British Educational Research Journal
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713406264

A synthesis of studies searching for school factors: implications for theory and research

Leonidas Kyriakides[a]; Bert Creemers[b]; Panayiotis Antoniou[a]; Demetris Demetriou[a]
[a] University of Cyprus, Cyprus [b] University of Groningen, Netherlands

First published on: 26 August 2009

To cite this Article Kyriakides, Leonidas , Creemers, Bert , Antoniou, Panayiotis and Demetriou, Demetris(2010) 'A synthesis of studies searching for school factors: implications for theory and research', British Educational Research Journal, 36: 5, 807 — 830, First published on: 26 August 2009 (iFirst)
To link to this Article: DOI: 10.1080/01411920903165603
URL: http://dx.doi.org/10.1080/01411920903165603

British Educational Research Journal
Vol. 36, No. 5, October 2010, pp. 807–830

A synthesis of studies searching for school factors: implications for theory and research

Leonidas Kyriakides[a*], Bert Creemers[b], Panayiotis Antoniou[a] and Demetris Demetriou[a]
[a]*University of Cyprus, Cyprus;* [b]*University of Groningen, Netherlands*

This paper reports the results of a meta-analysis in which the dynamic model of educational effectiveness is used as a framework to search for school factors associated with student achievement. The methods and results of a synthesis of 67 studies are presented. Findings reveal that effective schools are able to develop policies and take actions in order to improve their teaching practice and learning environment. Factors excluded from the dynamic model were found to be only weakly associated with outcomes. Implications for research on school effectiveness and for improvement of practice are drawn. It is illustrated that this approach of conducting meta-analysis helps us interpret the findings by providing support to the validity of the dynamic model and suggestions for its further development.

Introduction

Educational effectiveness research (EER) addresses the question of what works in education and why. Educational effectiveness research attempts to identify factors at different levels—student, teacher, school and system—associated with student outcomes. In this paper, the results of a synthesis of studies searching for school-level factors are presented. Although many studies of school factors have been conducted during the last two decades (Teddlie & Reynolds, 2000; Townsend, 2007) and each study resulted in a list of school factors associated with achievement, none of these studies can claim to be perfect. All studies contain measurement errors and, even independent of measurement error, no study's measures have perfect construct validity. Furthermore, there are typically other artefacts that distort study findings. Even if a hypothetical study suffered from none of these distortions, it would still contain sampling error. Given these inherent issues with individual studies, a large number of

*Corresponding author. Department of Education, University of Cyprus, Nicosia, Cyprus. Email: kyriakid@ucy.ac.cy

ISSN 0141-1926 (print)/ISSN 1469-3518 (online)/10/050807-24

DOI: 10.1080/01411920903165603

narrative reviews were conducted in the 1990s. Their main purpose was to compile state of the art knowledge in the EER field for the research community and policy-makers (e.g. Levine & Lezotte, 1990; Sammons, Hillman & Mortimore, 1995; Creemers & Reezigt, 1996). These reviews, however, were usually based on a collection of studies that were subjectively seen by the narrative review authors as good examples of research (e.g. Creemers & Reezigt, 1996; Sammons *et al.*, 1995) and the authors' judgements of methodological merit were often based on idiosyncratic ideas. On the other hand, some reviews were not selective at all, leading to a huge number of factors under consideration for which little empirical support was provided (Levine & Lezotte, 1990). As a consequence, the results of these reviews were questionable.

Although these reviews used schemes to categorise the factors included in the individual studies, the relation of each scheme and its categories with learning and learning outcomes was not made explicit. Thus, each category contained a heterogeneous group of unrelated factors that were not equally important in explaining learning outcomes. As a consequence, the importance of each category was underestimated due to the heterogeneity of the studies included. Thus, these reviews did not contribute significantly in the establishment of a knowledge base about the most important school-level factors (Scheerens & Bosker, 1997).

The main aims of the meta-analysis of school effectiveness studies

The meta-analysis reported here uses a dynamic model of educational effectiveness (Creemers & Kyriakides, 2008) as a framework to organise and structure the factors reported in the studies included in this review. This dynamic model is based on the assumption that school factors are expected to influence classroom-level factors, especially the teaching practice. The main assumptions of the model and its school factors are presented below. In order to estimate the effect size of school effectiveness factors on student achievement, we use a quantitative approach to synthesise the findings of multinational research studies conducted during the last 20 years. During the last two decades, many quantitative syntheses of studies searching for relationships between student achievement and factors such as homework, quality of teaching and parental involvement have been conducted (e.g. Seidel & Shavelson, 2007; Abrami *et al.*, 2008; Patall *et al.*, 2008). However, in this paper we draw attention to the importance of using a theoretical model to organise a meta-analysis. Through this meta-analysis, support for the validity of the model may emerge and ways to specify and develop the theoretical model further can be identified. Thus, the focus of the meta-analysis reported here is on the substantive findings that emerged from it. Specifically, we examine the extent to which the findings of the review justify the importance of school factors included in the dynamic model. In this way, this meta-analysis contributes to the establishment of a theoretical model by generating evidence supporting some of the factors accounted for and by pointing out possible weaknesses.

The second aim is to identify moderators that account for the variation in the observed effect sizes between studies, as effect sizes are likely to vary due to differences in procedures, instrumentation, study contexts and treatments. Identifying the impact

of these moderators (i.e. study characteristics) on the observed effect sizes gives further insight into the potential impact of school factors by clarifying the conditions under which each of them is able to influence effectiveness. For example, we can identify the extent to which the impact of factors upon student achievement depends on the criteria used for measuring effectiveness. Thus, the meta-analysis reported here may reveal school factors that are generic and/or other factors that have differential effects. In this way, the model may be developed further and its generic nature will be tested.

School factors in the dynamic model

The dynamic model of educational effectiveness is based on a critical review of the main findings of EER and of a critical analysis of the theoretical models of educational effectiveness, which were developed during the 1990s (e.g. Creemers, 1994; Scheerens, 1992; Stringfield & Slavin, 1992). Moreover, a synthesis of studies testing the validity of the comprehensive model of educational effectiveness (Creemers, 1994), which is considered the most influential theoretical construct in the field (Teddlie & Reynolds, 2000), reveals that some empirical support for this model has been provided through studies conducted in two different countries with more or less centralised educational systems (Kyriakides, 2008). It is also demonstrated that some characteristics of the comprehensive model can be seen as starting points for the development of the dynamic model, which is also able to establish strong links between effectiveness research and improvement of practice by taking into account the weaknesses of the comprehensive model. Finally, the results of these studies reveal the importance of the main characteristics of the dynamic model, which are as follows.

First, the dynamic model accounts for the fact that effectiveness studies conducted in several countries indicate that influences on student achievement are multilevel (Teddlie & Reynolds, 2000). As a result, the dynamic model refers to factors operating at the four levels shown in Figure 1. The teaching and learning situation and the roles of the two main actors (i.e. teacher and student) are analysed. Above these two levels, the dynamic model also refers to school- and context-level factors. It is expected that school-level factors influence the teaching-learning situation through the development and evaluation of school policies on teaching and on creating a learning environment. The context level refers to the influence of the educational system at large, especially through development and evaluation of the educational policy at the regional and/or national level. The model also accounts for the teaching and learning situation being influenced by the wider educational context in which students, teachers and schools are expected to operate. Factors such as the values of the society for learning and the importance attached to education play an important role both in shaping teacher and student expectations as well as in the development of perceptions of various stakeholders about effective teaching practice.

Figure 1 also illustrates the interrelationships between the model's components. In this way, the model supports that factors at the school and context level have both direct and indirect effects on student achievement since they are able to influence not only student achievement, but also the teaching and learning situations. This

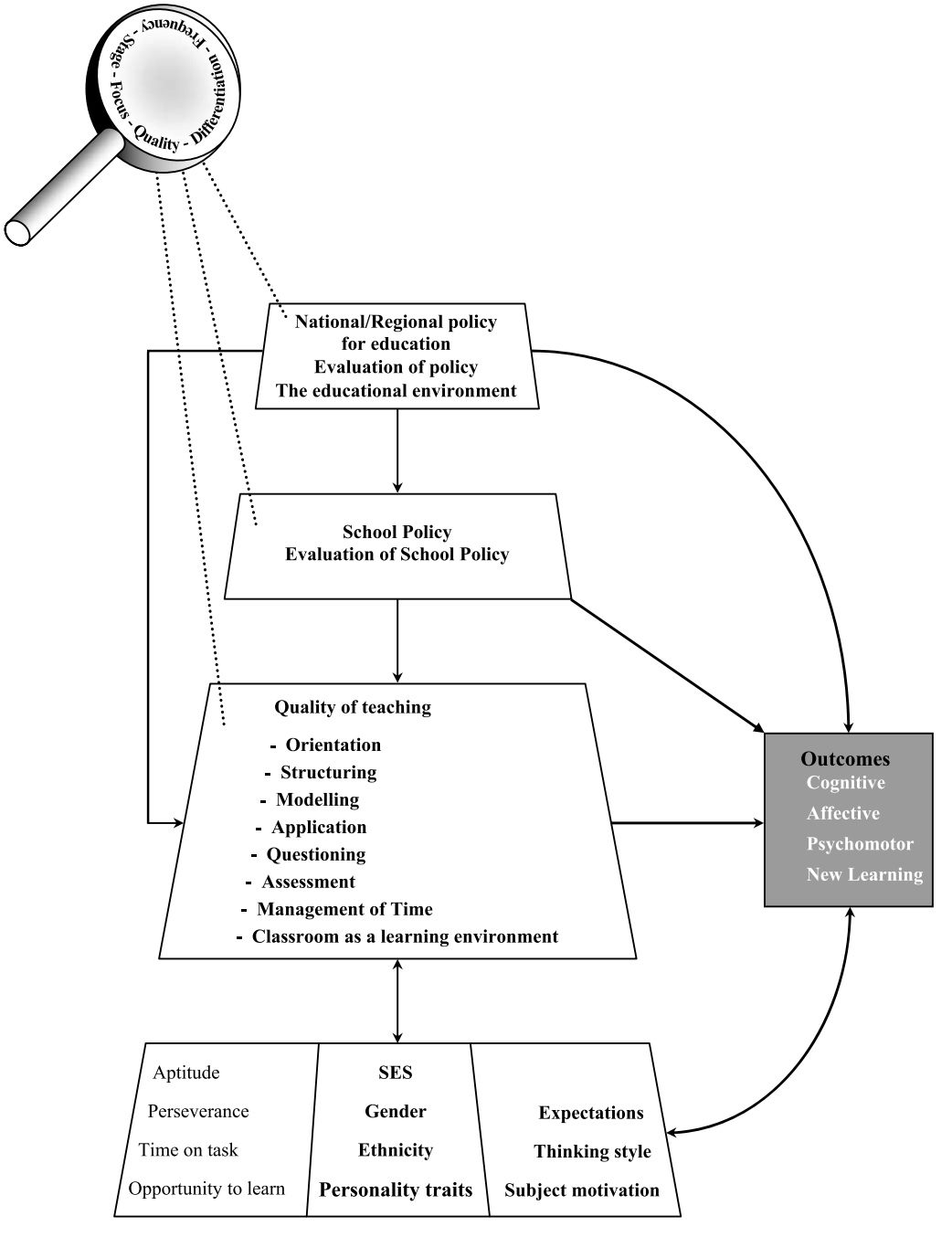

Figure 1. The dynamic model of educational effectiveness

assumption is supported by effectiveness studies conducted in order to test the validity of the comprehensive model (e.g. Kyriakides *et al.*, 2000; de Jong *et al.*, 2004; Kyriakides & Tsangaridou, 2008), which reveal that the relationships between factors at different levels might be more complex than assumed in the current integrated models. This is especially true for interaction effects among factors operating at the classroom and student levels, which reveal the importance of investigating differential effectiveness (Campbell *et al.*, 2003). Thus, the dynamic model demonstrates the complex nature of educational effectiveness.

Since this paper is concerned with school-level factors, a description of the dynamic model at the school level is provided below. Figure 1 reveals that the definition of the school level is based on the assumption that school factors are expected to have some direct effects, but mainly indirect effects, on student achievement. School factors are expected to influence classroom-level factors, especially the teaching practice. This assumption is based on EER studies that show the classroom level as more significant than the school and system levels (e.g. Teddlie & Reynolds, 2000; Kyriakides & Creemers, 2008) and that defining factors at the classroom level is a prerequisite for defining school- and system-level factors (Creemers, 1994). Therefore, the dynamic model refers to factors at the school level that are related to the key concepts of quantity of teaching, quality of teaching and provision of learning opportunities, which are the same factors used to define the classroom-level factors (see Creemers & Kyriakides, 2006). Specifically, the dynamic model emphasises two main aspects of school policy that affect learning at both the teacher and student level: (1) school policy for teaching and (2) school policy for creating a learning environment. These factors do not imply that each school should simply develop formal documents to articulate and install its policy; instead, the factors concerned with the school policy mainly refer to actions taken by the school to help teachers and other stakeholders have a clear understanding of what is expected from them. Support offered to teachers and other stakeholders to implement the school policy is also an aspect of these two factors.

Based on the assumption that the search for improvement underpins and defines the essence of a successful organization in the modern world (Kyriakides & Campbell, 2004), we also examine the processes and the activities that take place in the school to improve the teaching practice and its learning environment. For this reason, the processes used to evaluate school policy for teaching and the school learning environment are investigated. The following four factors at the school level are included in the model:

1. school policy for teaching and actions taken for improving teaching practice;
2. policy for creating a school learning environment and actions taken for improving the school learning environment;
3. evaluation of school policy for teaching and of actions taken to improve teaching; and
4. evaluation of the school learning environment.

1. School policy for teaching and actions taken for improving teaching. The dynamic model attempts to investigate aspects of school policy for teaching associated with quantity

of teaching, quality of teaching and provision of learning opportunities. Actions taken for improving the above three aspects of teaching practice, such as the provision of support to teachers for improving their teaching skills, are also taken into account. More specifically, the following aspects of school policy on quantity of teaching are taken into account:

- school policy on the management of teaching time (e.g. lessons start on time and finish on time; there are no interruptions of lessons for staff meetings and/or for preparation of school festivals and other events);
- policy on student and teacher absenteeism;
- policy on homework; and
- policy on lesson schedule and timetable.

School policy on provision of learning opportunities is measured by looking at the extent to which the school has an articulated and documented mission concerning the provision of learning opportunities. We also examine school policy on long-term and short-term planning and school policy on providing support to students with special needs. Furthermore, the extent to which the school attempts to make good use of school trips and other extra-curricular activities for teaching/learning purposes is investigated. Finally, school policy on the quality of teaching, which is closely related to the classroom-level factors of the dynamic model referring to the instructional role of teachers (see Creemers & Kyriakides, 2006), is examined.

Examining school policy for teaching in this context reveals that effective schools are expected to make decisions on maximizing the use of teaching time and the learning opportunities offered to their students. In addition, effective schools are expected to support their teachers in their attempt to help students learn by using effective teaching practices. In this context, the definition of the first school-level factor is such that we can identify the extent to which: (1) the school makes sure that teaching time is offered to students, (2) learning opportunities beyond those offered by the official curricula are offered to the students, and (3) the school attempts to improve the quality of teaching practice.

2. School policy for creating a school learning environment (SLE) and actions taken for improving the SLE. School climate factors have been incorporated in effectiveness models in different ways. Stringfield (1994) defines the school climate very broadly as the total environment of the school. This makes it difficult to study specific factors of the school climate and examine their impact on student achievement. On the other hand, Creemers (1994) defines climate factors more narrowly and expects them to exert influence on student outcomes in the same way as the effectiveness factors. The proposed dynamic model accounts for school effectiveness through the extent to which a learning environment has been created in the school. This element of school climate is seen as one of the most important predictors of school effectiveness since learning is the key function of a school. Moreover, EER has shown that effective schools are able to respond to the learning needs of both teachers and students and

to be involved in systematic changes of the school's internal processes. In this context, the following five aspects, which define the learning environment of the school, are taken into account:

1. student behaviour outside the classroom;
2. collaboration and interaction between teachers;
3. partnership policy (i.e. the relations of school with community, parents and advisors);
4. provision of sufficient learning resources to students and teachers; and
5. values in favour of learning.

The first three aspects refer to the rules that the school developed for establishing a learning environment inside and outside the classroom. Here the term *learning* does not refer exclusively to the student learning; for example, collaboration and interaction between teachers may contribute in their professional development (i.e. learning of teachers) but may also have an impact on teaching practice and thereby improve student learning. The fourth aspect refers to school policy on providing resources for learning. The availability of learning resources in schools may have not only an impact on student learning, but may also encourage the learning of teachers. For example, the availability of computers and software for teaching geometry may contribute to teacher professional development since it encourages teachers to find ways to make good use of the software in their teaching practice and thereby to become more effective. The fifth and final aspect of this factor is concerned with the strategies the school has established to encourage teachers and students to develop positive attitudes toward learning. Thus, the importance of the school climate is seen only in relation to the extent to which there is a learning environment within the school, which implies that values of the people not related with learning are not seen as effectiveness factors but may be related to schooling outcomes.

Following a similar approach to that concerned with school policy on teaching, the proposed dynamic model attempts to measure the school policy for creating a School Learning Environment (SLE). Actions taken for improving the SLE beyond the establishment of policy guidelines are also taken into account. More specifically, actions taken for improving the SLE can either be directed at: (1) changing the rules in relation to the first three aspects of the SLE factor mentioned above, (2) providing educational resources (e.g. teaching aids, educational assistance, new posts) or (3) helping students/teachers develop positive attitudes towards learning. For example, a school may have a policy promoting teacher professional development, but this might not be enough to improve the SLE, especially if some teachers do not consider professional development to be important. In that case, action should be taken to help teachers develop positive attitudes toward learning, which may help them become more effective.

The last two factors refer to the mechanisms used to evaluate the functioning of the first two factors concerned with school policy on teaching and the SLE. Creemers (1994) claims that control is one of the major principles operating in generating educational effectiveness. This implies that goal attainment and the school climate should be evaluated. It was, therefore, considered important to treat policy evaluation

for teaching and of other actions taken to improve teaching practice, as well as evaluation of the SLE, as factors operating at the school level. Data emerging from the evaluation mechanisms are expected to help schools develop their policies and improve the teaching practice at the classroom level as well as to improve their learning environment (Creemers & Kyriakides, 2008).

Methods

1. Selection of studies

Our research team conducted a search of documentary databases containing abstracts of empirical studies. The following databases were taken into account: Educational Resources Information Centre, Social Sciences Citation Index, Educational Administration Abstracts, SCOPUS, ProQuest 5000 and PsycArticles. We also paged through volumes of education-focused peer-reviewed journals with interest in EER, including the journals *School Effectiveness and School Improvement*, *British Educational Research Journal*, *Oxford Review of Education* and *Learning Environment Research*. Finally, relevant reviews of school effectiveness studies (e.g. Frazer *et al*., 1987; Levine & Lezotte, 1990; Sammons *et al*., 1995; Creemers & Reezigt, 1996; Scheerens & Bosker, 1997; Hallinger & Heck, 1998; Fan & Chen, 2001) and handbooks focused on effectiveness (e.g. Townsend *et al*., 1999; Teddlie & Reynolds, 2000) were examined for references to empirical studies.

2. Criteria for including studies

The following three criteria for selecting and including studies were used. First, we only selected studies conducted during the last 20 years that had been purposely designed to investigate the association of school factors with student outcomes. Most of the studies included in this meta-analysis were not designed to demonstrate causal relations between the factors and student achievement. Therefore, the effect sizes reported in this paper do not indicate the impact that school factors exercise on student achievement. Rather, the effect sizes reported refer to the strength of the association of school factors with student achievement. Second, the studies had to include explicit and valid measures of student achievement in relation to cognitive, affective or psychomotor outcomes of schooling. In addition, studies that used more global criteria for academic outcomes, such as drop out rates, grade retention and enrolment in top universities were also selected. These studies also had measures of specific school factors since they searched for relationships between these factors and student achievement. Thus, studies included in the meta-analysis had to provide information on the methods used to measure each school factor. As a result, we used minimal quality criteria for study selection with respect to the methods used to measure student outcomes and school factors, making it possible to include more studies in the meta-analysis. However, the method used to conduct the quantitative synthesis of these studies (see section 3, below) enables us to find out whether the variation in

the quality of measures has an impact on the observed effect sizes and to reconsider the selection criteria if needed. Finally, the meta-analysis reported here focuses on studies investigating the direct effects of school effectiveness factors upon student achievement. Only a few studies ($n = 5$) were found that attempted to examine the indirect effect of school effectiveness factors on achievement. Thus, the number of reported effect sizes ($n = 15$) in these five studies is insufficient for understanding the variation of indirect effect sizes for school level factors within and across studies. The studies included in the meta-analysis are listed in Appendix 1.

3. The computation of effect sizes

To indicate the effect of each school-level factor, we follow the same approach as Scheerens and Bosker (1997). Specifically, the Fisher's Z transformation of the correlation coefficient was used. Since not all studies presented their results in terms of correlations, all other effect size measures were transformed into correlations using the formulae presented by Rosenthal (1994). For small values of the correlation coefficient, Z_r and r do not differ much (see also Hunter & Schmidt, 2004). Furthermore, the Cohen's- d (Cohen, 1988) is approximately twice the size of the correlation coefficient when the latter is small (i.e. $-0.20 < r < 0.20$). Specifically, the three statistics d, t, and r are all algebraically transformable from one to the other.

The meta-analysis was conducted using MLwiN (Goldstein *et al.*, 1998). Specifically, we adopted the procedure suggested by Lamberts and Abrams (1995), which was also used in a number of meta-analyses conducted in the area of effectiveness (e.g. Scheerens & Bosker, 1997; Witziers *et al.*, 2003). More specific information on the multilevel model used to conduct the meta-analysis is given in Appendix 2. The multilevel analysis was conducted twice. In the first analysis, all studies were taken into account. In the second analysis, the so-called sensitivity analysis, outliers were removed from the samples to check the findings' robustness.

Findings

Table 1 provides information about the characteristics of studies investigating the relation of different school effectiveness factors with student achievement. In order to demonstrate the empirical support given to the factors of the dynamic model and the possible importance of factors not included yet, school factors are classified into those included in the dynamic model and those not included in the model. The average effect size of each factor is also provided. The following observations arise from this table. First, the values of the average effect sizes of the school effectiveness factors seem to support the argument that effective schools should develop a policy on teaching as well as a policy on establishing a learning environment. The six factors that belong to these two overarching school-level factors of the dynamic model were found to have an effect larger than 0.15. On the other hand, not enough data are available to support the importance of investigating the evaluation mechanisms that the schools develop to examine their own policies regarding teaching and their learning

environment. The lack of studies investigating the evaluation mechanisms of schools and the resulting improvement decisions may be attributable to the fact that only 8 out of 67 studies are longitudinal studies that took place for more than two school years. Moreover, the majority of the effectiveness studies (i.e. 55.2%) either investigated differences between schools in their effectiveness at a certain point in time or collected data at two time points only. Thus, there was no attempt to see the effect of school-level factors on changes in the school's effectiveness status.

Second, almost all the studies reviewed here collected data from either the primary and/or the secondary school level. Only four of them refer to effectiveness factors at the higher education level.

Third, the last part of Table 1 refers to the average effect size of factors not included in the dynamic model. We can see that a relatively high percentage of studies (i.e. 42.0%) measured the relationship between leadership and student achievement. However, the average effect size of this factor is very small. This implies that leadership has a very weak effect on student achievement. On the other hand, it could be claimed that the dynamic model should refer to the two school-level factors that were found to have an effect larger than 0.15: the school admission policy and teacher empowerment. However, only two studies tested the relationship between teacher empowerment and student achievement. Therefore, further studies are needed to investigate the generalisability of the observed effect size of this factor on student achievement. It should also be noted that very few studies have been conducted to assess the association between admission policy and student achievement ($n = 5$) and that this factor refers mainly to the school's input since the existence of an admission policy has a direct and strong impact on the students' prior knowledge before entering the school. Therefore, the school admission policy cannot be treated as a school effectiveness factor since it does not reveal how the functioning of schools can contribute to student achievement gains.

The next step in the meta-analysis was to use the multilevel approach to estimate the mean effect size of the following five factors: (1) policy on school teaching, (2) partnership policy (i.e. the relations of school with community, parents and advisors), (3) collaboration and interaction between teachers, (4) leadership and (5) school climate and culture. The multilevel analysis allows us to estimate the mean effect size of the first three factors of the dynamic model as well as the effect size of the two factors not included in the dynamic model that received significant attention in educational research (i.e. leadership and school climate and culture). In this way, we could justify our decision to include the first three factors in the model and exclude the other two. This approach also allowed us to examine whether the observed effect sizes vary across and within studies. For each of these five factors, a large variation in effect sizes within and across studies was identified. This implies that the multilevel analysis may also help us identify moderators that can explain variation across and within studies. Thus, the next step in our analysis examined which moderators can be held responsible for the variation in the effect sizes of the five school effectiveness factors. Table 2 shows the results of analyses that attempted to predict differences between effect sizes with such study characteristics as criteria of measuring effectiveness (i.e. the use of

Table 1. Characteristics of studies investigating the effect of school level factors on student achievement and types of effects identified

School-level factors	Average effect	Number of studies*	Outcomes** Cogn/affect/psych			Studies per sector Primary/secondry/both/other			
Factors included in the dynamic model									
1) Policy on teaching									
a) Quantity of teaching	0.16	18	14	6	2	9	6	2	1
b) Opportunity to learn	0.15	13	11	3	2	7	4	1	1
c) Quality of teaching	0.17	26	22	4	1	13	7	5	1
c1) Student assessment	0.18	12	10	4	0	8	3	1	0
2) Evaluation of policy on teaching	0.13	6	8	1	0	4	1	1	0
3) Policy on the school learning environment									
a) Collaboration	0.16	31	27	5	1	11	14	6	0
b) Partnership policy	0.17	21	14	9	0	8	10	3	0
4) Evaluation of policy on the learning school environment	–	0	0	0	0	0	0	0	0
Factors not included in the dynamic model									
1) Leadership	0.07	29	22	10	0	16	8	4	1
2) School climate	0.12	24	22	5	0	9	8	6	1
3) Autonomy	0.06	3	3	0	0	1	1	1	0
4) Teacher empowerment	0.17	2	2	0	0	0	2	0	0
5) Resources and working conditions (e.g. salary)	0.14	13	10	3	1	7	4	1	1
6) Admission policy, selection tracking	0.18	5	4	1	0	0	4	1	0
7) Staff experience (teachers and headteachers)	0.08	4	4	0	0	1	0	3	0
8) Job satisfaction	0.09	3	3	0	0	1	1	0	1

Notes. *Some studies reported more than one observed effect; **Some studies search for effects upon more than one type of outcomes of schooling.

Table 2. Predicting difference in effect sizes of each of the five school factors

Predictors	Policy on teaching estimate (*p* value)	Policy on SLE: partnership policy estimate (*p* value)	Policy on SLE: collaboration between teachers' estimate (*p* value)	Leadership estimate (*p* value)	School climate and school culture estimate (*p* value)
Intercept	0.18 (.001)	0.17 (.001)	0.16 (.001)	0.07 (.001)	0.12 (.001)
Language	−0.02 (.778)	−0.02 (.710)	0.01 (.834)	−0.01 (.812)	−0.02 (.792)
Drop out	0.01 (.482)	0.02 (.549)	0.01 (.793)	0.00 (.962)	0.02 (.781)
Non-cognitive	−0.05 (.005)*	−0.06 (.041)*	0.00 (.951)	0.04 (.082)	−0.02 (.542)
Secondary	−0.03 (.117)	−0.04 (.058)	0.01 (.741)	−0.05 (.038)*	0.03 (.41)
The Netherlands	−0.04 (.141)	0.01 (.812)	0.03 (.179)	−0.03 (.042)*	−0.07 (029)*
The UK	0.02 (.791)	0.00 (.955)	0.01 (.817)	0.01 (.816)	0.00 (.971)
Asian countries	0.02 (.548)	0.05 (.041)*	0.04 (.042)*	0.04 (.102)	0.03 (.041)*
All other countries	0.02 (.729)	0.01 (.805)	NA	0.04 (.115)	0.00 (.958)
Longitudinal	0.02 (.005)*	0.00 (.901)	0.01 (.829)	0.02 (.812)	0.01 (.712)
Experimental	NA	0.03 (.011)*	0.00 (.919)	NA	NA
Outlier	NA	0.01 (.887)	NA	−0.03 (.048)*	0.04 (.050)*
Unilevel versus multilevel	−0.06 (.103)	0.02 (.656)	0.01 (.804)	0.01 (.789)	0.02 (.612)

Notes. *A statistically significant effect at level .05 was identified; NA: It was not possible to test the effect of these explanatory variables since almost all the studies which assessed the impact of this factor belong to only one of the two groups that are compared; SLE = school learning environment.

different outcomes of schooling for measuring effectiveness), sector of education, country, study design employed and the use of multilevel rather than unilevel statistical techniques.

The following observations arise from Table 2. First, the results show that only in a few cases did moderators have a significant relationship with the effect size. Moreover, no moderator was found to have a significant relationship with the effect size of all five factors. However, looking at country differences, it appears that there are large discrepancies between several educational contexts. Specifically, the results for the two factors not included in the dynamic model are only true for studies conducted outside the Netherlands. In the Netherlands, the effect size of leadership is nearly zero and the effect size of school climate is very small (i.e. smaller than .06). Moreover, in the Asian countries, the effect sizes of the two main indicators of the factor concerned with the school policy for establishing a learning environment are higher. Country differences, however, were not identified in the case of the factor concerning the school policy on teaching, which is included in the dynamic model of educational effectiveness.

Second, the use of different criteria for measuring school effectiveness (i.e. criteria concerned with different types of schooling outcomes) is not associated with the effect size of three factors. However, two factors of the dynamic model (i.e. policy on teaching and partnership policy) were found to have a stronger relationship with achievement in cognitive outcomes rather than with achievement in non-cognitive outcomes. The only other significant effect pertains to the study design employed since it was found that longitudinal studies reported larger effect sizes for the first factor of the dynamic model concerning the school policy for teaching and experimental studies revealed larger effects of the factor concerned with partnership policy. Finally, on average, school leadership effects are almost absent in secondary education. School leadership effects are related to student achievement in primary schools but at a very small size.

As mentioned above, the meta-analysis' final step was to repeat the statistical procedure described above with the outliers removed from the samples to check the findings' robustness (i.e. the sensitivity study). The results showed that the effect of leadership style is greatly reduced when outliers are removed from the sample. Although this finding implies that there is still a positive and significant relationship between leadership and student outcomes, the indicator loses much of its relevance. On the other hand, the sensitivity study suggested that the other four factors do remain relevant even when outliers are removed from the sample.

Discussion

By looking at school factors examined by studies conducted in the last two decades, we can see that there are factors referring to the behaviour of persons in the organization (mainly the principal), the behaviour of persons related to the school (mainly parents), material conditions (e.g. quality of school curriculum), organizational elements (e.g. admittance policy) and school culture (e.g. orderly school climate).

Although all of these dimensions may be important for school effectiveness, it is necessary to arrange them according to theoretical notions to make their importance for student achievement easier to understand. The dynamic model attempts to provide such a framework by clarifying the relationships between school-level factors and classroom-level factors. It is emphasised that school factors are expected to influence classroom-level factors, especially the teaching practice. Thus, school-level factors are expected to influence learning that takes place in the classroom and/or the school. As a consequence, the dynamic model does not refer to a wide variety of factors as seen in earlier reviews.

The results of this meta-analysis reveal that at least the two overarching factors included in the dynamic model were associated with student achievement. Specifically, the estimated effect size of each of these two overarching factors was larger than 0.15. Moreover, the four components of the factor on policy for teaching as well as the two components of the factor concerned with the school policy on learning environment were significantly associated with student achievement and their effect sizes were larger than 0.15. Furthermore, the multilevel analysis attempting to identify the extent to which explanatory variables could predict variation in the estimated effect sizes of each factor reveal that these two factors of the dynamic model are generic in nature since the size of their effects does not vary in different countries. This implies that, irrespective of the country in which the study was conducted, effective schools are those that develop policies and take actions to improve teaching and the school learning environment. However, the effect size of the factor concerned with school policy on teaching was found to be larger when cognitive rather than non-cognitive outcomes are taken into account. Nevertheless, the effect size of this factor on non-cognitive outcomes is still larger than .10. This finding implies that this factor of the dynamic model is generic in nature but also has differential effects.

Moreover, the figures regarding the relation of student achievement to each of the eight factors not included in the dynamic model are not strong enough to suggest that they should be treated as relevant factors at the school level and be incorporated in a theoretical framework such as the dynamic model. An interesting case is leadership, which was considered an effectiveness factor in almost half of the studies conducted during the last two decades. The results of this meta-analysis revealed that leadership has a very weak effect on student outcomes. Moreover, its effect seems to disappear in secondary education and in some educational contexts. This finding is in line with the results of a meta-analysis conducted exclusively to identify the effect of leadership on student achievement (Witziers *et al.*, 2003), which applied multilevel modelling approaches to estimate the effect size of this factor. Similar results emerged from studies searching for indirect effects of leadership on student achievement (Leithwood & Jantzi, 2006). Therefore, school factors should not be concerned with who is in charge of designing and/or implementing the school policy, but with the content of the school policy and the type of activities that take place in school. Thus, support is provided to one of the major assumptions of the dynamic model, which is not focused on individuals as such but on the effects of the actions taking place at different levels of the educational system. At the school level, instead of measuring

the leadership of a principal, we should search for factors that refer to the impact of the end result of leadership, such as the development of school policy on teaching.

Only teacher empowerment and school admission policy had relatively large effect sizes, but the estimation of the average effect size of the teacher empowerment factor is based on only two studies. This implies a need for further studies investigating the impact of teacher empowerment on student achievement in order to have a more reliable estimation of this factor's effect. Regarding the impact of school admission policy, although this factor refers to organisational elements of the school, its functioning has a direct impact on the prior knowledge and skills of student achievement and thereby should be considered an input rather than a process variable (Glasman & Binianimov, 1981).

Regarding the two factors focused on the evaluation mechanisms of the school, the studies included in this meta-analysis were not concerned with them. Although among the main categories of school factors mentioned by several reviewers (e.g. Levine & Lezotte, 1990; Sammons *et al.*, 1995; Creemers & Reezigt, 1996; Scheerens & Bosker, 1997; Hallinger & Heck, 1998) is the monitoring of student progress, a significant number of studies were only concerned with student evaluation. Almost no studies investigated the impact of evaluation mechanisms of any other aspect of school policy. Moreover, no study attempted to examine systematically the processes used by schools to improve their teaching practice and their learning environment. It can, therefore, be claimed that the theoretical framework used for this meta-analysis validates the importance of looking at the effect of school evaluation mechanisms in a more systematic way rather than by looking at student results alone. The two relatively new school-level factors included in the dynamic model and the operational definition attached to them may have significant implications for the design of effectiveness studies. Specifically, our emphasis on actions taken to evaluate and change school policy implies that longitudinal studies should be conducted to measure the impact of these factors on the effectiveness status of schools rather than investigating the relation between existing practice and student achievement.

Finally, methodological implications for conducting meta-analyses can be drawn. Usually meta-analyses are conducted for two main reasons. First, researchers are interested in appraising the cumulative existing knowledge in a field and the main aim is, therefore, to give specific answers about the effect of certain factors or interventions on some other variables. In this way, both policy makers and practitioners can make use of the results. For example, a meta-analysis concerned with the impact and forms of homework on student achievement can help stakeholders to develop policies at the national or local level in order to improve teaching practice. Second, researchers may also be interested to use the findings to build a new theory or to design future studies. However, the approach used in the meta-analysis reported here was relatively new. In our attempt to conduct a meta-analysis, we used a theoretical framework to guide the structure and classification of factors included in the analysis and to interpret the findings. Based on the results, evidence supporting the validity of this framework was generated and suggestions for the further development of the model emerged. It can, therefore, be claimed that using this approach to conduct meta-analyses will help us

not only to integrate the findings across studies, but also systematically test the validity of a theory and thereby contribute to a better basis for theory development in the area of educational effectiveness.

References

Abrami, C. P., Bernard, M. R., Borokhovski, E., Wade, A., Surkes, A.M., Tamim, R. & Zhang, D. (2008) Instructional interventions affecting critical thinking skills and dispositions: a Stage-1 meta-analysis, *Review of Educational Research,* 78(4), 1102–1134.

Campbell, R.J., Kyriakides, L., Muijs, R.D. & Robinson, W. (2003) Differential teacher effectiveness: towards a model for research and teacher appraisal, *Oxford Review of Education,* 29(3), 347–362.

Cohen, J. (1988) *Statistical power analysis of the behavioral sciences* (2nd edn) (New York, Academic Press).

Creemers, B. P. M. (1994) *The effective classroom* (London, Cassell).

Creemers B. P. M. & Reezigt G. J. (1996) School level conditions affecting the effectiveness of instruction, *School Effectiveness and School Improvement,* 7(3), 197–228.

Creemers, B. P. M. & Kyriakides, L. (2006) A critical analysis of the current approaches to modelling educational effectiveness: the importance of establishing a dynamic model, *School Effectiveness and School Improvement,* 17(3), 347–366.

Creemers, B. P. M. & Kyriakides, L. (2008) *The dynamics of educational effectiveness: a contribution to policy, practice and theory in contemporary schools* (London, Routledge).

de Jong, R., Westerhof K. J. & Kruiter J. H. (2004) Empirical evidence of a comprehensive model of school effectiveness: a multilevel study in mathematics in the first year of junior general education in the Netherlands, *School Effectiveness and School Improvement,* 15(1), 3–31.

Fan, X. & Chen, M. (2001) Parent involvement and students' academic achievement: a meta-analysis. *Educational Psychology Review,* 13(1), 1–22.

Fraser, B. J., Walberg, H. J., Welch, W. W. & Hattie, J. A. (1987) Syntheses of educational productivity research, *International Journal of Educational Research,* 11, 145–252.

Glasman, N. S. & Biniaminov, I. (1981) Input-output analyses of schools, *Review of Educational Research,* 51, 509–539.

Goldstein, H., Rasbash, J., Plewis, I., Draper, D., Browne, W., Yang, M., Woodhouse, G. & Healy, M. (1998) *A user's guide to MlwiN* (London, Institute of Education).

Hallinger, P. & Heck, H. R. (1998) Exploring the principal's contribution to school effectiveness: 1980–1995, *School Effectiveness and School Improvement,* 9(2), 157–191.

Hunter, J. E. & Schmidt, F. L. (2004) *Methods of meta-analysis: correcting error and bias in research findings* (2nd edn) (Thousand Oaks, CA, Sage).

Kyriakides, L. (2008) Testing the validity of the comprehensive model of educational effectiveness: a step towards the development of a dynamic model of effectiveness, *School Effectiveness and School Improvement,* 19(4), 429–446.

Kyriakides, L. & Campbell, R. J. (2004) School self-evaluation and school improvement: a critique of values and procedures, *Studies in Educational Evaluation,* 30(1), 23–36.

Kyriakides, L., Campbell, R. J. & Gagatsis, A. (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), 501–529.

Kyriakides, L. & Creemers, B. P. M. (2008) A longitudinal study on the stability over time of school and teacher effects on student learning outcomes, *Oxford Review of Education,* 34(5), 521–545.

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), 807–883.

Lamberts, P. C. & Abrams, K. R. (1995). Meta-analysis using multilevel models, *Multilevel Modelling Newsletter,* 7(2), 17–19.

Leithwood, K. & Jantzi, D. (2006). Transformational school leadership for large-scale reform: effects on students, teachers and their classroom practices, *School Effectiveness and School Improvement,* 17, 201–227.

Levine, D. U. & Lezotte, L. W. (1990). *Unusually effective schools: a review and analysis of research and practice* (Madison, WI, National Center for Effective Schools Research and Development).

Patall, A. E., Cooper, H. & Robinson, C. J. (2008) Parent involvement in homework: a research synthesis, *Review of Educational Research,* 78(4), 1102–1134.

Raudenbush, S. W. (1994) Random effects models, in: H. Cooper & L. V. Hedges (Eds) *The handbook of research synthesis* (New York, Russell Sage), 301–323.

Raudenbush, S. W. & Bryk, A. S. (1985) Empirical Bayes meta-analysis, *Journal of Educational Statistics,* 10, 75–98.

Rosenthal, R. (1994) Parametric measures of effect size, in: H. Cooper & L. V. Hedges (Eds) *The handbook of research synthesis* (New York, Russell Sage), 231–245.

Sammons, P., Hillman, J. & Mortimore, P. (1995) *Key characteristics of effective schools: a review of school effectiveness research* (London, Office for Standards in Education and Institute of Education).

Scheerens, J. (1992) *Effective schooling: research, theory and practice* (London, Cassell).

Scheerens, J. & Bosker, R. J. (1997) *The foundations of educational effectiveness* (Oxford, Pergamon).

Seidel, T. & Shavelson, R. J. (2007) Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results, *Review of Educational Research,* 77(4), 454–499.

Stringfield, S. C. & Slavin, R. E. (1992) A hierarchical longitudinal model for elementary school effects. In B. P. M. Creemers & G. J. Reezigt (Eds) *Evaluation of educational effectiveness* (Groningen, ICO) 35–69.

Stringfield, S. (1994) A model of elementary school effects, in: D. Reynolds, B. P. M. Creemers, P. S. Nesselrodt, E. C. Schaffer, S. Stringfield & C. Teddlie (Eds) *Advances in school effectiveness research and practice* (Oxford, Pergamon), 153–187.

Teddlie, C. & Reynolds, D. (2000) *The international handbook of school effectiveness research* (London, Falmer Press).

Townsend, T. (Ed.) (2007) *International handbook of school effectiveness and improvement* (Dordrecht, the Netherlands,. Springer).

Townsend, T., Clarke, P. & Ainscow, M. (Eds) (1999) *Third millennium schools. a world of difference in school effectiveness and school improvement* (Lisse, Swets and Zeitlinger).

Witziers, B., Bosker, J. R. & Kruger, L. M. (2003) Educational leadership and student achievement: the elusive search for an association, *Educational Administration Quarterly,* 39(3), 398–425.

Appendix 1. Studies investigating the effect of school level factors on student achievement used for the meta-analysis reported in this paper

Aksoy, T. & Link, C. R. (2000) A panel analysis of student mathematics achievement in the US in the 1990s: does increasing the amount of time in learning activities affect math achievement? *Economics of Education Review*, 19, 261–277.

Bamburg, J. D. & Andrews, R. L. (1991) School goals, principals and achievement. *School Effectiveness and School Improvement*, *2*(3), 175–191.

Bosker, R. J., Kremers, E.J.J. & Lugthart, E. (1990) School and instruction effects on mathematics achievement, *School Effectiveness and School Improvement*, 1, 233–248.

Bosker, R. J. & Scheerens, J. (1994) Alternative models of school effectiveness put to the test, *International Journal of Educational Research*, 21, 159–181.

Brewer, D. J. (1993) Principals and student outcomes, *Economics of Education Review*, 12(4), 281–292.

Brown, R. & Evans, P. W. (2002) Extracurricular activity and ethnicity: creating greater school connections among diverse student populations. *Urban Education*, 37(1), 41–58.

Cheng, Y. C. (1993) Profiles of organizational culture and effective schools, *School Effectiveness and School Improvement*, 4(2), 85–110.

Cheng, Y. C. (1994) Principal's leadership as a critical factor for school performance: evidence from multilevels of primary schools, *School Effectiveness and School Improvement*, 5, 299–317.

de Jong R., Westerhof K. J. & Kruiter J. H. (2004) Empirical evidence of a comprehensive model of school effectiveness: a multilevel study in mathematics in the first year of junior general education in the Netherlands, *School Effectiveness and School Improvement*, 15(1), 3–31.

DiPaola, F. M. & Moran, M. T. (2005) Bridging or buffering? The impact of schools' adaptive strategies on student achievement, *Journal of Educational Administration*, 43(1), 60–71.

Driessen, G. & Sleegers, P. (2000) Consistency of teaching approach and student achievement: an empirical test, *School Effectiveness and School Improvement*, 11(1), 57–79.

Durland, M. & Teddlie, C. (1996) A network analysis of the structural dimensions of principal leadership in differentially effective schools, Paper presented at the *Annual Meeting of the American Educational Research Association*, New York, April.

Eberts, R. & Stone, J. (1988) Student achievement in public schools: do principals make a difference? *Economics of Education Review*, 7, 291–299.

Fan, X. (2001) Parental involvement and students' academic achievement: a growth modelling analysis, *Journal of Experimental Education*, 70(1), 27–61.

Fan, X. & Chen, M. (2001) Parent involvement and students' academic achievement: a meta-analysis, *Educational Psychology Review*, 13(1), 1–22.

Farrow, S., Tymms, P. & Henderson, B. (1999) Homework and attainment in primary schools, *British Educational Research Journal*, 25(3), 323–341.

Frazer, B. J., Walberg, H. J., Welch, W. W. & Hattie, J. A. (1987) Syntheses of educational productivity research, *International Journal of Educational Research*, 11, 145–252.

Friedkin, N. E. & Slater, M. R. (1994) School leadership and performance: a social network approach, *Sociology of Education*, 67, 139–157.

Goldring, E. & Paternak, R. (1994) Principals coordinating strategies and school effectiveness, *School Effectiveness and School Improvement*, 5(3), 239–253.

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 organizational models: implications for examining school effectiveness, *Elementary School Journal*, 110(1), 29–47.

Guay, F., Boivin, M. & Hodges, E. V. E. (1999) Predicting change in academic achievement: a model of peer experiences and self-system processes, *Journal of Educational Psychology*, 91(1), 105–115.

Hallinger, P., Bickman, L. & Davis, K. (1996) School context, principal leadership and student reading achievement, *Elementary School Journal*, 96(5), 527–549.

Heck, R. (1993) School context, principal leadership and achievement: the case of secondary schools in Singapore, *Urban Review*, 25, 151–166.

Heck, R., Larson, T. & Marcoulides, G. (1990) Principal instructional leadership and school achievement: validation of a causal model, *Educational Administration Quarterly*, 26, 94–125.

Heck, R., Marcoulides, G. & Lang, P. (1991) Principal instructional leadership and school achievement: the application of discriminant techniques, *School Effectiveness and School Improvement*, 2, 115–135.

Hofman, R. H., (1995) Contextual influences on school effectiveness: the role of school boards, *School Effectiveness and School Improvement*, 6(4), 308–331.

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, W. H. A. & Guldemond, H. (1999) Social and cognitive outcomes: a comparison of contexts of learning, *School Effectiveness and School Improvement*, 10(3), 352–366.

Houtveen, A. A. M., van de Grift, W. J. C. M. & Creemers, B. P. M. (2004) Effective school improvement in mathematics, *School Effectiveness and School Improvement*, 15(3), 337–376.

Hoy, W. K., Tater, J. C. & Bliss, J. R. (1990) Organizational climate, school health and effectiveness: a comparative analysis, *Educational Administration Quarterly*, 26(3), 260–279.

Jimerson, S., Egeland, B. & Teo, A. (1999) A longitudinal study of achievement trajectories: factors associated with change, *Journal of Educational Psychology*, 91(1), 116–126.

Karlya, T. & Rosenbaum, J. E. (1999) Bright flight: unintended consequences of detracking policy in Japan, *American Journal of Education*, 107(3), 210–230.

Kyriakides, L. (2005) Extending the comprehensive model of educational effectiveness by an empirical investigation, *School Effectiveness and School Improvement*, 16(2), 103–152.

Kyriakides, L., Campbell, R. J. & Gagatsis, A. (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), 501–529.

Kyriakides, L. & Charalambous, C. (2005) Using educational effectiveness research to design international comparative studies: turning limitations into new perspectives, *Research Papers in Education*, 20(4), 391–412.

Kyriakides, L. & Creemers, B. P. M. (2008) Using a multidimensional approach to measure the impact of classroom level factors upon student achievement: a study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, 19(2), 183–306.

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), 807–83.

Kythreotis, A. (2006) *The influence of school leadership styles and culture on students' achievement in Cyprus primary schools*, Unpublished doctoral dissertation, University of Cyprus, Cyprus.

Lee, V. E. & Bryk, A. S. (1989) A multilevel model of the social distribution of high school achievement, *Sociology of Education*, 62(3), 172–192.

Leithwood, K. & Jantzi, D. (1999) The relative effects of principal and teacher sources of leadership on student engagement with school, *Educational Administration Quarterly*, 35, 679–706.

Leithwood, K. & Jantzi, D. (2006) Transformational school leadership for large-scale reform: effects on students, teachers and their classroom practices, *School Effectiveness and School Improvement*, 17, 201–227.

Leitner, D. (1994) Do principals affect student outcomes? *School Effectiveness and School Improvement*, 5, 219–238.

Ma, X. (1999) Dropping out of advanced mathematics: the effects of parental involvement, *Teachers College Record*, 101(1), 60–81.

Mandeville, G. K. & Kennedy, E. (1991) The relationship of effective school indicators and changes in the social distribution of achievement, *School Effectiveness and School Improvement*, 2, 14–33.

Marks, H. M. & Printy, S. M. (2003) Principal leadership and school performance: an integration of transformational and instructional leadership, *Educational Administration Quarterly*, 39(3), 370–397.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D. & Ecob, R. (1988) *School matters: the junior years* (Somerset, UK, Open Books).

Mudler, L. & Van de Werf, G. (1997) Implementation and effects of the Dutch educational priority policy: results of four years of evaluation studies, *Educational Research and Evaluation*, 3(4), 317–339.

Opdenakker, M. C. & Van Damme, J. (2001) Relationship between school composition and characteristics of school process and their effect on mathematics achievement, *British Educational Research Journal*, 27(4), 407–432.

Prithchard, R. J., Morrow, D. & Marshal, J. C. (2005) School and district culture as reflected in student voices and student achievement, *School Effectiveness and School Improvement*, 16(2), 153–177.

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.

Scheerens, J., Vermeulen, C. J. & Pelgrum, W. J. (1989) Generalisability of instructional and school effectiveness indicators across nations, *International Journal of Educational Research*, 13, 789–799.

Shann, M. H. (1999) Academics and a culture of caring: the relationship between school achievement and prosocial and antisocial behaviours in four urban middle schools, *School Effectiveness and School Improvement*, 10(4), 390–413.

Sillins, H. & Mulford, B. (2004) Schools as learning organizations: effects on teacher leadership and student outcomes, *School Effectiveness and School Improvement*, 15, 443–466.

Smyth, E. (1999) Pupil performance, absenteeism and school drop-out: a multi-dimensional analysis, *School Effectiveness and School Improvement*, 10(4), 480–502.

Smyth, E. & Hannan, C. (2006) School effects and subject choice: the uptake of scientific subjects in Ireland, *School Effectiveness and School Improvement*, 17(3), 255–273.

Sterbinsky, A., Ross. M. S. & Redfield, D. (2006) Effects of comprehensive school reform on student achievement and school change: a longitudinal multi-site study, *School Effectiveness and School Improvement*, 17(3), 367–397.

Sweetland, R. S. & Hoy, K. W. (2000) School characteristics and educational outcomes: towards an organizational model of student achievement in middle schools, *Educational Administration Quarterly*, 36(5), 703–729.

Teddlie, C. & Stringfield, S. (1993) *Schools make a difference: lessons learned from a 10-year study of school effects* (New York, Teachers College Press).

Willms, J. D. & Somers, M. (2001) Family, classrooms, and school effects on children's educational outcomes in Latin America, *School Effectiveness and School Improvement*, 12(4), 409–445.

Van de Grift, W. (1990) Educational leadership and academic achievement in elementary education, *School Effectiveness and School Improvement*, 1(3), 26–40.

Van de Grift, W. J. C. M. & Houtveen, A. A. M. (2006) Underperformance in primary schools, *School Effectiveness and School Improvement*, 17(3), 255–273.

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., Creemers, B., De Jong, R. & Klaver E. (2000) Evaluation of school improvement through an educational effectiveness model: the case of Indonesia's PEQIP Project, *Comparative Education Review*, 44(3), 329–355.

Van Houtte, M. (2004) Tracking effects on school achievement: a qualitative explanation in terms of the academic culture of school staff, *American Journal of Education*, 110(4), 354–388.

Webster, B. J. & Fisher, D. L. (2000) Accounting for variation in science and mathematics achievement: a multilevel analysis of Australian data. Third International Mathematics and Science Study (TIMSS), *School Effectiveness and School Improvement*, 11, 339–360.

Webster, B. J. & Fisher, D. L. (2003) School-level environment and student outcomes in mathematics, *Learning Environment Research*, 6, 309–326.

Appendix 2. The multilevel model used to conduct the meta-analysis

The multilevel model was applied to analyse the observed effects of each study and the sources of variances among the findings emerged from different studies investigating the relation between the same school-level factor and student achievement (Raudenbush & Bryk 1985). Specifically, studies that investigated the relationship of student achievement with another factor, such as the school policy on quantity of teaching, are considered to be a sample from the population of studies investigating the relationship between this factor and student achievement. Nested under each study are the schools. Each study can then be viewed as an independent replication. This concept, however, does not solve the problem of multiple results from one study, such as when effect sizes are reported for more than one outcome of schooling (e.g. mathematics and language achievement or mathematics and development of positive attitudes towards the school) in one study while using the same sample of schools and students. To deal with this problem, the two-level model for meta-analysis was expanded to a three-level model. As a consequence, the highest level of study results is referred to as the across-replication level and the multiple results within a study as the within-replication level. The main advantage of the statistical meta-analysis employed here is that the information from each study is weighted by the reliability of the information, in this case the sample size. Moreover, the multilevel model helps us identify factors responsible for the variation in observed effect sizes for each of the main school-level factors on student achievement that emerged from this synthesis of school effectiveness studies. Therefore, differences in reported effect sizes are modelled as a function of study characteristics, such as differences in the type of outcomes used to measure student achievement, the level of education at which the study was conducted and the nature of the study. Further information about the statistical modelling technique is given below.

The multilevel model for the meta-analysis (Raudenbush & Bryk, 1985; Raudenbush, 1994), starting with the 'within-replications model,' is given by the following equation:

$$d_{rs} = \delta_{rs} + e_{rs}$$

The above equation implies that the effect size *d* in replication *r* in study *s* (d_{rs}) is an estimate of the population parameter (δ_{rs}) and the associated sampling error (e_{rs}). The sampling error is attributed to the fact that in each replication only a sample of schools is studied. Concerning the between-replications model, the following equation is used:

$$\delta_{rs} = \delta_s + u_{rs}$$

In the above model, it is acknowledged that the true replication effect size is a function of the effect size in study s and sampling error u_{rs}. Finally, the between-studies model is formulated as follows:

$$\delta_s = \delta + u_{0s}$$

The above formula implies that the true unknown effect size as estimated in study *s* (δ_s) is a function of the effect size across studies (δ_0) with random sampling error v_s, which is attributed to the fact that the studies are sampled from a population of studies.

To assess the effects of the study characteristics, we extended the between-replication model to one that took into account the effect of explanatory variables, which are presented below. Since all explanatory variables were grouping variables, they were entered as dummies with one of the groups as baseline. Specifically, the following explanatory variables, which refer to the characteristics of the studies included in the meta-analysis, were entered in the model.

Outcomes of schooling

A grouping variable was available to search the impact of the type of schooling outcomes employed in each study. Specifically, it was possible to identify four types of outcomes that were taken into account by a relatively large number of studies: (1) mathematics achievement, (2) language achievement, (3) general measures of academic outcomes (e.g. drop out rates, rates of grade retention) and (4) measures of psychomotor and affective outcomes of schooling (e.g. self-concept, attitudes towards mathematics, attitudes to teachers, attitudes to peers). Therefore, three dummy variables were entered and mathematics achievement was treated as the baseline variable. It is important to note that since few studies used outcomes of schooling associated with either the new learning or with the achievement of psychomotor aims, we treated them as belonging to the same group of studies as those that measured affective outcomes of schooling. In this way, the studies in the fourth group did not consider achievement of any cognitive aim as criterion for measuring effectiveness.

Level of education

It was possible to classify the studies into groups according to the age of the students participating in the study. Specifically, one dummy variable was entered into the model and the studies could be classified into those conducted in primary schools and those conducted in secondary schools.

Country in which the study was conducted

The studies were classified into the following five groups according to the country where the study took place: the USA, the Netherlands, the United Kingdom, Asian countries and other countries. Studies conducted in the USA were treated as the baseline group.

The study design employed

We also examined the study design employed. It was possible to classify the studies into the following four groups: cross-sectional studies, longitudinal studies,

experimental studies and outlier studies. The cross-sectional studies were treated as the baseline group. Our initial intention was to classify the longitudinal studies into two groups: studies that lasted up to two years and longitudinal studies that lasted more than two years. However, the number of longitudinal studies that lasted more than two years was very small. For this reason, all longitudinal studies were treated as one group.

Statistical techniques employed

The statistical techniques employed to analyse data were taken into account. Specifically, we examined whether multilevel or unilevel analysis was used to investigate the relation of each factor with student achievement. For this reason, a relevant dummy variable was entered into the model in order to see whether the statistical technique employed could predict variation on the effect sizes.

In order to assess the impact of the above explanatory variables, we extend the between replication model into the model shown below:

$$\delta_{rs} = \delta_0 + \gamma_1 \text{outcome} - \text{language}_{rs} + \gamma_2 \text{general} - \text{measures}_{rs} + \gamma_3 \text{outcome} - \text{affective}_{rs} + \gamma_4 \text{secondary}_{rs} + \gamma_5 \text{netherlands}_{rs} + \gamma_6 \text{uk}_{rs} + \gamma_7 \text{asiancountries}_{rs} + \gamma_8 \text{othercountries}_{rs} + \gamma_9 \text{longitudinal}_{rs} + \gamma_{10} \text{experimental}_{rs} + \gamma_{11} \text{outlier}_{rs} + \gamma_{12} \text{multilevel}_{rs} + v_s$$

where:

outcome-language	0 = else, and 1 = language achievement
general-measures	0 = else, and 1 = general measures of academic outcomes
outcome-affective	0 =else, and 1 = measure of non-cognitive outcomes of schooling (i.e. achievement of affective or psychomotor aims)
secondary	0 = primary, and 1= secondary
Netherlands	0= else, and 1= conducted in the Netherlands
UK	0= else, and 1= conducted in the United Kingdom,
Asiancountries	0= else, and 1= conducted in Asian countries
othercountries	0= else, and 1= conducted in countries other than the United States, the Netherlands, the UK or any of the Asian countries
longitudinal	0 = else, and 1= longitudinal study
experimental	0 = else, and 1= experimental study
outlier	0 = else, and 1= outlier study.
multilevel	0= unilevel analysis, and 1= multilevel analysis.