, a considerable number show relationships with significance. This indicates that the school factor Pupil Composition of the School is of importance for creation of a 'good school' for Pupil Groups with no specified SES.

This conclusion calls for an elaboration. On the one hand, it is well known that the pupils' socio-economic background plays a key role in pupil achievement, but that aspect does not form part of this systematic review. On the other hand, it is conceivable that e.g. politicians decide to adjust the proportion of low SES, bilingual or ethnic groups in the individual schools. In this case the factor can be considered 'malleable'. In keeping with the last interpretation the school factor is considered malleable and supported as important for creation of the 'good school'.

10. Parental Relationship

This factor is in total treated in five high weight of evidence studies and 18 medium weight of evidence studies.

Looking across the studies, a considerable number show relationships with significance. This indicates that the school factor Parental Relationship is of importance for creation of a 'good school' for Pupil Groups with no specified SES.

The factor has to be considered in this context: Schools employ different means of communication and interaction with the parents.

Parents are invited to be active on School Boards, and are given the opportunity to participate in leadership decisions. Schools give parents tips about homework and encourage them to participate in focus groups and surveys to uncover children's and parents' needs. Parents' support of children and involvement in school matters and community partnership are important to the 'good school' defined as a school with 'high academic achievement'.

Summary

Figure 4.5 summarizes the result of this systematic review. The white boxes contain the 'malleable' school factors/subcategories that– in the sense indicated – have been shown to be of importance for creating a school with 'high academic achievement' for Pupil Groups with low and/or no specified SES.

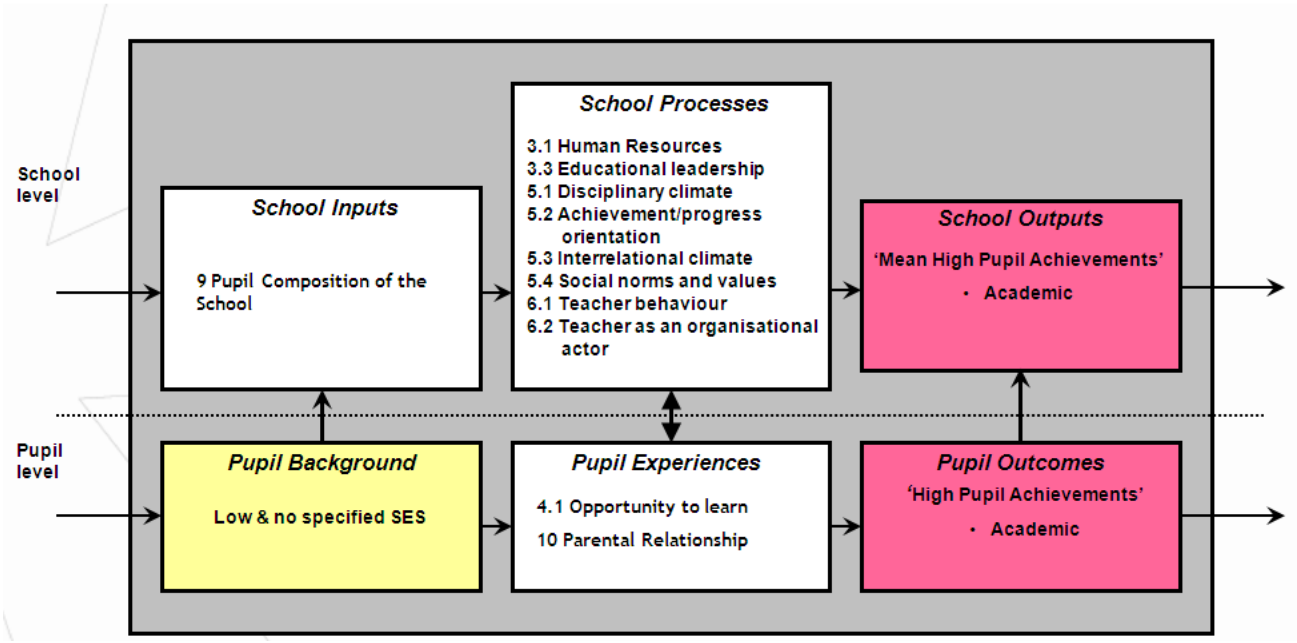

Figure 4.5: Summary of school factors/subcategories of importance for the 'good school'

It is worth noting that the box 'School Inputs' in this figure has changed role from context to 'malleable' school factor.

5 Conclusion

5.1 The results of the systematic review

In this systematic review the following systematic review questions are answered:

> *What empirical research has been carried out to examine the relationship between factors in primary and lower secondary schools (inputs and processes) and the learning achieved by primary and lower secondary school pupils (outputs and outcomes)?*
>
> *What are the results with weight of evidence of this empirical research?*

The answer is provided by conducting a research mapping and a narrative synthesis on the basis of the last 20 years' empirical pedagogical and educational research in school effectiveness.

The result is that the following 11 school factors and subcategories are important for creating a 'good school', where the 'good school' is defined as a school with 'high pupil achievements'. In the list a subcategory is followed by the name of the school factor to which it belongs. The list is not prioritized:

- Human Resources (Management and Leadership)
- Educational Leadership (Management and Leadership)
- Opportunity to Learn (Curriculum/scheduling)
- Disciplinary Climate (School Culture and School Climate)
- Achievement/progress Orientation (School Culture and School Climate)
- Interrelational Climate (School Culture and School Climate)
- Social norms and values (School Culture and School Climate)

- Teacher behaviour (Teacher)
- Teacher as an Organisational Actor (Teacher)
- Pupil Composition of the School
- Parental Relationship

In the report's Sections 4.3 and 4.6 an account is given of the details of this answer. The following riders can be added about the strength of this assertion:

The answer is based on the best evidence available from pedagogical and educational research in the period 1990-2008

The answer is based upon a research mapping and a research assessment of that research, with particular attention to the School Effectiveness Research

The answer has been arrived at by undertaking syntheses generated on the basis of a data extraction carried out by a review group and Clearinghouse.

The answer invites the following comments:

- The answer is of interest both in terms of what it directs attention towards and in terms of what it does not direct attention towards. It does not highlight factors that are not already familiar. But it indicates that it is precisely these factors — and not others — that according to our best evidence have been shown to be of importance for creation of the 'good school' at a school with 'high academic achievement'.
- The answer conceals the complexity of the individual school factors and subcategories. To create process indicators on the basis of the school factors revealed in the synthesis, it is necessary to uncover the conceptual and empirical landscape behind the

enumerated school factors. The report summarizes this conceptual and empirical landscape.

- Due to the method of synthesis it has not been possible to indicate the relative weight of the individual factors/subcategories. Neither has it been possible to calculate effect size.
- Some school factors (School size and Class size) have measurements that do not permit a precise interpretation, since for instance an average school size is relative to the country studied.

5.2 Recommendation for practice, policy and research

In conclusion consideration should be given to the recommendations for practice, policy and research that derive from the results of the systematic review carried out here.

5.2.1 Practice

The school leader should realize that a number of aspects of his or her work are important for pupil learning: the more he or she is available for teachers the better; the more the principal's policies are concerned with teachers' growth, the better; the more teachers and parents are involved in school decisions the better. The principal should demonstrate strong leadership, above all in the areas of curriculum and instruction, and be able to involve other staff members in leadership activities and position. The principal's behaviour ought to be supportive and egalitarian and neither directive nor restrictive, and should be 'resource supportive', e.g. in deciding textbooks and contents of the teaching.

The teacher's efficiency of organising the instruction process, measured by the percentage of time teachers reported spending on the planning of their lessons for the following day, the making of a weekly teaching plan, keeping to the timetable, and by the assigned time spent on lessons, improve pupil learning. It also includes

homework hours, which are total hours pupils spent on homework both in school and out of school per week.

In a good school an orderly atmosphere prevails, and also an ordered environment in which appropriate pupil behaviours are present. A good school for pupils is a school where pupils do not feel unsafe since the proportion of pupils who feel unsafe has a significant negative effect on pupil achievement.

A good school focuses on academic achievement and high expectations, high pupil engagement exists and negative peer pressure is absent. Teachers rate attentiveness and have established a 'learning climate'.

In a good school staff and pupils show affiliation and support/respect, and a warm teacher/pupil relationship exists. Teachers can obtain assistance, advice and encouragement and are made to feel accepted by their colleagues. Pupils develop positive relationships with each other.

Good schools employ various means of communication and interaction with the parents. Parents are invited to be active on School Boards, and are given the opportunity to participate in leadership decisions. Schools give parents tips about homework and encourage them to participate in focus groups and surveys to uncover children's and parents' needs. Parents' support of children and involvement in school matters and community partnership are important.

5.2.2 Policy

Policymakers can influence pupil learning through choices of school size, class size, and the pupil composition of the school.

Policymakers can promote pupil achievement by helping to identify strengths and weakness in school by developing indicator systems for malleable school factors and subcategories.

5.2.3 Research

Although research in 'the good school' to a certain extent is based on high quality data and sophisticated statistical models, taking into account the fact that data is sampled as clusters (students within classes and classes within schools) and thus reporting the correct standard errors, it is also evident that no studies in this review seriously address causality in terms of using experimental or quasi-experimental data, or statistical methods that allow for causal interpretation. It seems that there is a complete lack of interest in establishing causal directions in "good school" research. Referring to the problems noted in section 4.4 regarding the causal direction of school inputs it is clearly crucial that future research takes causality more seriously. There is a huge and growing interest in causal effects in the economics of education, see e.g. the overview by Blundell and Costa Dias (2007). This research makes extensive use of statistical methods and points to how one can use "natural" experiments to establish the causal effect of various school inputs, such as teacher characteristics and school resources. Taking causality seriously also means that new requirements must be made concerning the data, requirements that are not always met by the existing data sources. Thus the research community must also convince policy makers that there is a need for a new causal agenda in school research.

6 Appendix 1 – An example of data extraction

6.1 EPPI-Centre tool for education studies V2.0 — editable version

Item: Van Damme, Jan; De Fraine, Bieke; Van Landeghem, Georges; Opdenakker, Marie-Christine; Onghena, Patrick (Dec 2002) A Study on Educational Effectiveness in Secondary Schools in Flanders: An Introduction.

Section A: Administrative details

A.1 Name of the reviewer	Details Jaap Scheerens
A.2 Date of the review	Details 23rd of January 2009
A.3 Please enter the details of each paper which reports on this item/study and which is used to complete this data extraction.	Paper (1) Journal article Unique Identifier: *1650885 Van Damme* Authors: Jan Van Damme; Bieke De Fraine; Georges Van Landeghem; Marie-Christine Opdenakker; Patrick Onghena. Title: A Study on Educational Effec-

	tiveness in Secondary Schools in Flanders: An Introduction. Paper (2) With respect to filling in this EPPI-reviewer, three articles are considered: Unique Identifier: Paper 1: 1654438 (The effect of schools and classes on language achievement) Paper 2: 1650884 (The effect of schools and classes on mathematics achievement) Paper 3: 1650883 (The effect of schools and classes on noncognitive outcomes) Authors: Paper 1: Jan Van Damme; Bieke De Fraine; Georges Van Landeghem; Marie-Christine Opdenakker; Patrick Onghena. Paper 2: Jan Van Damme; Bieke De Fraine; Georges Van Landeghem; Marie-Christine Opdenakker; Patrick Onghena. Paper 3: Jan Van Damme; Bieke De Fraine; Georges Van Landeghem; Marie-Christine Opdenakker; Patrick Onghena.

	Title: Paper 1: The effect of schools and classes on language achievement. Paper 2: The effect of schools and classes on mathematics achievement. Paper 3: The effect of schools and classes on noncognitive outcomes.
A.4 Main paper. Please classify one of the above papers as the 'main' report of the study and enter its unique identifier here.	Unique Identifier: *1650885 Van Damme*
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.6 If the study has a broad focus and this data extraction focuses on just one component of the study, please specify this here.	Specific focus of this data extraction (please specify) The study consists of public (state-run) schools and private (catholic schools, city school) schools. The data extraction is focused on the public schools.
A.7 Language (please specify)	Details of Language of report *English*

Section B: Study Aims and Rationale

B.1 What are the broad aims of the study?	Explicitly stated (please specify) Our aim is to search for interesting correlates (student, class, or school characteristics) of the noncognitive outcomes at he end of the second grade that have some predictive power on top of the background characteristics and the group composition variables.
B.2 What is the purpose of the study?	A: Description *Please edit*
B.3 When was the study carried out?	Explicitly stated (please specify) *1991-1992*
B.4 What are the study research questions and/or hypotheses?	Explicitly stated (please specify) Research question with respect to the noncognitive outcomes; The first research question inquires about the 'raw' effect of the school and of the first-and second grade class on the noncognitive outcomes; The second research question asks about net effects of the secondary school and classes; Our third objective is achieved simultane-

	ously, namely a description of the relationships between the student's background characteristics and the noncognitive outcomes at the end of the second grade. A more strict brand of net effects is defined by additionally controlling for group composition in terms of the background characteristics (fourth research questions) Research question with respect to the outcomes in mathematic: (1)How important is the class/teacher and the school in explaining differences in mathematics achievement at the end of the second grade.? (2)To what extent are differenced between classes within schools with respect to mathematics achievement attributable to differences in student intake? (3)What characteristics of classes, teaching practice, and schools are linked to the mathematics achievement of students? (4)Are there indications of differential effectiveness of classes or schools? Language: 1: Does a student's language

	achievement depend upon the school where he or she is taught? 2; Are class and teacher more important than school with regard to the language achievement? 3; To what extent are differences between schools and between classes a result of differences in student intake? 4; What characteristics of schools and classes are linked to the language achievement of the students? 5; Are some schools or classes more effective for particular groups of students (with respect to students' ethnic background, gender and ability)?

Section C: Study Policy or Practice Focus

C.1 What is the curriculum area, if any?	N/A (not on a specific curriculum area) *Please edit* Literacy - first languages *Please edit* Maths Please edit Coding is based on: Authors' de-

	scription *Please edit*
C.2 What is/are the educational setting(s) of the study?	Lower secondary school *Please edit* Secondary school *Please edit*
C.3 In which country or countries was the study carried out?	Explicitly stated (please specify) *Belgium*
C.4 Please describe in more detail the specific phenomena, factors, services or interventions with which the study is concerned.	Details The effect of schools and classes upon mathematics + literacy achievement in the second grade of second education is addressed. The reported data in this contribution stem from the LOSO-project.

Section D: Phenomena/Factors in School Addressed in the Study

D.1 Which phenomena/factors in school are addressed in the study?	School culture Socio-economic composition of the pupils in the schools Other Please edit other variables addressed are gender, classroom and school climate,(to be seen

	as aspects of school culture), performance feedback and ability grouping

Section E: Pupil Result Focus

E.1 Are academic effects involved?	Yes, achievement performance. Please specify mathematics achievement at secondary school level Dutch language achievement in Belgium secondary schools
E.2 Are non-academic effects involved?	Yes, psychical. Please specify -the degree to which students feel at home in the school environment, - the extent to which the student does his/her best for the school work; - academic self concept - social integration in the class
E.3 Does the study focus on effects on a specific group of pupils?	Yes, pupils with low SES Yes, boys

Section F: Actual sample

F.1 Who or what is/ are the sample in the study?	Schools (please specify)

	In total: 57 secondary schools With respect to the study of the mathematics achievement: A subsample of 2,552 students following the general track in the second grade of secondary education (and who did not have to retake the first grade), belonging to 150 mathematics classes and to 57 secondary schools, was used. Teachers (please specify) In every school, a representative sample of 15 teachers in the first cycle completed a school characteristics questionnaire. Language: in the large sample there are 155 Dutch teachers and 275 second grade classes, so some teachers taught more than one class. Pupils (please specify) *6411 students*
F.2 What was the total number of participants in the study (the ACTUAL sample)?	Explicitly stated (please specify) With respect to the study of the mathematics achievement: A dataset of 2105 students belonging to 147 classes and 56 schools

	was available. Analyses on the effect of class and school variables are based on a dataset with 1,119 students, 74 classes and 33 schools. The analysis with the relevant student, class, and school variables included in the model of mathematics achievement is based on a dataset with 1,936 students, 131 classes and 47 schools. With respect to the study of the noncognitive outcomes: The LOSO-cohort of 6,411 students who started secondary education in September 1990 contains a subsample of 4,759 students in 57 schools who were enrolled in the first grade in 1990-1991 (in 291 classes) and in the second grade of the general track in 1991-1992 (276 classes). Of those 4,759 students, 150 changed schools during their first 2 years in secondary education. Our analyses are based on the remaining 4,609 students. With respect to the study of language achievement: A subsample of 2569 students in 152 Dutch

	classes in 55 schools.
F.3 Please specify any other useful information about the study participants.	Details With respect to the study of the mathematics + language achievement and noncognitive outcomes: Five student-level explanatory variables, measured at the start of the first grade, were used: initial cognitive ability (COGN), socioeconomic status of the family (SES), achievement motivation (AM), immunity to stress (STRESSIMM), and sex and language spoken at home (DUTCHHOME). Prior mathematics achievement is based on a dataset with 1,936 students, 131 classes and 47 schools. Specifically to the noncognitive outcome the age of the start of secondary education were also described. At class level, the following group composition variables were used: mean initial cognitive ability (CL-COGN), mean SES (CL-SES), mean achievement motivation (CL-AM), mean immunity to stress

	(CL-STRESSIMM), proportion of girls in the class (CL-SEX), and proportion of students who speak Dutch at home (CL-DUTCHHOME). At the school level, group composition variables comparable with the described group composition variables at the class level were used to describe the student population of a school.

Section G: Results and conclusions

G.1 What are the results of the study as reported by the authors?	Details The study stands out for its attention for school and class composition effects, in terms of SES, initial cognitive ability, achievement motivation, language spoken at home and the age at the start of secondary education. These variables are used as individual student background control variables, but also as compositional effects at class and school level. Results are presented in three areas: mathematics, Dutch language

	and non-cognitive outcomes. Mathematics - a learning climate that is focused on learning and cohesive with a teacher that has positive expectations towards the achievement of students (the effects of other school and classroom process variables like opportunity to learn, disappeared after taking the compositional effects into consideration; climate was also overlapping with composition in terms of average SES and average initial ability in explaining mathematics achievement) - paying attention to differences between students had a negative effect - consultation between teachers had also a negative effect NO effects of structured teaching and performance feed back - ability grouping was found to be positive for low achieving students Language (also for language high effects of student background variables and compositional effects; specifically a pronounced effect of gender composition: a high proportion

	of girls is indicative of high achievement); - learning climate was the only school factor that had an effect over and above the individual background variables and compositional effects Non cognitive outcomes +Learning climate had a positive significant effect on three of the four non cognitive effect measures: environment, work, and peers - Feedback had a negative significant effect on self image
G.2 What do the author(s) conclude about the findings of the study?	Details Please edit The study provides a wealth of information on school composition effects, particularly the effects of mean initial ability level and mean SES (at school and classroom level). Very interestingly are the interaction or joint effects of school composition and school climate. AN orderly work oriented climate was the most relevant school factor, operational at classroom and at school level.
G.3 Which answer(s) does the	Please specify

study offer to the review question?	The substantive school factor that had the expected positive effect on all outcome variables was a work oriented school climate. Relevant is also that strong candidates from other studies like: structured teaching, opportunity to learn and performance feedback had none or negative effects.

Section H: Study Method

H.1 Study Timing	Prospective *Please edit*
H.2 What is the method used in the study?	Cohort study *Please edit*
H.3 Study design summary	Please specify The mathematics and the language (i.e., Dutch) achievement was measured by means of curriculum relevant multiple choice tests at the start of the secondary school and at the end of the first, the second, the fourth, and the sixth grade. As already mentioned, the LOSO-cohort was followed through secondary school,

	but also afterwards. This makes it possible to consider another type of effectiveness criterion by studying the effects of secondary schools upon dropout. Data about the students' primary school career were also collected because some school effectiveness studies indicate that the primary school can have long-term effects upon achievement in secondary school. They are though not considered in these articles.

Section I: Methods-groups

I.1 If Comparisons are being made between two or more groups*, please specify the basis of any divisions made for making these comparisons	Not applicable (not more than one group) *Please edit*
I.2 How do the groups differ?	Not applicable (not in more than one group) *Please edit*
I.3 Number of groups	Not applicable (not more than one group) *Please edit*

Section J: Methods - Sampling strategy

J.1 Are the authors trying to produce findings that are representative of a given population?	Explicitly stated (please specify) To all the Flemish secondary schools.
J.2 What is the sampling frame (if any) from which the participants are chosen?	Explicitly stated (please specify) *Flemish secondary schools*
J.3 Which method does the study use to select people, or groups of people (from the sampling frame)?	Implicit (please specify) The set of schools is to a certain extent representative of the Flemish secondary schools.
J.4 Planned sample size	Not stated/unclear (please specify) *Please edit*
J.5 How representative was the actual sample (as recruited at the start of the study) in relation to the aims of the sampling frame?	Medium (please specify) With respect to the mathematics and language achievement: The sample size can be reduced throughout the analyses due to missing values on some of the variables involved. With respect to the noncognitive outcomes: Of the 4,759 students, 150 changed schools during their first 2 years in secondary education. Our

	analyses are based on the remaining 4,609 students.
J.6 If the study involves studying samples prospectively over time, what proportion of the sample dropped out over the course of the study?	Implicit (please specify) Not clearly specified, though, it is mentioned that the noncognitive data have an inherent three-level structure and the noncognitive outcomes of all 4759 students are analysed.
J.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?	Yes (please specify) With respect to the noncognitive outcomes: Of the 4,759 students, 150 changed schools during their first 2 years in secondary education. Our analyses are based on the remaining 4,609 students.
J.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?	Yes (please specify) intelligence, initial achievement and SES

Section K: Methods - Data Collection

K.1 Which methods were used to collect the data?	Curriculum-based assessment *Please edit* Self-completion questionnaire Because of the large sample size, data on schools, classes, teachers and students had to be collected by means of questionnaires. All the noncognitive outcomes in this study were derived from one questionnaire (104 five-point items), the well-being questionnaire. The questionnaire was administered four times during the students' secondary school career: at the end of the first, second, fourth and sixth grade. Coding is based on: Author's description *Please edit*
K.2 Do the authors' describe any ways they addressed the repeatability or reliability of their data collection tools/methods?	Details The mathematics achievement test used at the end of the first grade (MATH1) covers set and relations theory, theory of numbers and geometry. The reliability is 0.76. The mathematics achievement test used at the

	end of the second grade (MATH2) covers theory of numbers and geometry. The reliability (Cronbach's a) is 0.70
K.3 Do the authors describe any ways they have addressed the validity or trustworthiness of their data collection tools/methods?	Details The level of content validity is high, because of the several items asked in the tests. The aggregation included all students in a class and not only those students belonging to the LOSO - cohort.

Section L: Methods - data analysis

L.1 Which methods were used to analyse the data?	Explicitly stated (please specify) The data were analysed by means of correlations and multilevel analysis. Three levels were identified: the student level, the school level, and an intermediate level that combines the mathematics teacher and the class group. The multi-level analysis is done by the MLwiN-software. Student-level measures were aggregated and used as descriptive indicators of the

	composition of the group of students in the classroom. A mean score is calculated for each class separately: mean initial cognitive ability (CL-COGN), mean SES (CL-SES), mean achievement motivation (CL-AM), mean immunity to stress (CL-STRESSIMM), proportion of girls in the class (CLSEX)and proportion of students who speak Dutch at home (CL-DUTCHHOME). The group means are calculated over all students for which the particular variable is available. The aggregation included all students in a class and not only those students belonging to the LOSO-cohort. Students with missing data are mostly omitted from the multilevel analyses, but in the aggregation process even students with values on other variables are included. On the other hand, we did not calculate group composition scores that are based on too small fraction of the group. If less than 50 % of the scores on a student level

	variable in a class were available, the aggregated score for that class was not calculated. This procedure is expected to reduce the overall measurement error on the independent variables
L.2 Which statistical methods, if any, were used in the analysis?	Details the tests were converted into IRT-scores Noncognitive outcomes are analysed by factor analysis.
L.3 Do the authors describe strategies used in the analysis to control for bias from confounding variables?	Yes (please specify) Only students who did not have to retake the first grade were included in the study. That is only students from the A-stream that is students that stayed 2 consecutive years were considered.
L.4 Do the authors describe any ways they have addressed the repeatability or reliability of data analysis?	Details Language: Internal consistency was high (0.90). Students with missing data were omitted from the multilevel models, but in the aggregation process even students with missing values on

	other variables were included. If less then 50% of the scores on a student-level variable in a class was available, the aggregated score for that class was not calculated. This procedure was expected to reduce the overall measurement error on these independent variables.
L.5 Do the authors describe any ways that they have addressed the validity or trustworthiness of data analysis?	Details The content validity of the achievement tests for mathematics and language was assessed by teacher ratings of the extent to which students have had the opportunity to learn the content represented in the individual test items. A test item is scored "1" by the mathematics/Dutch teacher when the item is not covered in the curriculum. A score of "2" refers to items that students should be able to solve on the basis of the content covered, although the formulation of the item differs from the usual presentation in the class. A score of 3 indicates that the item is a typical question for the

	student in the class (would have appeared in a regular examination). The aggregation included all students in a class and not only those students belonging to the LOSO - cohort. Students with missing data are mostly omitted from the multilevel analyses, but in the aggregation process even students with missing values on other variables are included. On the other hand, we did not calculate group composition scores that are based on too small a fraction of the group. If less than 50% of the scores on a student-level variable in a class were available, the aggregated score for that class was not calculated. This procedure is expected to reduce the overall measurement error on the independent variables.
L.6 If the study uses qualitative methods, how well has diversity of perspective and content been explored?	Details Not a study with qualitative methods.

L.7 If the study uses qualitative methods, how well has the detail, depth and complexity (i.e. the richness) of the data been conveyed?	Details Not a study with qualitative methods.
L.8 If the study uses qualitative methods, has analysis been conducted such that context is preserved?	Details Not a study with qualitative methods.

Section M: Quality of study - reporting

M.1 Is the context of the study adequately described?	Yes (please specify) The study context is described in Van Damme et al., 2002 Sampling is explicitly described
M.2 Are the aims of the study clearly reported?	Yes (please specify) Their study has attempted to be a fully fledged longitudinal school effectiveness study, paying particular attention to student background variables and compositional effect
M.3 Is there an adequate description of the sample used in the study and how the sample was identified and recruited?	Yes (please specify) Yes, the sample is described in Vab, Damme et al 2002 p 386

M.4 Is there an adequate description of the methods used in the study to collect data?	Yes (please specify) Questionnaires and scales administered to teachers, student assessment in reading, math and non cognitive outcomes
M.5 Is there an adequate description of the methods of data analysis?	Yes (please specify) Yes, multi level modelling was used, applying among others three level analyses,
M.6 Is the study replicable from this report?	Yes (please specify) Yes, procedures are described explicitly
M.7 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?)	Yes (please specify) There is no sign of biased reporting, also none effects and counter intuitive effects were reported

Section N: Quality of the study - Weight of evidence

N.1 Were users / relatives of users involved in the design or conduct of the study?	No, involvement is not relevant (please specify) *No, involvement in never relevant*
N.2 Was the choice of research design appropriate for addressing the research question(s) posed?	Yes, completely (please specify) The design was completely appropriate. In fact this series of

	studies is among the best in the field, from a methodological perspective
N.3 Have sufficient attempts been made to establish the repeatability or reliability of data collection methods or tools?	Yes, good (please specify) Yes, procedures are explicit, reliabilities are reported
N.4 Have sufficient attempts been made to establish the validity or trustworthiness of data collection tools and methods?	Yes, good (please specify) See previous point. Psychometric qualities of instruments are reported
N.5 Have sufficient attempts been made to establish the repeatability or reliability of data analysis?	Yes (please specify) Yes, see previous points
N.6 Have sufficient attempts been made to establish the validity or trustworthiness of data analysis?	Yes, good (please specify) *Yes, see previous answers*
N.7 To what extent is the research design and methods employed able to rule out any other sources of error/bias which would lead to alternative explanations for the findings of the study?	A lot (please specify) As stated before, in its longitudinal design, its elaborate controls and state of the art analysis this is an exemplarily study in its field

N.8 How generalisable is the study results?	Details The sample might be representative for Flemish secondary schools
N.9 In light of the above, do the reviewers differ from the authors over the findings or conclusions of the study?	Not applicable (no difference in conclusions) No difference with the conclusions of the authors
N.10 Weight of evidence A: Taking account of all quality assessment issues, can the study findings be trusted in answering the study question(s)?	High trustworthiness Again this is a very good set of studies
N.11 Weight of evidence B: Appropriateness of research design and analysis for addressing the question, or sub-questions, of this specific systematic review.	High The studies considered relevant school and class factors, only one or two factors had a positive effect. The study is of great relevance, since it is the best example of applying controls in term of background and compositional effects that I am aware of.
N.12 Weight of evidence C: Relevance of particular focus of the study (including conceptual focus, context, sample and measures) for addressing the question,	High The study is very severe in controlling for non malleable variables; as a result the yield in terms of policy amenable vari-

or sub-questions, of this specific systematic review	ables is very limited. This could be seen as a "negative" outcome, but in fact it is not. School effectiveness is an academic field of research, and falsifying expectations is as important as finding support in the expected direction. The studies teach us that policy malleable variables have much smaller impact than "given" background variables
N.13 Weight of evidence D: Overall weight of evidence	High Again: very high level research

7 Appendix 2

In this appendix all school factors applied in the data extraction are presented. To each factor comments are added about relevant aspects, such as how many studies address this factor with high or medium weight of evidence, which other factors are treated in combination with this factor, and – if subcategories are enumerated – whether it has been possible to work with these subcategories in a reliable manner.

7.1 School size

The school factor 'School size' is concerned with the number of pupils in the school.

No subcategories are made regarding this factor.

In total, 19 of the 71 studies rated with high or medium weight of evidence examined the school factor "School size".

15 studies apply a quantitative approach. 3 studies apply qualitative methods, and 2 studies apply a mixed method approach, i.e. both qualitative and quantitative methods.

Of importance in relation with this school factor is the systematic review by Garrett: Secondary school size - a systematic review. The systematic review examines the relationship between secondary school size and outcomes (both student outcomes and others).

The studies dealing with the school factor 'school size' are the following:

High weight of evidence (N=3): Opdenakker, van der Werf, Greetje.

Medium weight of evidence (N=15): Bondi, Silins, Fullarton, Lamb S, Dronkers, Foley, Heck, Coates, Smyth, Grisay, Mandeville, Franklin, Postlethwaite, Martin, Perez.

The studies that deal with the school factor 'school size' deal with the other school factors in the following way:

Attribute	Number
Class size	10
Curriculum/scheduling	13
Management and Leadership	12
Other factors/phenomena (please specify)	11
Parental Relationship	5
Physical environment	3
Pupil composition of the school	12
School culture and climate	13
Support teams	3
Teacher	17

Table 7.1: Primary studies that examine School Size together with other named school factors

7.2 Class size

Class size concerns the number of pupils in the class. Besides the number of pupils, this factor also comprises dimensions such as student-teacher ratios, teacher aid and teaching assistance.

No subcategories are made regarding this school factor.

Of the 71 studies rated with high or medium weight of evidence, 15 studies deal with the school factor 'Class size'. All the studies make use of quantitative methods in the data analysis; three of the studies apply qualitative methods. Three studies apply a mixed method approach, i.e. both quantitative and qualitative methods.

The studies dealing with the school factor 'class size' are the following:

High weight of evidence (n=3): Rogers, Young, Woessmann.

Medium weight of evidence (n=12): Thomas, Lamb, Perez, Coates, Smyth, Grisay, Franklin, Postlethwaite, Bain, Martin, Ringsmose, Grøgaard.

The studies that deal with the school factor 'class size' deal with the other school factors in the following way:

School factor	Number of studies
Curriculum/scheduling	11
Management and Leadership	9
Other factors/phenomena (please specify)	7
Parental Relationship	7
Physical environment	5
Pupil composition of the school	9
School culture and climate	10
School size	9
Support teams	6
Teacher	15

Table 7.2: Primary studies that examine Class Size together with other named school factors

7.3 Management and leadership

The scope of the school factor 'management and leadership' is defined as follows:

The concepts of management and leadership are often used interchangeably in the study of schools. Leadership could however be seen

as the broader concept relative to the two narrower concepts: management and educational leadership.

Management concerns the local school level as the decision-making authority. It is related to decisions concerning curricula, instructional technologies, and other school initiatives. Three areas of decision-making can be school based: budget (e.g. decisions regarding personnel, equipment, materials, and staff development), personnel (e.g. recruitment), and curriculum (e.g. decisions regarding the curriculum and instructional strategies at the school level within a framework of district or state goals).

Educational leadership is traditionally associated with people in positions such as principals and superintendents. Accordingly, principals and superintendents are the parties most responsible for crafting the essential educational agreements upon which schools either succeed or fail.

All studies with a bearing on this factor/phenomenon have been classified on the two following dimensions: [1]

External orientation of leadership, internal orientation of leadership

Content of leadership: Human resources, rational goal leadership (Quinn & Rohrbauch, 1983), educational leadership, administrational leadership, etc.

Several studies employed more than one measure to assess leadership/management. In such cases each measure has been classified according to this system.

[1] An attempt was also made to classify the studies on the following leadership/management dimensions: Transactional, transformational, distributive, not applicable. As this led to 3 out of 4 of the studies being classified as 'not applicable', these dimensions have been left out of the analysis.

In the systematic review we have 39 studies with medium or high weight of evidence with a bearing on this school factor.

In addition, one systematic review examines the effects of transformational leadership on student outcome (Leithwood & Jantzi), and one examines the impact of head teachers and principals on student outcome (Bell).

Among the 39 studies are 26 which apply a design with quantitative elements and 6 which apply a design with qualitative elements. 7 studies have a 'mixed methods approach'.

The studies dealing with the school factor 'Leadership and management' are the following:

High weight of evidence (N=6): Ross(2006a), Tarter, Reezigt, Teddlie, Ross(2006b), Woessmann

Medium weight of evidence(N=33): Hofman(1996), Webster, Silins, Waxman, Dronkers, Foley, Heck, Perez, Lindsay, Kitchen, Pressley, Hofman(2002), Picucci, Towns, Mosenthal, Sweetland, Texas, Thomas, Zigarelli, Florida, Grisay, Hill, Kennedy, Postlethwaite, Hoy, Witte, Pressley, Traufler, Senkbeil, Sammons, Kyriakides, Stringfield, Ringsmose, Grøgaard.

The studies can be distributed on the abovementioned leadership and management dimensions as follows:

Leadership/management dimension	Number of studies
External orientation	8
Internal orientation	33
Human resources	19
Rational goal leadership	3
Educational Leadership	27
Administrational leadership	14
Other	6

Table 7.3: Distribution of Management and Leadership into Subcategories

The 46 studies address the other 10 school factors as follows:

School factors/phenomena	Number of studies
School culture and climate	32
Teacher	33
Curriculum/scheduling	29
Parental Relationship	25
Pupil composition of the school	15
School size	11
Physical environment	9
Class size	9
Support teams	9
Other	18

Table 7.4: Primary studies that examine Management and Leadership together with other named school factors

7.4 Curriculum/scheduling:

The scope of the school factor 'curriculum/scheduling' is defined as follows:

Curriculum is often defined as covering only those topics actually taught by teachers. However, the definition of curriculum can range from virtually everything that takes place in a classroom to the topics that are defined as instructional requirements in the legal regulation of an educational system. Curriculum can further be subdivided into three components: the intended, the implemented, and the attained. Typical examples could be opportunity to learn, homework, coordination and alignment of the curriculum, and learning goals.

All studies with a bearing on this factor/phenomenon have been classified on the following subcategories:

Opportunity to learn: This subcategory consists of the curriculum actually offered to the students. (Homework is placed in 'Opportunity to learn').

Alignment: 'Alignment' is about coordination, i.e., bringing purpose and means together. (Differentiation on an organisational level such as 'single gender classroom' is placed in this subcategory).

Learning goals

Other (School resources such as books are categorised in the 'Other' category).

In the systematic review we have 41 studies with medium or high weight of evidence with a bearing on this school factor.

In addition, one systematic review deals with curriculum in the form of social information processing interventions (Wilson). This review focuses on programs used in school settings that address one or more aspects of students' social information processing difficulties. It examines the effects of universal school-based social information processing interventions (training in social information processing steps, teaching generic thinking skills, use of structured tasks and activities) on the aggressive and disruptive behaviour of school-age children.

Among the 42 studies are 30 which apply a design with quantitative elements and 7 which apply a design with qualitative elements. 5 studies have a 'mixed methods approach'.

The studies dealing with the school factor 'Curriculum/scheduling' are the following:

High weight of evidence (N=9): Ross (2006a), Campbell, Meelissen, Opdenakker (2007), Rogers, Rumberger, Reezigt, Van der Werf (1997), and Woessmann.

Medium weight of evidence (N=33): Lamb, Waxman, Foley, Dronkers, Dumay, Perez, Lindsay(2006), Bottoms, Kitchen, Pressley (2007), Coates, Hofman (2002), Towns, Mosenthal, Sweetland, Texas Education Agency, Smyth, Thomas(1997), Grisay, Hill, Mandeville, Postlethwaite, Pressley (2004), Traufler, Meijnen, Martin, Kyriakides, Stringfield(1993), Grøgaard, Taylor, Willis, Ringsmose, Young (2001).

The studies can be distributed into the above-mentioned subcategories as follows:

Curriculum/scheduling subcategories	Number of studies
Opportunity to learn	**33**
Alignment	**11**
Learning goals	**5**
Other	**5**

Table 7.5: Distribution of Curriculum/Scheduling into Subcategories

The 42 studies address the other 10 school factors as follows:School factors/phenomena	Number of studies
School size	12
Class size	10
Management and leadership	29
School culture and climate	35
Teacher	38
Support teams	7
Physical environment	7
Pupil composition of the school	114
Parental Relationship	14
Other factors/phenomena	11

Table 7.6: Primary studies that examine Curriculum/Scheduling together with other named school factors

7.5 School culture and climate

"School culture and climate" is understood in terms of the feel, atmosphere, tone, ideology, or milieu of a school. The concepts of school climate and school culture are often used interchangeably in the study of schools. Some authors, however, make a distinction between the two.

While much of the school climate literature focuses on the structural dimensions of schools, culture looks beyond structural elements, both the formal and informal specifics, to the meanings those specifics hold for the participants and how they make use of them.

When school climate and school culture are seen as synonyms, the indicators of school culture/climate can range from perceptions and normative views to behavioural characteristics and factual circumstances (e.g. shared visions, goals and values, monitoring progress,

achievement orientation, internal relationships, evaluative potential, feedback reinforcement and behavioural rules).

All studies with a bearing on this factor/phenomenon have been classified into the following subcategories:

Disciplinary climate

Achievement/progress orientation (This subcategory includes an evaluative culture; it also includes the students' attitude towards the school and school work as well as the students' self-concept regarding the school work)

Interrelational climate

Social norms and values (Pupil involvement is assigned to this subcategory)

Other.

In the systematic review we have 57 studies with medium or high weight of evidence with a bearing on this school factor.

In addition, one systematic review (Dyson) examines the relationship between an inclusive culture and student participation. The inclusive culture is here understood as consensus around values of respect for difference and a commitment to offering all students access to learning opportunities;there is a high level of staff collaboration and joint problem solving, and similar values and commitments may extend into the student body and into parent and other community stakeholders in the school.

Among the 57 studies are 53 which apply a design with quantitative elements and 18 which apply a design with qualitative elements. 11 studies have a 'mixed methods approach'.

The studies dealing with the school factor "School culture and climate" are the following:

High weight of evidence (N=14): Ross (2006a), Meelissen, Opdenakker(2007), Rogers, Rumberger, Tarter, Van Damme, Yu, Reezigt, Van der Werf(1997), Van der Werf, Teddlie, Young.

Medium weight of evidence (N=43): Hofman(1996), Webster, Silins, Young, Lamb, Waxman, Dronkers, Dumay, Foley, Lindsay, Bottoms, Choi, Kitchen, Pressley, Griffith, Hofman(2002), Picucci, Towns, Mosenthal, Sweetland, Texas Education Agency, Opdenakker, Smyth, Thomas, Zigarelli, Grisay, Hill, Kennedy, Hoy, Witte, Pressley, Griffith, Willis, Traufler, Senkbeil, Martin, Sammons, Kyriakides, Papanastasiou, Stringfield, Ringsmose, Grøgaard, Taylor(2000).

The studies can be distributed into the above-mentioned subcategories as follows:

School culture and climate subcategories	Number of studies
Disciplinary climate	**27**
Achievement/progress orientation	**37**
Interrelational climate	**33**
Social norms and values	**22**
Other	**2**

Table 7.7: Distribution of School Culture and Climate into Subcategories

The 57 studies address the other 10 school factors as follows:

School factors/phenomena	Number of studies
School size	13
Class size	10
Management and leadership	42
Curriculum/scheduling	36
Teacher	51
Support teams	10
Physical environment	7
Pupil composition of the school	21
Parental Relationship	27
Other factors/phenomena	18

Table 7.8: Primary studies that examine School Culture and Climate together with other named school factors

7.6 Teacher

Teacher is understood in terms of the teacher as an individual and/or the teacher as part of an organisation.

All studies with a bearing on the 'teacher as an individual teacher' have been classified into the following subcategories:

Teacher behaviour: covers the ways teachers ensure that pupils behave in an appropriate manner both towards each other/the teacher, and in relation to the learning that is to take place in the school. It is about getting the teaching right (e.g. by differentiation/ using a variety of teaching strategies). Examples of teacher behaviour are:

Classroom management: teacher's organisation and structuring of teaching

Behaviour management: Correction of student misbehaviour e.g. rewards truly praiseworthy behaviour.

Classroom climate: Contribution from the teacher to the classroom climate e.g. high expectations, teacher enthusiasm, avoids criticism.

Teacher beliefs: represents teacher's theories about how pupils function, i.e. their beliefs about what constitutes 'good teaching'.

Subject knowledge: is about the teacher's content knowledge of his/her subject.

Teacher self-efficacy beliefs: this is covered by two concepts:

Teachers' self-concept (a person's perception of him-/herself, formed through interaction with the environment)

Teachers' self-efficacy (a teacher's judgment of his/her capabilities to bring about desired outcomes of the student engagement and learning)

The scope of 'teacher as an organisational actor' is determined as follows:

The aspect could contain teacher groups/teams, the teachers' job satisfaction, teachers' gender, teacher corps stability, teachers' formal competence (certified/uncertified teacher/teaching assistant).

In the systematic review we have 62 studies with medium or high weight of evidence with a bearing on this.

Moreover, one systematic review (Nordenbo et al) examines the relationship between teacher competencies (i.e. what teachers know, value and do in the classroom context) and how pupils are affected by this, i.e. student learning.

Among the 62 studies are 56 that apply a design with quantitative elements and 17 applying a design with qualitative elements. 9 of the studies apply a 'mixed methods approach'.

The studies dealing with the school factor 'Teacher' are the following:

High weight of evidence (N=16): Ross(2006a), Meelissen, Opdenakker (2007), Rogers, Rumberger, Tarter, Van Damme, Yu, Taylor,

Reezigt, van der Werf, van der Werf, Teddlie, Young, Ross (2006b), Woessmann.

Medium weight of evidence (N= 47): Thomas, Webster, Silins, Fullarton, Lamb, Waxman, Dronkers, Foley, Heck, Perez, Lindsay, Bottoms, Choi, Kitchen, Pressley, Coates, Griffith, Hofman, Towns, Mosenthal, Sweetland, Texas, Opdenakker, Smyth, Thomas, Zigarelli, Grisay, Hill, Mandeville, Kennedy, Franklin, Postlethwaite, Bain, Hoy, Pressley, Griffith, Willis, Papanastasiou,Traufler, Senkbeil, Meijnen, Martin, Sammons, Kyriakides, Picucci, Ringsmose, Grøgaard.

The studies can be distributed on the above-mentioned teacher dimensions as follows:

Teacher	Number of studies
Teacher behaviour	**49**
Teacher beliefs	**5**
Teacher self-efficacy beliefs	**6**
Teacher subject knowledge	**2**
Teacher as an organisational actor	**40**

Table 7.9: Distribution of the factor Teacher into subcategories

The 63 studies address the other 10 school factors as follows:

Attribute	Number
Class size	17
Curriculum/scheduling	38
Management and Leadership	42
Other factors/phenomena (please specify)	21
Parental Relationship	27
Physical environment	9
Pupil composition of the school	23
School culture and climate	54
School size	18
Support teams	11

Table 7.10: Primary studies that examine Teacher together with other named school factors

7.7 Support teams

The scope of 'support teams' is defined as follows:

'Support teams' is concerned with non-instructional services or extra-curricular activities with the goal of addressing students´ needs, e.g., school dentist, nurse, advisors, and leisure-time activities.

No subcategories have been defined for this factor.

In the systematic review we have 12 studies with medium weight of evidence with a bearing on this school factor.

In addition, one systematic review (Zif) examines the effects of after-school programs (programs that combine recreation/youth development/academic support services) on youth context, participation in activities, and behavioural, social and emotional and academic outcomes.

Among the 12 studies are 11 which apply a design with quantitative elements and 6 applying a design with qualitative elements. 5 studies have a 'mixed methods approach'.

The studies dealing with the school factor 'Support teams' are the following:

Medium weight of evidence (N=12): Thomas(1995), Foley, Perez, Bottoms, Choi, Kitchen, Picucci, Florida State Dept. of Education, Grisay, Kyriakides, Ringsmose, Grøgaard.

No studies with overall high weight of evidence deal with the school factor 'Support teams'.

The 12 studies address the other 10 school factors as follows:

School factors/phenomena	Number of studies
School size	3
Class size	6
Management and leadership	9
Curriculum/scheduling	8
School culture and climate	10
Teacher	11
Physical environment	3
Pupil composition of the school	7
Parental Relationship	3
Other factors/phenomena	3

Table 7.11: Primary studies that examine Support Teams together with other named school factors

7.8 Physical environment

Studies grouped within the school factor 'Physical environment' all deal with the physical characteristics of the school. Examples of such

characteristics are facilities such as furnishing, materials and supplies, equipment and information technology, characteristics of the school building and various aspects of the school layout such as athletic fields and playgrounds.

No subcategories have been defined for this school factor.

Ten of the studies rated with medium or high weight of evidence addressed the school factor 'Physical environment'. Of these ten studies, seven make use of quantitative methods in the data analysis, three make use of qualitative methods and three studies adopt a 'mixed methods' approach. The studies dealing with the school factor 'Physical environment' are the following:

Medium weight of evidence (N=10): Dronkers, Pressley (2004), Grisay, Postlethwaite, Pressley(2007), Ringsmose, Grøgaard, Willis (1996), Towns (2001), Texas (2000).

The 10 studies that deal with the school factor 'Physical environment' address the other school factors as follows:

School factor	Number of studies
School size	3
Class size	5
Management and leader-ship	10
Curriculum/scheduling	9
School culture and climate	9
Teacher	10
Support teams	3
Pupil composition of the school	4
Parental Relationship	6
Other factors	4

Table 7.12: Primary studies that examine Physical Environment together with other named school factors

7.9 Pupil composition of the school

Studies that are grouped in the category 'Pupil composition of the school' all deal with the effects of percentages of different groups of pupils in the school (e.g. social economic status, special educational needs and ethnicity). This factor is more accurately defined as *"the aggregate characteristics of a student group on a student's learning over and above the effects on learning associated with that student's individual characteristic"* (Wilkinson, 2002 in Dumay & Dupriez, 2007). Hence, this factor is not to be confused with the inclusion criteria which every study has met in order to be included in this review: "control is present for differences in pupils' socio-economic background" or "control is present for differences in pupils' scholastic aptitude". *(NB! It has sometimes been difficult to determine whether the studies addressing this factor are dealing with pupil composition of the school as a control variable or as a malleable school factor)*

No subcategories have been defined for this factor.

24 of the studies rated with medium or high weight of evidence addressed the school factor 'Pupil composition of the school'. All of these studies make use of quantitative methods in the data analysis. However, six studies also make use of qualitative methods.

The studies dealing with the school factor 'Pupil composition of the school' are the following:

High weight of evidence (N=6): Meelissen, Opdenakker, Tarter, Young, van Damme (1997), Rogers.

Medium weight of evidence (N=18): Thomas (1995), Bondi, Fullarton, Young (2001), Kennedy, Dronkers, Dumay, Heck, Perez, Choi, Smyth, Grisay, Mandeville, Witte, Senkbeil, Martin, Ringsmose, Grøgaard

The 24 studies that deal with the school factor 'Pupil composition of the school' address the other school factors as follows:

School factor	Number of studies
School size	11
Class size	9
Management and leadership	12
Curriculum/scheduling	12
School culture and climate	19
Teacher	22
Support teams	7
Physical environment	4
Parental Relationship	7
Other factors	10

Table 7.13: Primary studies that examine Pupil Composition of the School together with other named school factors

7.10 Parental Relationship

The school factor 'Parental Relationship' is used to group studies that deal with the emphasis on parental involvement in school policy and contact with parents. Illustrative examples of the school's role in encouraging parental involvement include practices such as conducting workshops for families, and communicating to parents about their children's education.

No subcategories have been defined for this school factor.

Of the 71 studies rated with high or medium weight of evidence, 26 studies focus on 'Parental Relationship'. 26 studies make use of quantitative methods in the data analysis, and 12 studies make use of qualitative methods. Nine of the studies apply both qualitative and qualitative methods.

The studies dealing with the school factor 'Parental Relationship' are the following:

High weight of evidence (N=7): Ross(2006a), Rogers, Taylor, Van der Werf(1997), Van der Werf(1996), Teddlie¸ Ross(2006b)

Medium weight of evidence(N=19): Hofman(1996), Young(2001), Foley, Lindsay, Sweetland, Griffith, Hofman(2002), Towns, Texas, Zigarelli, Hill, Kennedy, Postlethwaite, Witte, Pressley, Traufler, Senkbeil, Sammons, Ringsmose.

The 26 studies that deal with the school factor 'Parental Relationship' address the other school factors as follows:

School factor	Number of studies
School size	4
Class size	5
Management and leadership	23
Curriculum/scheduling	13
School culture and climate	25
Teacher	25
Support teams	3
Physical environment	4
Pupil composition of the school	6
Other factors	11

Table 7.14: Primary studies that examine Parental Relationship together with other named school factors

7.11 Other

The school factor 'Other' is used if the included studies address factors that are incompatible with the other ten school factors listed in the EPPI-reviewer.

23 of the studies rated with medium or high weight of evidence address school factors that cannot be placed in any one of the ten specified school factors listed in the EPPI-reviewer. All 23 studies make use of quantitative methods in the data analysis. However, two of the studies also make use of qualitative methods.

The studies dealing with school factors that are not directly compatible with the remaining ten factors listed in the EPPI-reviewer are the following:

High weight of evidence (N=5): Ross(2006a), Opdenakker, Rumberger, Van der Werf, Woessmann,

Medium weight of evidence (N=18): Hofman(1997), Bondi, Silins, Young, Lamb, Dronkers, Foley, Heck, Choi, Hofman(2002), Sweetland, Kennedy, Postlethwaite, Bain, Hoy, Senkbeil, Martin, Grøgaard.

A more detailed examination of the school factor 'Other' shows that more than half of the studies (N=12) placed in this category deal with the factor 'location', i.e. whether the school is located in an inner-city, suburban, urban or rural area.

Dimensions within the overall factor: Other	Number of studies
Location	12
Denominational status	5
Sector (public/private)	4
Overall school measure	7
Other	5

Table 7.15: Dimensions within the overall factor "Other"

Next, five of the studies focus on the effects of the school's "denominational status", i.e. whether the school is catholic or non-catholic,

and four of the studies deal with the factor 'sector', i.e. whether the school is public or private. **(NB! Whether the factors 'location',' denominational status' and 'sector' are to be considered as malleable school factors is obviously debatable. This matter should be further discussed at the review meeting).*

Another group of studies (N=7) placed in the 'Other' category are characterised by the fact that they all make use of an 'overall school measure' which typically consists of a number of very different school characteristics. Examples of such overall measurements or indexes are: 'The quality of instructional systems' (Foley); 'School quality of learning opportunities' (Heck); and 'Institutional integrity' (Hoy).

The five remaining studies in the 'other' category in the table above address factors that are very diverse in scope. The factors addressed in these studies are: parents as context, selection of pupils, interaction between home and school, child/youth school, and average performance level.

The 23 studies that deal with the overall school factor 'Other' address the other ten school factors in the following way:

School factor	Number of studies
School size	11
Class size	7
Management and leader-ship	18
Curriculum/scheduling	11
School culture and climate	18
Teacher	21
Support teams	3
Physical environment	4
Pupil composition of the school	10
Parental Relationship	11

Table 7.16: Primary studies that examine "Other" together with other named school factors

8 Appendix 3

In this Appendix we briefly discuss an attempt to synthesise some of the quantitative findings in the studies included in this review. Initially this was intended as a meta-analysis (cf. Borenstein et al., 2009), but for several reasons (see section 8.1 below) this was not feasible. Instead we show the distribution of the significance of the different school factors. This analysis tells us the magnitude of significance of different types of school factors. This enables us to assess which of the different school factors have the greatest statistical significance. It does not say much about the size of the impact of the different factors, however, which is usually the purpose of meta-analysis. However, we believe that looking at the distribution of the significance of different factors might be useful both in terms of the direction of future studies - as our results might point to promising research directions - and in terms of assessing the significance of the different school factors discussed in this systematic review.

In the analysis below we only include studies that include pupil achievement in maths, science and reading as outcome. This includes the vast majority of studies in this systematic review; cf. Chap. 7, Appendix 2. However, these outcomes are not seen as relevant only in relation to the subjects mentioned, but in most cases – when outcomes are greater than expected – as an indication of the 'good school' in general.

8.1 Why not a meta-analysis?

There are three main reasons why it is not possible to carry out a meta-analysis in the context of the present systematic review.

Firstly, the factors in the studies included in this review are multi-facetted in design, cf. p. 28. That is, studies are selected for this re-

view by the fact that they look at several school factors. This means that there is no single school factor covered by all studies and hence that no single factor can be extracted from the studies in the review. This implies that studies on a single feature of the school, for example 'teacher effectiveness' or 'the competence of school leaders', were not included.

Secondly, many of the studies provide only limited information about the variables in the statistical analysis. To compare effect size across different studies, information is needed not only about actual effect size but also about sample size and standard deviations for both dependent and independent variables. This information does not appear systematically throughout the studies in this review, so meta-analysis is difficult or even impossible. [1]

Thirdly, because all studies are based on non-experimental or quasi-experimental data, cf. Table 3.10, the relationships between outcomes and factors are correlations and do not necessarily have a causal interpretation. Furthermore, there is no common set of controls in the studies in the review and hence the effect and significance of the different factors are the result of a very heterogeneous set of statistical controls.

In sum, there is only a very limited scope and potential, if any, for carrying out a meta-analysis based on the studies in this review. We therefore turn to the more modest aim of discussing the significance of different school factors.

[1] Clearinghouse has made an attempt to extract the necessary data from the relevant studies in a Working Paper (Clearinghouse, 2009). The present assessment is based on the data collected for this report.

8.2 The factors included in the analysis

This analysis of the statistical significance of school factors is divided into five broader categories: Teacher, School Culture and Climate, Curriculum/Scheduling, Management and Leadership and 'Other Factors', and six narrower categories: School Size, Class Size, Support Teams, Physical Environment, Pupil Composition of the School and Parental Relationship. Prototypical elements of the different factors are indicated in Frame 8.1, p. 220, and in Chap. 7: Appendix 2.

D.2.1 School size

The study deals with the number of pupils in the school

D.2.2 Class size

The study deals with the number of pupils in the class or with pupil-teacher ratios

D.2.3 Management and Leadership

Keywords:

1. External orientation, internal orientation (at least one of these must always be applied)
2. Human resources, rational goal leadership (from Grim and Rohrbauch), educational leadership, administrational leadership, other (at least one of these must always be applied)
3. Transactional/transformational/distributive/not applicable (at least one of these must always be applied)

Scope:
The study deals with management and leadership. The concepts of

management and leadership are often used interchangeably in the study of schools. Leadership could be seen as the broader concept to the two narrower concepts: management and educational leadership.

Management concerns the local school level as the decision-making authority. It is related to decisions concerning curricula, instructional technologies, and other school initiatives. Three areas of decision-making can be school-based: budget (e.g. decisions regarding personnel, equipment, materials, and staff development), personnel (e.g. recruitment), and curriculum (e.g. decisions regarding the curriculum and instructional strategies at the school level within a framework of district or state goals).

Educational leadership is traditionally associated with people in positions such as principals and superintendents. Accordingly, principals and superintendents are the parties most responsible for crafting the essential educational agreements upon which schools either succeed or fail.

Transformational leaders seek to motivate, influence, empower and develop the skills of others (Adamson, 1996). Leadership is a function of capacity and motivation, meaning that people are more motivated by affective factors than cognitive factors.

'Distributed leadership is characterised as a form of collective leadership in which teachers develop skills and expertise through working collaboratively'. The ideology within this paradigm shifts the 'doing' and 'thinking' from one to many. It is about the division of labour and creating a workplace that requires and facilitates collaboration, teamwork and cooperation.

Transactional leadership qualities include behaviours that emphasise exchanges or bargains between manager and follower, and focus on how current needs of subordinates can be met.

D.2.4 Curriculum/scheduling

Keywords:
Opportunity to learn, alignment, learning goals, other

Scope:
The study deals with curriculum/scheduling in this scope: Curriculum is often defined as only those topics actually taught by teachers. However, the definition of curriculum can range from virtually everything that takes place in a classroom to the topics that are defined as instructional requirements in the legal regulation of an educational system. Curriculum can further be subdivided into three components: the intended, the implemented, and the attained. Typical examples could be opportunity to learn, homework, coordination and alignment of the curriculum, and learning goals.

D.2.5 School culture and climate

Keywords:
Disciplinary climate, achievement/progress orientation, interrelation climate, social norms and values, other.

Scope:
The study deals with school culture/climate in terms of the feel, atmosphere, tone, ideology, or milieu of a school. The concepts of school climate and school culture are often used interchangeably in the study of schools. Some authors, however, make a distinction between the two. While much of the school climate literature focuses on the structural dimensions of schools, culture looks beyond structural elements, both the formal and informal specifics, to the meanings those specifics hold for the participants and how they make use of them. When school climate and school culture are seen as syno-

nyms, indicators on the school culture/climate range from perceptions and normative views to behavioural characteristics and factual circumstances (e.g., shared visions, goals and values, monitoring progress, achievement orientation, internal relationships, evaluative potential, feedback reinforcement, and behavioural rules)

D.2.6 Teacher

Keywords:
teacher behaviour, teacher beliefs, subject knowledge, teacher self-efficacy beliefs (the individual teacher): and teacher as an organisational actor.

Scope:
This study deals with teacher in terms of teachers as an individual teacher and/or the teacher as part of an organisation.

A. Teacher as individual covers the following:

TEACHER BEHAVIOUR:

This aspect covers the way teachers ensure that pupils behave in an appropriate manner both towards each other and the teacher and in relation to the learning that is to take place at school. It is about getting the teaching right (e.g. by differentiation/ using a variety of teaching strategies).

Teacher behaviour covers:

Classroom management: teacher's organisation and structuring of teaching.
Behaviour management: Correction of pupil misbehaviour e.g. rewards truly praiseworthy behaviour.

Classroom climate: Contribution from the teacher to the classroom

climate e.g. high expectations, teacher enthusiasm, avoids criticism.

TEACHER BELIEFS

Teacher beliefs represent teacher's theories about how pupils function, i.e. their beliefs about what constitutes 'good teaching'.

SUBJECT KNOWLEDGE

The teacher's content knowledge of his/her subject.

TEACHER SELF-EFFICACY BELIEFS

This is covered by two concepts: Teachers' self-concept (a person's perception of him/herself, formed through interaction with the environment) and teachers' self efficacy (a teacher's judgment of his or hers capabilities to bring about desired outcomes of the pupil engagement and learning).

B. TEACHER AS AN ORGANISATIONAL ACTOR

This aspect could contain teacher groups/teams, the teachers' job satisfaction, teacher's gender, teacher corps' stability, teacher's formal competence (certified/uncertified teacher/teaching assistant)

D.2.7 Support teams

The study deals with non-instructional services or extra-curricular activities with the goal of addressing pupils´ needs (e.g., school dentist, nurse, advisors, leisure time activities)

D.2.8 Physical environment

The study deals with physical characteristics of the school (e.g. facilities such as furnishing, materials and supplies, equipment and information technology, characteristics of the school building, and

various aspects of the building grounds such as athletic fields and playgrounds)

D.2.9 Pupil composition of the school

The study deals with the effects of percentages of different groups of pupils in the school (e.g. Social Economic Status, Special Educational Needs and ethnicity).
This factor is the aggregate characteristics of a pupil group on a pupil's learning over and above the effects on learning associated with that pupil's individual characteristic".

This factor should not be confused with the inclusion criteria which every included study has lived up to: "Control is present for differences in pupils´ socioeconomic background" or "control is present for differences in pupils´ scholastic aptitude".

D.2.10 Parental Relationship

The study deals with parental involvement, emphasis on parental involvement in school policy and contact with parents. The schools´ role in encouraging parental involvement can include specific practices such as holding workshops for families and communicating to parents about their children's education

D.2.11 Other factors/phenomena (please specify)

Frame 8.1: Final definition of school factors applied in the data extraction

In the following section we show how these factors differ in statistical significance.

8.3 The statistical significance of different school factors

The statistical significance of the five broad categories of school factors mentioned in the previous section has been calculated on the basis of the data extracted in the Work Paper referred to in footnote 1, p. 214. Ideally, rather than just assessing the significance of school factors, one would also want to look at effect size, that is, the effect of a malleable school factor on a student outcome. In order to compare (weight) the effect sizes for the same school factor on comparable student outcome across studies (meta analysis) one needs a number of statistics from each study: regression coefficients, sample statistics and sample sizes. However, in most of the reported studies within this review, not all of this information was available for all studies. Hence it is not possible to compare weighted effect sizes across studies. Instead we report the somewhat less informative significance level of different school factors across studies in the review. Significance in this study is reported in terms of z-scores. Z-scores indicate the effect sizes in relation to their standard errors. The ratio of effect size with its standard error follows a standard normal distribution, assuming that the true effect size is zero. The larger the ratio, the less likely is this assumption. Hence large z-scores indicate significant effect sizes. There are two reasons why z-scores may be large. The first is that the effect, –i.e. the nominator in the ratio of the effect size to its standard error – the z-score - may be large. However, the other reason is that the standard error of the effect size – the denominator of the z-score -may be small. This again depends on the residual error – what is left unexplained by the statistical analysis and the sample size of the model. Residual error depends both on the effect size of the school factor but also on the explanatory power of other independent variables in the statistical analysis. Hence the z-score for a particular school factor may be large due to a large sample size and a good fit to the student outcome from other independent

variables in the model. Therefore, z-scores are a less informative measure of the relationship between a school factor and a student outcome variable. However, they do provide some evidence of the statistical stringency of the relationship between student outcome and school factors. School factors associated with large z-scores are from studies of higher statistical strength than school factors associated with lower z-sores.

We show below how the z-scores of the different prototypical factors are distributed within each group of these school factors. The z-scores by school factor group are illustrated in Figure 8.1, p. 222.

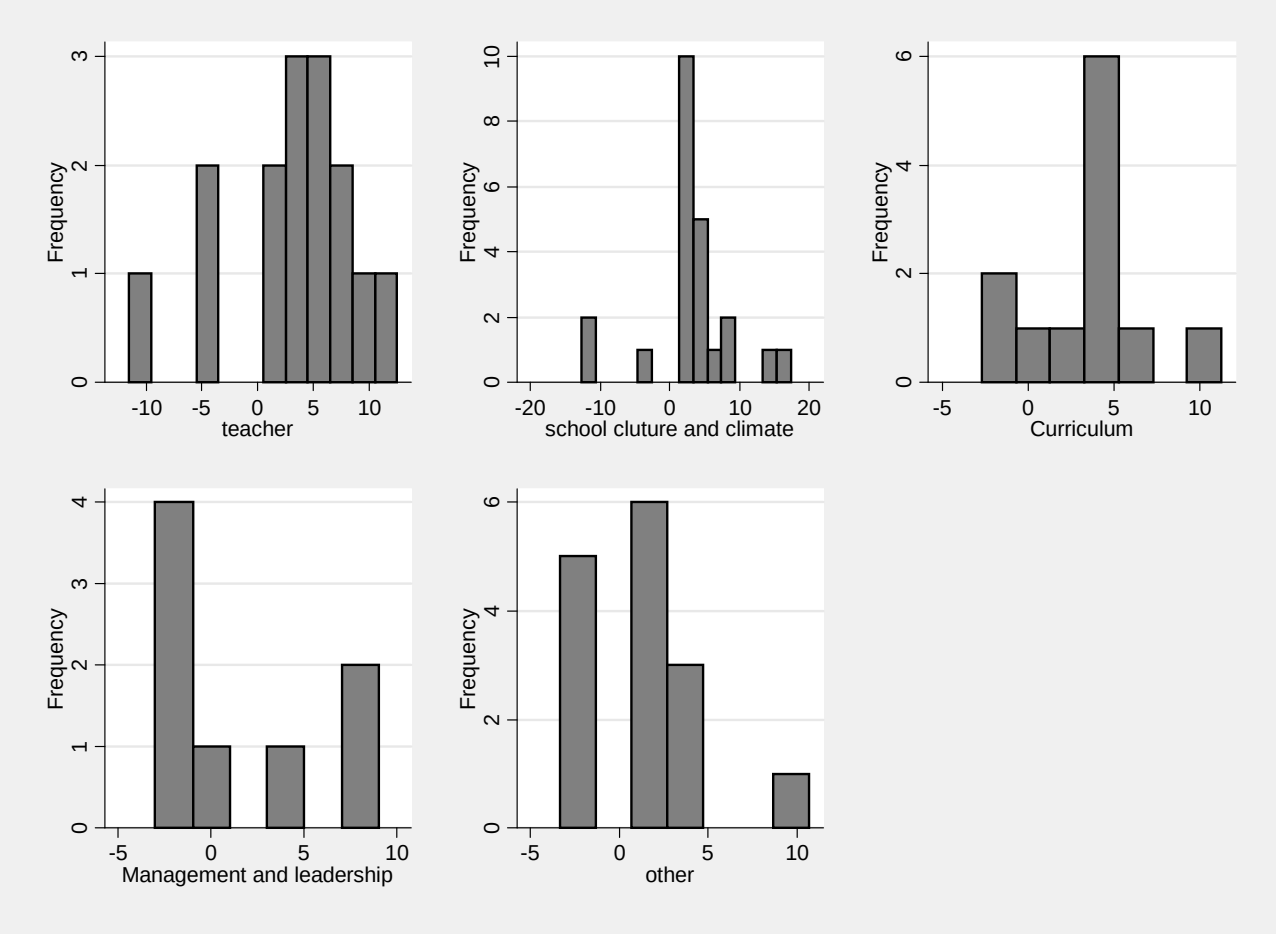

Figure 8.1: Z-scores by five broad school factor groups

From the figure it is seen that the most prototypical factors are positive in all groups. The negative factors that appear in some studies arise because some variables, like Class Size, need to be negative in order to have the expected effect, i.e. the smaller the class/school the better the pupil outcome. From the figure we find that the range of the statistical significance is more or less the same across all school factors, except for school culture and climate, where the z-scores are more dispersed. This could either be due to large effect sizes for this type of school factor or stronger statistical evidence (large samples or better statistical fit due to other variables). The fact that 'School Culture and Climate' has large z-scores could therefore be an indication that the 'School Culture and Climate' factor on average has larger effect sizes than other factors. We also find from the figures that the 'Management and Leadership' factor is less significant than other factors - which again might be due to smaller effect sizes (or larger samples etc). Therefore if on the basis of the limited amount of information available in the studies in this review one wishes to point to the most important factor, this is School Culture and Climate. Finally one should bear in mind that most studies are based on non-experimental data and thus only based on statistical control. Causality is therefore not firmly established.

9 Appendix 4: Power calculation

The following Table 9.1 is applied in connection with the estimation of the completed synthesis of the quantitative studies in Section 4.3.1.[1] For every school factor/subcategory in this section it is recorded how many studies are significant, insignificant, or intractable in this respect. In the tables it is also indicated how many of the studies have high or medium weight of evidence, respectively. However, in the synthesis this last distinction is not taken into account.

The following table (ss) indicates the hypothetical value (H_0). The level of significance 0.05 is chosen for testing the H_0-value on the basis of a common practice for statistical tests of significance. This has moreover been the general level of significance that has been found in the analysed studies. The row 'alternative' specifies the alternative value (H_1) to H_0. The row (n) gives the number of studies addressing the school factor in question. In row four (m) indicates the minimal number of studies sufficient to reject hypothesis H_0 (that the frequency of significant studies among n is larger than 5 %, which indicates the type 1 error). Row five reports the statistical power with the specified values, which test of H_0 against H_1 gives rise to.

The table shows only the calculated scenarios where the power is equal or greater than 80 %.

The calculation is done as calculations of power in the light of a binomial distribution with the parameter n=10, 11, 12, 13, ..., 100, where a hypothesis H_0: ss=0.05 is tested against an alternative H_1: ss=0.10 0.15 0.20. As already mentioned, the figure 0.05 is chosen

[1] The table has been calculated by Peter Allerup, Professor in Statistics, Department ofEducation, Aarhus University.

since this has been the general level of significance applied in the studies analysed.

ss	alternative	n	m	Power
0.05	0.10	10	2	0.93
0.05	0.10	11	2	0.91
0.05	0.10	12	2	0.89
0.05	0.10	13	2	0.87
0.05	0.10	14	2	0.84
0.05	0.10	15	2	0.82
0.05	0.10	16	3	0.93
0.05	0.10	17	3	0.92
0.05	0.10	18	3	0.90
0.05	0.10	19	3	0.89
0.05	0.10	20	3	0.87
0.05	0.10	21	3	0.85
0.05	0.10	22	3	0.83
0.05	0.10	23	3	0.81
0.05	0.10	24	4	0.91
0.05	0.10	25	4	0.90
0.05	0.10	26	4	0.89
0.05	0.10	27	4	0.87
0.05	0.10	28	4	0.86
0.05	0.10	29	4	0.84
0.05	0.10	30	4	0.82
0.05	0.10	31	4	0.81
0.05	0.10	32	5	0.91
0.05	0.10	33	5	0.89
0.05	0.10	34	5	0.88
0.05	0.10	35	5	0.87
0.05	0.10	36	5	0.85
0.05	0.10	37	5	0.84
0.05	0.10	38	5	0.83
0.05	0.10	39	5	0.81
0.05	0.10	40	6	0.90
0.05	0.10	41	6	0.89
0.05	0.10	42	6	0.88
0.05	0.10	43	6	0.87
0.05	0.10	44	6	0.85
0.05	0.10	45	6	0.84
0.05	0.10	46	6	0.83
0.05	0.10	47	6	0.81
0.05	0.10	48	7	0.90
0.05	0.10	49	7	0.89
0.05	0.10	50	7	0.88
0.05	0.10	51	7	0.87
0.05	0.10	52	7	0.86
0.05	0.10	53	7	0.84
0.05	0.10	54	7	0.83
0.05	0.10	55	7	0.82
0.05	0.10	56	7	0.81
0.05	0.10	57	8	0.89
0.05	0.10	58	8	0.88
0.05	0.10	59	8	0.87
0.05	0.10	60	8	0.86
0.05	0.10	61	8	0.85
0.05	0.10	62	8	0.84
0.05	0.10	63	8	0.83
0.05	0.10	64	8	0.81
0.05	0.10	65	8	0.80
0.05	0.10	66	9	0.88
0.05	0.10	67	9	0.87
0.05	0.10	68	9	0.86
0.05	0.10	69	9	0.85
0.05	0.10	70	9	0.84
0.05	0.10	71	9	0.83

ss	alternative	n	m	Power
0.05	0.10	72	9	0.82
0.05	0.10	73	9	0.81
0.05	0.10	74	10	0.88
0.05	0.10	75	10	0.87
0.05	0.10	76	10	0.86
0.05	0.10	77	10	0.86
0.05	0.10	78	10	0.85
0.05	0.10	79	10	0.84
0.05	0.10	80	10	0.83
0.05	0.10	81	10	0.82
0.05	0.10	82	10	0.81
0.05	0.10	83	11	0.88
0.05	0.10	84	11	0.87
0.05	0.10	85	11	0.86
0.05	0.10	86	11	0.85
0.05	0.10	87	11	0.84
0.05	0.10	88	11	0.83
0.05	0.10	89	11	0.82
0.05	0.10	90	11	0.81
0.05	0.10	91	11	0.80
0.05	0.10	92	12	0.87
0.05	0.10	93	12	0.86
0.05	0.10	94	12	0.86
0.05	0.10	95	12	0.85
0.05	0.10	96	12	0.84
0.05	0.10	97	12	0.83
0.05	0.10	98	12	0.82
0.05	0.10	99	12	0.81
0.05	0.10	100	12	0.80
0.05	0.15	10	2	0.82
0.05	0.15	11	3	0.93
0.05	0.15	12	3	0.91
0.05	0.15	13	3	0.88
0.05	0.15	14	3	0.85
0.05	0.15	15	3	0.82
0.05	0.15	16	4	0.92
0.05	0.15	17	4	0.90
0.05	0.15	18	4	0.88
0.05	0.15	19	4	0.86
0.05	0.15	20	4	0.83
0.05	0.15	21	4	0.80
0.05	0.15	22	5	0.90
0.05	0.15	23	5	0.88
0.05	0.15	24	5	0.86
0.05	0.15	25	5	0.84
0.05	0.15	26	5	0.82
0.05	0.15	27	6	0.90
0.05	0.15	28	6	0.88
0.05	0.15	29	6	0.87
0.05	0.15	30	6	0.85
0.05	0.15	31	6	0.83
0.05	0.15	32	6	0.81
0.05	0.15	33	7	0.89
0.05	0.15	34	7	0.87
0.05	0.15	35	7	0.86
0.05	0.15	36	7	0.84
0.05	0.15	37	7	0.82
0.05	0.15	38	8	0.89
0.05	0.15	39	8	0.88
0.05	0.15	40	8	0.86
0.05	0.15	41	8	0.85
0.05	0.15	42	8	0.83

ss	alternative	n	m	Power
0.05	0.15	43	8	0.81
0.05	0.15	44	9	0.89
0.05	0.15	45	9	0.87
0.05	0.15	46	9	0.86
0.05	0.15	47	9	0.84
0.05	0.15	48	9	0.83
0.05	0.15	49	9	0.81
0.05	0.15	50	10	0.88
0.05	0.15	51	10	0.87
0.05	0.15	52	10	0.85
0.05	0.15	53	10	0.84
0.05	0.15	54	10	0.82
0.05	0.15	55	10	0.81
0.05	0.15	56	11	0.87
0.05	0.15	57	11	0.86
0.05	0.15	58	11	0.85
0.05	0.15	59	11	0.83
0.05	0.15	60	11	0.82
0.05	0.15	61	11	0.80
0.05	0.15	62	12	0.87
0.05	0.15	63	12	0.86
0.05	0.15	64	12	0.85
0.05	0.15	65	12	0.83
0.05	0.15	66	12	0.82
0.05	0.15	67	12	0.80
0.05	0.15	68	13	0.87
0.05	0.15	69	13	0.86
0.05	0.15	70	13	0.84
0.05	0.15	71	13	0.83
0.05	0.15	72	13	0.82
0.05	0.15	73	13	0.80

ss	alternative	n	m	Power
0.05	0.15	74	14	0.86
0.05	0.15	75	14	0.85
0.05	0.15	76	14	0.84
0.05	0.15	77	14	0.83
0.05	0.15	78	14	0.82
0.05	0.15	79	14	0.80
0.05	0.15	80	15	0.86
0.05	0.15	81	15	0.85
0.05	0.15	82	15	0.84
0.05	0.15	83	15	0.83
0.05	0.15	84	15	0.81
0.05	0.15	85	15	0.80
0.05	0.15	86	16	0.86
0.05	0.15	87	16	0.85
0.05	0.15	88	16	0.84
0.05	0.15	89	16	0.83
0.05	0.15	90	16	0.81
0.05	0.15	91	16	0.80
0.05	0.15	92	17	0.86
0.05	0.15	93	17	0.85
0.05	0.15	94	17	0.84
0.05	0.15	95	17	0.83
0.05	0.15	96	17	0.81
0.05	0.15	97	17	0.80
0.05	0.15	98	18	0.86
0.05	0.15	99	18	0.85
0.05	0.15	100	18	0.84
0.05	0.20	10	3	0.88
0.05	0.20	11	3	0.84
0.05	0.20	12	4	0.93
0.05	0.20	13	4	0.90

ss	alternative	n	m	Power
0.05	0.20	14	4	0.87
0.05	0.20	15	4	0.84
0.05	0.20	16	5	0.92
0.05	0.20	17	5	0.89
0.05	0.20	18	5	0.87
0.05	0.20	19	5	0.84
0.05	0.20	20	5	0.80
0.05	0.20	21	6	0.89
0.05	0.20	22	6	0.87
0.05	0.20	23	6	0.84
0.05	0.20	24	6	0.81
0.05	0.20	25	7	0.89
0.05	0.20	26	7	0.87
0.05	0.20	27	7	0.84
0.05	0.20	28	7	0.82
0.05	0.20	29	8	0.89
0.05	0.20	30	8	0.87
0.05	0.20	31	8	0.85
0.05	0.20	32	8	0.83
0.05	0.20	33	9	0.89
0.05	0.20	34	9	0.87
0.05	0.20	35	9	0.85
0.05	0.20	36	9	0.83
0.05	0.20	37	9	0.81
0.05	0.20	38	10	0.88
0.05	0.20	39	10	0.86
0.05	0.20	40	10	0.84
0.05	0.20	41	10	0.82
0.05	0.20	42	11	0.88
0.05	0.20	43	11	0.86
0.05	0.20	44	11	0.85
0.05	0.20	45	11	0.83
0.05	0.20	46	11	0.80
0.05	0.20	47	12	0.87
0.05	0.20	48	12	0.85
0.05	0.20	49	12	0.83
0.05	0.20	50	12	0.81
0.05	0.20	51	13	0.87
0.05	0.20	52	13	0.86
0.05	0.20	53	13	0.84
0.05	0.20	54	13	0.82
0.05	0.20	55	13	0.80
0.05	0.20	56	14	0.86
0.05	0.20	57	14	0.85
0.05	0.20	58	14	0.83
0.05	0.20	59	14	0.81
0.05	0.20	60	15	0.87
0.05	0.20	61	15	0.85
0.05	0.20	62	15	0.84
0.05	0.20	63	15	0.82
0.05	0.20	64	15	0.80
0.05	0.20	65	16	0.86
0.05	0.20	66	16	0.85
0.05	0.20	67	16	0.83
0.05	0.20	68	16	0.81
0.05	0.20	69	17	0.87
0.05	0.20	70	17	0.85
0.05	0.20	71	17	0.84
0.05	0.20	72	17	0.82
0.05	0.20	73	17	0.80
0.05	0.20	74	18	0.86
0.05	0.20	75	18	0.84

ss	alternative	n	m	Power
0.05	0.20	76	18	0.83
0.05	0.20	77	18	0.81
0.05	0.20	78	19	0.86
0.05	0.20	79	19	0.85
0.05	0.20	80	19	0.84
0.05	0.20	81	19	0.82
0.05	0.20	82	19	0.81
0.05	0.20	83	20	0.86
0.05	0.20	84	20	0.84
0.05	0.20	85	20	0.83
0.05	0.20	86	20	0.81
0.05	0.20	87	21	0.86
0.05	0.20	88	21	0.85
0.05	0.20	89	21	0.84
0.05	0.20	90	21	0.82
0.05	0.20	91	21	0.81
0.05	0.20	92	22	0.86
0.05	0.20	93	22	0.84
0.05	0.20	94	22	0.83
0.05	0.20	95	22	0.82
0.05	0.20	96	22	0.80
0.05	0.20	97	23	0.85
0.05	0.20	98	23	0.84
0.05	0.20	99	23	0.82
0.05	0.20	100	23	0.81

Table 9.1: Power Calculation

10 Appendix 5: Indicators in education

Originally indicators were used in education to 'Measure the outcomes of educational institutions, programs, and practices, where resulting statistics are intended to inform educational policy'.[1] The intention to inform the policy is still there, but today indicators in education refer to context, input, process and outcome variables.[2] Most countries in the industrialized world apply some kind of indicator system in education. The international organisations working within education have taken up a leading role in the development of indicator systems. UNESCO is still active in this.[3] Today, however, OECD has taken the leading role in the field with its INES and TALIS projects.

Below we present a brief outline of indicator systems in education as they can be found in OECD and in individual countries.

10.1 OECD indicators

OECD publishes annual 'Education at a glance' OECD indicators. (OECD, 2009a) The annual report is the result of cooperation between OECD member state governments, INES, which is OECD's programme on indicators in education, and the OECD Secretariat. The publication offers an overview of the state of affairs in education

[1] This is actually the present Scope Note of 'Educational Indicators' offered in The ERIC Thesaurus.

[2] Scheerens (1990) gives a good introduction to the problems about relating process variables to process indicators in relation to the needs of policy-makers.

[3] The project World Education Indicators:

http://www.uis.unesco.org/ev.php?ID=5263_201&ID2=DO_TOPIC. A joint UNESCO OECD programme. See also UNESCO, 2005

in the OECD member states. The overview is given as data and text on a number of indicators.

In the choice of indicators the OECD project seems to have balanced pragmatically different points of view: firstly to assure internationally reliable and comparative data while aiming at data on matters of importance on the education policy agenda, and secondly to secure simplicity in presentation while aiming at retaining the possibility of reflecting the complexity in the matters described. The number of indicators is deliberately kept low.[4]

The OECD education indicator system has three levels:

On level one are Themes, namely these four:

Output and impact

Resources invested

Access, participation and progression

Organisation- and learning environment of schools

On level two, questions are found. Every theme is unfolded in a number of questions. For instance, the theme 'output and impact' has questions like: To what level have adults studied? How many students finish secondary education and access tertiary education?

The questions denote the indicators.

On level three are answers. Every question is answered under different definitions. For instance, 'To what level have adults studied?' is answered as: educational attainment of adult population, population

[4] This is how the project is described in the foreword of OECD,2009

with at least upper secondary education, population with tertiary education etc.

The Two upper levels (Themes and Questions) of the indicator system are:

A. Output of educational institutions and the impact of learning (9 indicators)

Indicator A1 To what level have adults studied?

Indicator A2 How many students finish secondary education and access tertiary education?

Indicator A3 How many students finish tertiary education?

Indicator A4 What is the profile of 15-year-old top performers in science?

Indicator A5 What are the top performers' attitudes and motivations for science in PISA 2006?

Indicator A6 How does participation in education affect participation in the labour market?

Indicator A7 What are the economic benefits of education?

Indicator A8 What are the incentives to invest in education?

Indicator A9 What are the social outcomes of education?

B. Financial and human resources invested in education (7 indicators)

Indicator B1 How much is spent per student?

Indicator B2 What proportion of national wealth is spent on education?

Indicator B3 How much public and private investment is there in education?

Indicator B4 What is the total public spending on education?

Indicator B5 How much do tertiary students pay and what public subsidies do they receive?

Indicator B6 On what resources and services is education funding spent?

Indicator B7 Which factors influence the level of expenditure?

C. Access to Education, Participation and Progression (3 indicators)

Indicator C1 Who participates in education?

Indicator C2 Who studies abroad and where?

Indicator C3 How successful are students in moving from education to work?

D. The Learning Environment and Organisation of Schools (6 indicators)

Indicator D1 How much time do students spend in the classroom?

Indicator D2 What is the student-teacher ratio and how big are classes?

Indicator D3 How much are teachers paid?

Indicator D4 How much time do teachers spend teaching?

Indicator D5 How much appraisal and feedback do teachers receive, and what is the impact?

Indicator D6 How do teacher practices, beliefs and attitudes measure up?

Frame 10.1: The Two upper levels (Themes and Questions) of the OECD indicator system

Themes B and D contain matters which are treated in the present systematic review.

Themes A and C, however, are in the present systematic review only treated as criteria for 'the good school', i.e. to give information about whether a school phenomenon or factor has positive effects on pupils.

It can be discussed to what extent the indicators actually selected by the OECD in themes B and D are grounded in evidence. Are the indicators selected which have the largest impact on school output or outcome? Or are the indicators selected which just have some impact on output or outcome? Or can we be sure that the indicators selected have an impact on output or outcome?

The OECD indicator project has been criticised for not leaving sufficient room for taking into account the efforts of the teacher, and the teaching/learning process in the classroom.[5] The TALIS project (OECD, 2009b) has opened up for this and has also taken in aspects from school management. However, the TALIS project does not yet cover 'the school' in general with indicators.

The first report from TALIS (OECD, 2009b) describes the project in the foreword as "a programme of surveys, with successive rounds designed to address policy-relevant issues chosen by countries."

There is a focus on the following aspects in lower secondary education:

School leadership

[5] OECD's TALIS project can be seen as a reaction to this criticism. (OECD,2005 & OECD, 2009b)

Appraisal and feedback to teachers

Teaching practices, beliefs and attitudes

Teachers' professional development

It is the aim of TALIS to develop indicators in these four fields. The independent variables applied in the first version of the project can be seen as the first proposal for new indicators. The independent variables chosen in the project are:

School socio-economic background

Teacher level: ability of students in class lower than the average at the same grade level

Teacher level: ability of students in class higher than the average at the same grade level

Teacher level: percentage of students in class speaking instruction language

Teacher level: percentage of students in class with at least one parent with completed ISCED 5 or higher

School level: percentage of students in school speaking instruction language

School level: percentage of students in school with at least one parent with completed ISCED 5 or higher

School level: ability of students in class lower than the average

School level: ability of students in class higher than the average

Bloc 1: Teacher characteristics

Female teacher

Teacher employed full-time

Teacher employed on a permanent contract

Teacher's education: above bachelor degree

Number of years for teaching

Bloc 2: Teacher professional development

Number of days for professional development

School providing induction process for teachers

School providing mentor for new teachers

Bloc 3: Teacher beliefs and practices

Index of teacher-student relations

Index of classroom teaching practice: structuring

Index of classroom teaching practice: student-oriented

Index of classroom teaching practice: enhanced activities

Index of direct transmission beliefs about instruction

Index of constructivist beliefs about instruction

Index of exchange and co-ordination for teaching

Index of professional collaboration

Bloc 4: Teacher appraisal and feedback

Never received appraisal or feedback from any source

Never received a school evaluation within the last 5 years

Teacher perceives that effective teachers receive more monetary or non-monetary rewards in the school

Important aspect for teacher appraisal: student test scores

Important aspect for teacher appraisal: innovative teaching practices

Important aspect for teacher appraisal: professional development the teacher has undertaken

Teacher appraisal and feedback impact: a change in salary

Teacher appraisal and feedback impact: opportunities for professional development activities

Teacher appraisal and feedback impact: public-private recognition from the principal and/or your colleagues

Teacher appraisal and feedback impact: changes in the teacher's work responsibilities that make the job more attractive (1=moderate or large change; 0=others)

School evaluation published

Important aspect for school evaluations: student test scores

Bloc 5: School leadership

Index of management-school goals

Index of instructional management

Index of direct supervision of instruction in the school

Index of accountable management

Index of bureaucratic management

Bloc 6: School autonomy and resources

Index of school climate: student delinquency

Index of school climate: teachers' working morale
Index of a lack of personnel
Index of school resources: shortage of materials
Index of school autonomy in hiring teachers, determining salaries
Index of school autonomy in budgeting (formulating and allocating the school budget)
Index of school autonomy: student policy and textbooks
Index of school autonomy in curriculum (courses offered, course content)
School average class size
Public school

Frame 10.2: The independent variables chosen in the TALIS project

TALIS does indeed offer new possible indicators. These could be added to the ones from 'Education at a Glance'. However no evidence is available for selecting one over the other, based on the indicators' effect on output or outcomes in pupils. Because of its survey character, the TALIS project does not offer 'cause-effect' explanations. Furthermore, the informants to the study dealing with teachers and school managers are to a large degree teachers and school managers themselves. So there is a risk of circularity in the research. In future, though, we may see attempts to introduce more of the indicators from 'Education at a glance' into TALIS.

10.2 National indicators

A recent systematic review offers an overview of outcome indicators used in high-performing education systems (Husbands et al., 2008).[6] What follows will be based on this.

Almost all analyzed countries use achievement as a child outcome indicator. Achievement refers often - but not always - to specific subjects. Achievement in subjects like literacy in the national language, mathematics and science are used as indicators in many countries. Several countries also apply the data on achievement from comparative surveys like PIRLS, TIMMS and PISA as indicators.

In Ireland and The Netherlands a broader range of indicators is applied: an 'Educational Careers Cohort 5-18 Survey on pupils' social and emotional development'; environmental factors about the home and school and citizenship competencies.

Most of the countries have national compulsory testing while some countries also offer voluntary tests. When it comes to timing of the achievement measurement, this is most commonly done at the end of compulsory schooling. However several countries apply achievement measures several times: after primary and lower secondary and after upper secondary. Only a few of the countries measure children on entry to school.

Another outcome indicator widely used is participation in education or employment. An indicator on enrolment in school or pre-school is common. Less frequently, indicators like actual attendance, truancy, grade repetition and suspensions are used.

---

[6] The use of output/outcome indicators was studied in Australia, Belgium, Denmark, Finland, Hungary, Ireland, Japan, Korea, The Netherlands, New Zealand, Singapore, Sweden and Switzerland. Indicators on children's education, health and well-being were all analysed. Only indicators in education are dealt with in this appendix.

In secondary education the highest number of participation indicators can be found: participation in education and employment/unemployment outcomes are both measured. Indicators in widespread use are: school completion rates, drop-out rates, participation in post-secondary education and employment outcomes of secondary education.

Some countries apply equity indicators i.e. they make it possible to have outcome indicators in relation to pupils with special educational needs. Such indicators could also be based on the gender and ethnic origin of the pupils.

Indicators on progress in achievement are also in use in some countries. Value-added measures which apply regression analysis can be used here. The calculated residual effects on pupils' achievement, which does not come from pupils' socio-economic background, gender or ethnic background are used as an indicator of what the school contributes to achievement. The value-added model focuses on the contributions in outcome and output which actually come from the schools. As such it can be considered a concrete clarification which could be added to any outcome or output indicator.[7]

10.3 The selection of indicators revisited

The discussion whether indicators should be selected and applied in education is for obvious reasons left aside here.

When the question of indicators comes up in education we can detect a tendency to concentrate on output or outcome variables. However, the OECD endeavours to take process indicators into consideration as well.

[7] A recent 'best practice study' of the 'value added model' is OECD, 2008.

When indicators are being selected, several considerations are involved:

Is the content of the indicator something which is high on the education policy agenda?

Is it possible to create a system of indicators which has simplicity as an important feature?

Can reliable measures be found of what the indicators refer to ?

Is it possible to obtain the data on the measures without putting undue workloads on the education systems and the schools?

Is it possible to create indicators and measures which have definitional precision as an important feature, accompanied by the openness which would make it possible for different schools and education systems to provide data which in a reliable way could represent the differences?

The present systematic review does not have a bearing on the selection of output or outcome indicators. If we concentrated on process indicators, then to the considerations already mentionedcould be added:

Do the selected process-indicators consist of phenomena that have an adequate evidence base for yielding positive outcomes in pupils?

This systematic review looks for the evidence base of possible process-indicators. This is so because, other things being equal, indicators consisting of phenomena and factors in schools that are important for the results of schooling should be selected rather than indicators with low or no importance for the results.

11 Complete overview of references included in the research mapping

This is the total list of the 150 references included. Two references, indicated with *, were not available in time for them to be subjected to coding. The remaining 148 references refer to 111 studies analysed in this report, i.e. in many cases several references report different aspects of the same study. A reference containing an ED number refers to a report published by ERIC. Such reports are to be found in the database ERIC by searching for the ED-number indicated.

The systematic review is based on studies, not references. Therefore, sometimes more references in the the list refer to only one study. In those cases where only the name of one author without publication year is indicated in the main text, this name refers to the study, i.e. all references in this list by that author.

Adler, M. A., & Fisher, C. (2000). Early Reading Programs in High-Poverty Schools: Emerald Elementary Beats the Odds. Michigan: University of Michigan.

Adler, M. A., Taylor, B.M., Pearson, D.P. (2002). Serna Elementary School. In Teaching reading effective schools, accomplished teachers. Mahwah, NJ: LEA.

Bain, H.P.B., Fulton, B.D., Zaharias, J.B. & Johnston, J. (1990). Student/Teacher Achievement Ratio (STAR) Project. Tennesee: Tennessee State Department of Education.

Bain, H. P., & Jacobs, R. (1990). The Case for Smaller Classes and Better Teachers. Streamlined Seminar, 9(1 Se).

Bamburg, J. D., & Andrews, R. L. (1990). Instructional Leadership, School Goals, and Student Achievement: Exploring the Relationship between Means and Ends (No. (ED319783)).

Barth, P., Haycock, K., Jackson, H., Mora, K., Ruiz, P., Robinson, S. et al. (1999). Dispelling the Myth: High Poverty Schools Exceeding Expectations (No. (ED445140)).

Bearden, D., Bembry, K. & Babu, S. (1995). Effective Schools: Is There a Winning Combination of Administrators, Teachers, and Students? (No. (ED386801)).

Benton, T. (2003). Study of the performance of maintained secondary schools in England: National Foundation for Educational Research.

Binkowski, K., Cordeiro, P. & Iwanicki, E. (1995). A Qualitative Study of Higher and Lower Performing Elementary Schools (No. (ED386311)).

Bondi, L. (1991). Attainment at primary schools: an analysis of variations between schools. British Educational Research Journal, 17(3), p203-217.

Booker, K. C., Invernizzi, M.A. & McCormick, M. (2007). "Kiss Your Brain": A Closer Look at Flourishing Literacy Gains in Impoverished Elementary Schools. Reading Research and Instruction, 46(4), 315-339.

Borman, G. D., & Rachuba, L. T. (2000). The Characteristics of Schools and Classrooms Attended by Successful Minority Students (No. (ED443933)).

Borman, G. D., & Rachuba, L. T. (2001). Academic Success among Poor and Minority Students: An Analysis of Competing Models of School Effects (No. CRESPAR-R-52).

Bottoms, G., Han, L., & Presson, A. (2006). Urban Students Achieve When High Schools Implement Proven Practices. Research Brief: Southern Regional Education Board (SREB).

Bottoms, G., Presson, A., & Han, L. (2004). Rigor, Relevance and Relationships Improve Achievement in Rural High Schools: High School Reform Works When Schools Do the Right Things. Challenge to Lead: Southern Regional Education Board (SREB).

Breaux, G., Danridge, J., Pearson, P. D., Taylor, B. M., & Pearson, D. P. (2002). Scott Elementary School: Home grown School improvement in the flesh. In Teaching reading effective schools, accomplished teachers. Mahwah, NJ: LEA.

Brown, K. E., & Medway, F. J. (2007). School climate and teacher beliefs in a school effectively serving poor South Carolina (USA) African-American students: A case study. Teaching and Teacher Education, 23(4), 529-540.

Bulach, C. R., Malone, B. & Castleman, C. (1995). An Investigation of Variables Related to Student Achievement. Mid-Western Educational Researcher, 8(2), 23-29.

Calaff, K. P. (2008). Supportive Schooling: Practices that Support Culturally and Linguistically Diverse Students' Preparation for College. NASSP Bulletin, 92(2), 95-110.

Campbell, R. J., Gagatsis, A., & Kyriakides, L. (2000). The significance of the classroom effect in primary schools: an application of Creemers' comprehensive model of educational effectiveness. School Effectiveness and School Improvement, 11(4), p 501-529.

Ceperley, P. E. (1999). Implementation of Title I Schoolwide Programs in Four Rural Virginia Schools (No. (ED438993)).

Choi, K., & Kim, J. (2006). Closing the Gap: Modeling Within-School Variance Heterogeneity in School Effect Studies (No. 689): National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Coates, D. (2003). Education Production Functions Using Instructional Time as an Input. Education Economics, 11(3), 273-292.

Coco, A., Frid, S., Galbraith, P., Gholam, M., Goos, M., Horne, M., Jolly, L., Kostogriz, A. Lincoln, D., Lowrie, T. & Short, T. (2004). Home, school and community partnerships to support children's numeracy: Dept. of Education, Science and Training.

Coe, P., Keyes, M., Meehan, M., Orletsky, S., Lewis, S., Rigney, S., Runge, C. & Whitaker, J. (1999). Development and Validation of Successful Writing Program Indicators Based on Research in Continuously Improving and Continuously Declining Schools: Report of the Kentucky State Writing Project. Study of Writing Instruction in Kentucky Schools (No. (ED442602)).

Colt, J., Mills, R., Taylor, B. M., & Pearson, D. P. (2002). Rocky Mountains Elementary schools, Longmont, Colorado. In Teaching reading effective schools, accomplished teachers. Mahwah, NJ: LEA.

Cooke, S. (2008). School-wide implementation of elements of effective classroom instruction: lessons from high-performing, high-powerty urban schools. University of Southern California.

Creemers, B. P. M., Reynolds, D. & Swint, F.E. (1996). The International School Effectiveness Research Project Quantitative and Class Study Data, 1992-1994 (No. (ED395380)).

De Fraine, B., Van Damme, J., Van Landeghem, G., Opdenakker, M.C., & Onghena, P. (2003). The Effect of Schools and Classes on Language Achievement. British Educational Research Journal, 29(6), 841-859.

De Fraine, B., van Landeghem, G., Van Damme, J., & Onghena, P. (2005). An Analysis of Well-Being in Secondary School with Multilevel Growth Curve Models and Multilevel Multivariate Models. Quality and Quantity, 39(3), 297-316.

Dronkers, J., & Robert, P. (2008). Differences in Scholastic Achievement of Public, Private Government-Dependent, and Private Independent Schools: A Cross-National Analysis. Educational Policy, 22(4), 541-577.

Dumay, X. & Dupriez, V. (2007). Accounting for Class Effect Using the TIMSS 2003 Eighth-Grade Database: Net Effect of Group Composition, Net Effect of Class Process, and Joint Effect. School Effectiveness and School Improvement, 18(4), 383-408.

Fend, H. (1998). Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung: Juventa.

Florida State Dept. of, E. (1994a). Chapter 1 Successful Schools. Pilot Project Report. 1993-94 (No. (ED383816)). Tallahassee: Florida State Dept. of Education.

Florida State Dept. of, E. (1994b). Chapter 1 Successful Schools. Technical Papers. 1993-94 (No. (ED383815)). Tallahassee: Florida State Dept. of Education.

Foley, E. M., Klinge, A., & Reisner, E. R. (2007). Evaluation of New Century High Schools: Profile of an Initiative to Create and Sustain Small, Successful High Schools. Final Report: Policy Studies Associates.

Franklin, B. J., & Crone, L. J. (1992). School Accountability: Predictors and Indicators of Louisiana School Effectiveness. Mid-South Educational Association (No. (ED354261))

Fullarton, S. (2004). Closing the gaps between schools : accounting for variation in mathematics achievement in Australian schools using TIMSS 95 and TIMSS 99. Paper presented at the IRC-2004 Conference, Cyprus.

Glidden, H. G. (1999). Breakthrough Schools: Characteristics of Low-Income Schools That Perform as Though They Were High-Income Schools. ERS Spectrum, 17(2), 21-26.

Griffith, J. (2002). A multilevel analysis of the relation of school learning and social environments to minority achievement in public elementary schools. Elementary School Journal, 102(5), 349-366.

Griffith, J. (2003). Schools as Organisational Models: Implications for Examining School Effectiveness. Elementary School Journal, 104(1), 29-47.

Grisay, A. (1994). Effective and Less Effective Junior Schools in France: A Longitudinal Study on the School Environment Variables Influencing the Student's Academic Achievement, Study Skills, and Socio-Affective Development (No. (ED380864)).

Grøgaard, J.B., Helland, H., & Lauglo, J. (2008). Elevenes læringsutbytte: Hvor stor betydning har skolen? Oslo: NIFO STEP.

Heck, R. H. (2007). Examining the Relationship between Teacher Quality as an Organisational Property of Schools and Students' Achievement and Growth Rates. Educational Administration Quarterly, 43(4), 399-432.

Hill, P. W., Holmes-Smith, P. & Rowe, K.J. (1993). School and Teacher Effectiveness in Victoria. Key Findings from Phase 1 of the Victorian Quality Schools Project (No. (ED367067)): International Congress for School Effectiveness and Improvement

Hill, P. W., Rowe, K. J., Holmes-Smith, P., & Russell, V. J. (1996). The Victorian Quality Schools Project : a study of school and teacher effectiveness: Centre for Applied Educational Research, Faculty of Education, University of Melbourne.

Hofman, R. H., Hofman, W. H. A., & Guldemond, H. (2002). School Governance, Culture, and Student Achievement. International Journal of Leadership in Education, 5(3), 249-272.

Hofman Roelande, H. (1996). Variation in effectiveness between private and public schools: the impact of school and family networks. Educational Research and Evaluation, 2(4), p366-394.

*Holtzapple, E. A. (1992). School Effects on Student SES/Achievement Relationships in Southwestern Pennsylvania, Conference of the American Sociological Association, .

Hopkins, M. S. (1999). Effective School Practices: What Works, International Conference on Effective Schools. Houston.

Hoy, W. K., Tarter, C.J. & Bliss, J.R. (1990). Organisational Climate, School Health, and Effectiveness: A Comparative Analysis. Educational Administration Quarterly, 26(3), 260-279.

Hughes, M. F. (1995). Achieving Despite Adversity: Why Are Some Schools Successful in Spite of the Obstacles They Face? A Study of the Characteristics of Effective and Less Effective Elementary Schools in West Virginia Using Qualitative and Quantitative Methods (No. (ED398004)).

Hughes, M. F. (1999). Similar Students--Dissimilar Opportunities for Success: High- and Low-Achieving Elementary Schools in Rural, High Poverty Areas of West Virginia. Journal of Research in Rural Education, 15(1), 47-58.

Jesse, D., Davis, A., & Pokorny, N. (2004). High-Achieving Middle Schools for Latino Students in Poverty. Journal of Education for Students Placed at Risk, 9(1), 23-45.

Johnson Jr, J. F., Taylor, B. M., & Pearson, D. P. (2002). High performing, high powerty, urban elementary schools. In Teaching reading effective schools, accomplished teachers. Mahwah, NJ: LEA.

Jones, S., Tanner, H., & Treadaway, M. (2000). Raising standards in mathematics through effective classroom practice. Paper presented at the Conference of the Australian Association for Research in Education, Sydney.

Jones, S. E. (2004). Studying "Success" at an "Effective" School: How a Nationally Recongnized Public School Overcomes Racial, Ethnic and Social Boundaries and Creates a Culture of Success. University of California.

Kennedy, E., Teddie, C. & Stringfield, S. (1993). A Multilevel Analysis of Phase II of the Louisiana School Effectiveness Study (No. (ED361902)).

Kitchen, R. S., DePree, J., Celedon-Pattichis, S., & Brinkerhoff, J. (2006). Mathematics Education at Highly Effective Schools that Serve the Poor: Strategies for Change.

Kyriakides, L., & Tsangaridou, N. (2008). Towards the development of generic and differentiated models of educational effectiveness: a study on school and teacher effectiveness in physical education. British Educational Research Journal, 34(6), 807-839.

Lamb, S., & Fullarton, S. (2002). Classroom and school factors affecting mathematics achievement: a comparative study of Australia and the United States using TIMSS. Australian Journal of Education, 46(2), 154-171.

Lamb, S., Rumberger, R., Jesson, D., & Teese, R. (2004). School performance in Australia: results from analyses of school effectiveness: Dept of Premier and Cabinet.

Lindsay, G., & Muijs, D. (2006). Challenging Underachievement in Boys. Educational Research, 48(3), 313-332.

Mahimuang, S. (2005). Factors influencing academic achievement and improvement: a value-added approach. Educational Research for Policy and Practice, 4(1), 13-26.

Mandeville, G. K., & Kennedy, E. (1993). A Longitudinal Study of the Social Distribution of Mathematics Achievement for a Cohort of Public High School Students (No. (ED365517)).

Martin, M. O., Mullis, I.S.V, Gregory, K.D., Hoyle, C. & Shen, C. (2001). Effective schools in science and mathematics. Iea's third international mathematics and science study.: Chestnut Hill, MA: Boston College.

*McCarthy, L. (2004). Structures and cultures found in effective schools : implications for school improvement. University of Newcastle, Newcastle.

Meelissen, M., & Luyten, H. (2008). The Dutch Gender Gap in Mathematics: Small for Achievement, Substantial for Beliefs and Attitudes. Studies in Educational Evaluation, 34(2), 82-93.

Meijnen, G. W., Lagerweij, N. W., & Jong, P. F. d. (2003). Instruction characteristics and cognitive achievement of young children in elementary schools. School effectiveness and school improvement, 14(2), 159-187.

Meijnen, G. W., & Sontag, L. (1997). Effective education for young children in primary schools. Educational studies, 23(1), 87-104.

Miles, K. H., & Darling-Hammond, L. (1997). Rethinking the Allocation of Teaching Resources: Some Lessons from High Performing Schools (No. (CPRE-RR-38)).

Mosenthal, J., Lipson, M., Mekkelsen, J., Russ, B., & Sortino, S. (2001). Elementary Schools Where Students Succeed in Reading (No. (ED459443)).

Mosenthal, J., Lipson, M., Torncello, S., Russ, B., & Mekkelsen, J. (2004). Contexts and practices of six schools successful in obtaining reading achievement. Elementary School Journal, 104(5), 343-367.

Mullis, I. V. S., Jenkins, F. & Johnson, E.G. (1994). Effective Schools in Mathematics: Perspectives from the NAEP 1992 Assessment. Research and Development Report (No. (NAEP-23-RR-01)).

Neumann, M., Schnyder, I., & Trautwein, U. (2007). Schulformen als differenzielle Lernmilieus. Institutionelle und kompositionelle Effekte auf die Leistungsentwicklung im Fach Französisch.; Paralleltitel: School types as differential learning environments: Institutional and compositional effects on achievement gains in French as a foreign language. Zeitschrift für Erziehungswissenschaft, 10(3), 399-420.

Opdenakker, M.C., & Damme, J.V. (2006). Differences between secondary schools. A study about school context, group composition, school practice, and school effects with special attention to public and Catholic schools and types of schools. School effectiveness and school improvement, 17(1), 87-117.

Opdenakker, M.C., & Damme, J.V. (2000). Effects of Schools, Teaching Staff and Classes on Achievement and Well-Being in Secondary Education: Similarities and Differences between School Outcomes. School Effectiveness and School Improvement, 11(2), 165-196.

Opdenakker, M.-C., & Van Damme, J. (2007). Do School Context, Student Composition and School Leadership Affect School Practice and Outcomes in Secondary Education? British Educational Research Journal, 33(2), 179-206.

Papanastasiou, C. (2008). A residual analysis of effective schools and effective teaching in mathematics. Studies in Educational Evaluation, 34(1), 24-31.

Perez, M., Anand, P., Speroni, C., Parrish, T., Esra, P., Socias, M. & Gubbins, P. (2007). Successful California Schools in the Context of Educational Adequacy: American Institutes for Research.

Picucci, A. C., Brownson, A., Kahlert, R., & Sobel, A. (2002a). Driven To Succeed: High-Performing, High-Poverty, Turnaround Middle Schools. Volume I: Cross-Case Analysis of High-Performing, High-Poverty, Turnaround Middle Schools (No. (ED476107)).

Picucci, A. C., Brownson, A., Kahlert, R., & Sobel, A. (2002b). Driven To Succeed: High-Performing, High-Poverty, Turnaround Middle Schools. Volume II: Case Studies of High-Performing, High-Poverty, Turnaround Middle Schools (No. (ED476108)).

Pollanen, S. (1991). Equity of Educational Achievement and School Effectiveness (No. (ED333021)).

Pollard-Durodola, S. (2003). Wesley Elementary: A Beacon of Hope for At-Risk Students. Education and Urban Society, 36(1), 94-117.

Postlethwaite, T. N., & Ross, K. N. (1992). Effective Schools in Reading: Implications for Educational Planners. An Exploratory Study (No. (ED360614)).

Pressley, M., Mohan, L., Raphael, L. M., & Fingeret, L. (2007). How does Bennett Woods Elementary School produce such high reading and writing achievement? Journal of Educational Psychology, 99(2), 221-240.

Pressley, M., Raphael, L., Gallagher, J. D., & DiBella, J. (2004). Providence-St. Mel School: How a School That Works for African American Students Works. Journal of Educational Psychology, 96(2), 216-235.

Pustjens, H., Van de gaer, E., Van Damme, J., & Onghena, P. (2008). Curriculum choice and success in the first two grades of secondary education: students, classes, or schools? School Effectiveness and School Improvement, 19(2), 155-183.

Reezigt, G. J., Guldemond, H., & Creemers, B. P. M. (1999). Empirical Validity for a Comprehensive Model on Educational Effectiveness. School Effectiveness and School Improvement, 10(2), 193-216.

Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., Schaffer, G., Bellin, W. & (2002). World class schools: international perspectives on school effectiveness. London: RoutledgeFalmer.

Ringsmose, C., & Mehlbye, J. (2004). Elementer i god skolepraksis: De gode eksempler. København: AKF Forlaget.

Rogers, W. T., Ma, X., Klinger, D. A., Dawber, T., Hellsten, L., Nowicki, D. & Tomkowicz, J. (2006). Examination of the Influence of Selected Factors on Performance on Alberta Learning Achievement Tests. Canadian Journal of Education, 29(3), 731-756.

Ross, J. A., & Gray, P. (2006b). School leadership and student achievement. The mediating effects of teacher beliefs. Canadian journal of education, 29(3), 798-822.

Ross Steven, M., Redfield, D. & Sterbinsky, A. (2006a). Effects of Comprehensive School Reform on student achievement and school change: a longitudinal multi-site study. School Effectiveness and School Improvement, 17(3), p367-397.

Rubenstein, M. C., & Wodatch, J. K. (2000). Stepping Up to the Challenge: Case Studies of Educational Improvement in Title I Secondary Schools. Manpower Demonstration Research (No. (PES-2000-12))

Rudd, P. (2002). High performing specialist schools: what makes the difference? Slough, Berkshire: National Foundation for Educational Research.

Rumberger, R. W., & Palardy, G. J. (2005). Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance. American Educational Research Journal, 42(1), 3-42.

Sammons, P., Thomas, S., Mortimore, P., Walker, A., Cairns, R. & Bausor, J. (1998). Understanding Differences in Academic Effectiveness: Practitioners' Views. School Effectiveness and School Improvement, 9(3), 286-309.

Sammons, P., Thomas, S., Mortimore, P., & Walker, A. (1997). Forging links: effective schools and effective departments. London: Paul Chapman.

Senkbeil, M., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., et al. (2006). Die Bedeutung schulischer Faktoren für die Kompetenzentwicklung in Mathematik und in den Naturwissenschaften. In PISA 2003. (pp. 277-308).

Silins, H., & Mulford, B. (2004). Schools as learning organisations: effects on teacher leadership and student outcomes. School Effectiveness and School Improvement, 15(3-4), 443-466.

Slinis, H. C., Mulford, W. R., & Zarins, S. (2002). Organisational learning and school change. Educational Administration Quarterly, 38(5), 613-642.

Smith, P. A., Johnson, J. E., Jones, B., Taylor, B. M., & Pearson, D. P. (2002). In pursuit of academic excellence: The story of Gladys Noon Spellman Elementary. In Teaching reading effective schools, accomplished teachers. Mahwah, NJ: LEA.

Smyht, E. (1999). Pupil Performance, Absenteeism and School Drop-out: A Multi-dimensional Analysis. School Effectiveness and School Improvement, 10(4), 480-502.

Smyth, E. (2000). Dimensions of School Effectiveness: Academic and Non-Academic Outcomes among Pupils in the Republic of Ireland (No. (ED443168)).

Spencerport Central Schools, N. Y. (1990). More Effective Schools/Teaching Project. Seventh Annual Report (No. (ED328524)): Spencerport Central Schools, NY.

Stringfield, S., Reynolds, D., & Schaffer, E. (2008). Improving secondary students' academic achievement through a focus on reform reliability: 4- and 9-year findings from the High Reliability Schools project. School Effectiveness and School Improvement, 19(4), 409-429.

Stringfield, S., & Teddlie, C. (1991). School, Classroom, and Student Level Indicators of Rural School Effectiveness. Journal of Research in Rural Education, 7(3), 15-28.

Sweetland, S. R., & Hoy, W. R. (2000). School Characteristics and Educational Outcomes: Toward an Organisational Model of Student Achievement in Middle Schools. Educational Administration Quarterly, 36(5), 703-729.

Tanner, H., Jones, S., & Treadaway, M. (2000). The role of the middle managers in raising standards in Mathematics, Conference of Australian Association for research in education (pp. 17). Sydney: Australian Association for research in education.

Tarter, C. J., & Hoy, W. K. (2004). A Systems Approach to Quality in Elementary Schools: A Theoretical and Empirical Analysis. Journal of Educational Administration, 42(5), 539-554.

Taylor, B., Pearson, D., Clark, K., & Walpole, S. (2000). Beating the Odds in Teaching All Students To Read: Lessons from Effective Schools and Accomplished Teachers (No. (ED450352)).

Taylor, B. M., Critchley, C., & Pearson, D. P. (2002). Sunnyside elementary school, mounds view, Minnesota. In Teaching reading effective schools, accomplished teachers. Mahwah, NJ: LEA.

Taylor, B. M., Pearson, D., Clark, K., Walpole, S., & Pearson, D. P. (2002). Effective schools and accomplished Teachers: Lessons about Primary grade reading instruction in low income schools. In Teaching reading effective schools, accomplished teachers. Mahwah, NJ: LEA.

Taylor, B. M., Pearson, P. D., Clark, K., & Walpole, S. (2000). Effective Schools and Accomplished Teachers: Lessons about Primary-Grade Reading Instruction in Low-Income Schools. Elementary School Journal, 101(2), 121-165.

Taylor, B. M., Pearson, P. D., Clark, K. F., & Walpole, S. (1999). Beating the Odds in Teaching All Children To Read (No. (CIERA-R-2-006)).

Teddlie, C., & Stringfield, S. (1992). Case Histories from a Longitudinal Study of School Effects (No. (ED346603)).

Teddlie, C., & Stringfield, S. (1993). Schools Make a Difference: Lessons Learned from a 10-Year Study of School Effects (No. (ED366054)).

Texas Education Agency (2000). The Texas Successful Schools Study: Quality Education for Limited English Proficient Students. Texas Education Agency, Austin (No. (AD00-300-01))

Thomas, S. (1995). Considering primary school effectiveness: an analysis of 1992 key stage 1 results. Curriculum Journal, 6(3), p279-295.

Thomas, S. (1997). Differential secondary school effectiveness: examining the size, extent and consistency of school and departmental effects on GCSE outcomes for different groups of students over three years. School Field, 8(1-2), p57-95.

Thomas, S., Mortimore, P., & Sammons, P. (1997). Differential school effectiveness: departmental variations in GCSE attainment. School Field, 8(1-2), p97-125.

Thomas, W. P., & Collier, V. (1997). School Effectiveness for Language Minority Students (No. (ED436087)).

Towns, D. P., Cole-Henderson, B., & Serpell, Z. (2001). The Journey to Urban School Success: Going the Extra Mile. Journal of Negro Education, 70(1-2), 4-18.

Traufler, V. J., & A. (1992). The relationship between student achievement and individual correlates of effective schools in selected schools of South Carolina. University of South Carolina.

Van Damme, J., De Fraine, B., Van Landeghem, G., Opdenakker, M.C., & Onghena, P. (2002). A Study on Educational Effectiveness in Secondary Schools in Flanders: An Introduction. School Effectiveness and School Improvement, 13(4), 383-397.

van der Werf, G. (1997). Differences in School and Instruction Characteristics between High-, Average-, and Low-Effective Schools. School Effectiveness and School Improvement, 8(4), 430-448.

van der Werf, G., & Weide, M. (1996). Differences in Quality of Instruction and Compensating Services between Effective and Ineffective Schools for Ethnic Minorities. Journal of At-Risk Issues, 3(1), 35-44.

Van Landeghem, G., et. al (2002a). The Effect of Schools and Classes on Mathematics Achievement. School Effectiveness and School Improvement, 13(4), 399-427.

Van Landeghem, G., et. al (2002b). The Effect of Schools and Classes on Noncognitive Outcomes. School Effectiveness and School Improvement, 13(4), 429-451.

Walpole, S., Taylor, B. M., & Pearson, D. P. (2002). Stevenson Elementary: Schoolwide succes. In Teaching reading effective schools, accomplished teachers. Mahwah, NJ: LEA.

Waxman, H. C., Garcia, A., & Read, L. L. (2008). Classroom Learning Environment & Student Motivational Differences between Exemplary, Recognized, & Acceptable Urban Middle Level Schools. Middle Grades Research Journal, 3(2), 1-21.

Webster, B. J., & Fisher, D. L. (2003). School-level environment and student outcomes in mathematics. Learning Environments Research, 6(3), 309-326.

Webster, B. J., & Fisker, D. L. (2001). School-level environment and student outcomes in mathematics achievement, Conference of the Australian Association for Research in Education. Melbourne: Australian Association for Research in Education

Willis, M. G. A. (1996). We're family: Creating success in an African American public elementary school. Georgia State University, Atlanta.

Witte, J. F., & Walsh, D. J. (1990). A Systematic Test of the Effective Schools Model. Educational Evaluation and Policy Analysis, 12(2), 188-212.

Woessmann, L. (2003). Educational production in East Asia: The impact of family background and schooling policies on student performance. Munich: University of Munich.

Young, D. (2001). The Western Australian school effectiveness study: Academe Consultancies.

Young, D. J. (1998a). Characteristics of Effective Rural Schools: A Longitudinal Study of Western Australian Rural High School Students.

Young, D. J. (1998b). Rural and Urban Differences in Student Achievement in Science and Mathematics: A Multilevel Analysis. School Effectiveness and School Improvement, 9(4), 386-418.

Young, D. J. (1999a). Case studies of Australian schools : what makes a school effective, Annual Conference of the British Educational Research Association (pp. 35). Brighton: BERA.

Young, D. J. (1999b). Student Progress in Australian Schools: A Multilevel Multivariate Model (No. (ED431020)).

Young, D. J. (1999c). The Usefulness of Value-Added Research in Identifying Effective Schools (No. (ED440810)).

Young, D. J., & Fisher, D. L. (1996). Western Australian School Effectiveness Study: a longitudinal study of rural schools. In C. B. a. P. d'Plesse (Ed.), Rural education: quality provision, quality experience quality outcomes: proceedings (pp. 101-129). Hobart: SPERA.

Young, D. J., & Fraser, B. J. (1992). School Effectiveness and Science Achievement: Are There Any Sex Differences? (No. (ED356946)).

Yu, L., & White, D. B. (2002). Measuring Value Added School Effects on Ohio Six-Grade Proficiency Test Results Using Two-Level Hierarchical Linear Modeling (No. (ED481653)).

Zigarelli, M. A. (1996). An Empirical Test of Conclusions from Effective Schools Research. Journal of Educational Research, 90(2), 103-110.

12 References

Listed here are only those references that are included in the textual commentary of the the systematic review. A complete list of references included in the research mapping is given in Chapter 11.

Adamson, L. (1996). Tranformational Leadership among Occupational Therapy Managers. In: K.w. Parry (Ed). *Leadership Research and Practice: Emerging Themes and New Challenges*. South Melbourne: Pitman Publishing/Woodslane, 55-64.

Borenstein, M. et al. (2009). Introduction to meta-analysis. Oxford : Wiley-Blackwell.

Blundell, R. & Costa Dias, M. (2007). Alternative Approaches to Evaluation in Empirical Microeconomics. Institute for Fiscal Studies.

Clearinghouse (2009): Work Paper on Effect size - all quantitative studies; November.

Concept Note. Danish Clearinghouse for Educational Research, http://www.dpu.dk/site.aspx?p=10634 (Verified 20 April 2010)

Creemers, B. et al. (1992). National and International School-Effectiveness Research in Retrospect and Prospect. (ERIC Report: ED347651.)

Creemers, B. & Kyriakides, L. (2008). The dynamics of educational effectiveness: a contribution to policy, practice and theory in contemporary schools. Abingdon: Routledge.

EPPI-Reviewer (http://eppi.ioe.ac.uk)

Husbands C., Shreeve A., Jones N.R. (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.

OECD (2009a) Education at a glance: OECD Indicators. Paris: OECD

OECD (2009b) Creating Effective Teaching and Learning Environments: First Results from TALIS. Paris: OECD

Petticrew, M. & H. Roberts (2003). Evidence, hierarchies, and typologies: horses for courses. Journal of Epidemiology and Community Health, 57, 527-9.

Petticrew, M. & H. Roberts (2006). Systematic Reviews in the Social Sciences. A Practical Guide. Malden: Blackwell Publishing.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., with Roen, K., & Duffy, W. (2006). Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. Version 1, April 2006. A Product from the ESRC Methods Programme.

Quinn, R.E. & Rohrbauch, J. (1983). A spatial model of effectiveness criteria: towards a competing values approach to organizational effectiveness. *Management Science*, 29, 363-377.

Rieper, O. & H. Foss Hansen (2007). Metodedebatten om evidens. København: AKF Forlaget.

Rumberger, R. W., & Palardy, G. J. (2005). Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance. American Educational Research Journal, 42(1), 3-42.

Scheerens, J. (1990). School Effectiveness Research and the Development of Process Indicators of School Functioning. School Effectiveness and School Improvement. 1, 61-80.

Scheerens, J. (1997). The foundations of educational effectiveness. Oxford: Pergamon.

Scheerens, J. (2000). Improving school effectiveness. *Fundamentals of Educational Planning – No. 68.* Paris: UNESCO.

Skolverket (2009). Vad påverkar resultaten I svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer. Stockholm: Skolverket.

Teddlie, C. & Reynolds, D. (2000). The international handbook of school effectiveness research. London: Falmer Press.

Tjeldvoll, A. (1998). Education and the Scandinavian welfare state in the year 2000: Equality, policy, and reform. New York: Garland Publishing.

Townsend, T. (2007). International handbook of school effectiveness and improvement. Dordrecht: Springer.

Uljens, M. (1997). School Didactics and Learning. Hove, East Surrex: Psychology Press.

UNESCO (2005) Education trends in perspective: Analysis of the world education indicators. Hamburg: UNESCO

Weiss, C.H. (1998). Evaluation. Second Edition. Printice Hall, Upper Saddle River, New Jersey.

Wholey, J. S. (1987). Organisational excellence: Stimulating quality and communicating value. Lexington Books, Lexington, MA.