Situational effects of the school factors included in the dynamic model of educational effectiveness

Bert Creemers and Leonidas Kyriakides

b.p.m.creemers@rug.nl

We present results of a longitudinal study in which 50 schools, 113 classes and 2,542 Cypriot primary students participated. We tested the validity of the dynamic model of educational effectiveness and especially its assumption that the impact of school factors depends on the current situation of the school and on the type of problems/difficulties the school is facing. Reference is made to the methods used to test this assumption of the dynamic model by measuring school effectiveness in mathematics, Greek language, and religious education over two consecutive school years. The main findings are as follows. School factors were found to have situational effects. Specifically, the development of a school policy for teaching and the school evaluation of policy for teaching were found to have stronger effects in schools where the quality of teaching at classroom level was low. Moreover, time stability in the effectiveness status of schools was identified and thereby changes in the functioning of schools were found not to have a significant impact on changes in the effectiveness status of schools. Implications of the findings for the development of the dynamic model and suggestions for further research are presented.

Introduction

Teaching and learning are dynamic processes that are constantly adapting to changing needs and opportunities. Slater and Teddlie (1992) argue that effective schools/teachers are expected to change in order to remain effective as their contexts change; they must adapt their schooling to the changing context. Therefore, effective schooling should be treated as a dynamic, ongoing process. This idea is consistent with the contingency theory (Donaldson, 2001; Mintzberg, 1979) and with the main assumptions upon which the dynamic model of educational effectiveness is based (Creemers & Kyriakides, 2008). One of the major assumptions of the model which reveals its essential difference from other integrated models of educational effectiveness (e.g. Creemers, 1994; Scheerens, 1992) is that schools and educational systems that are able to identify their weaknesses and develop a policy on aspects associated with teaching and the learning environment of the school are also able to improve the functioning of classroom-level factors and their effectiveness status (Creemers & Kyriakides, 2008). Only changes in those factors for which schools face significant problems are expected to be associated with the improvement of school effectiveness. For example, at those schools where teacher and/or student absenteeism rarely occur, change in their policy on

absenteeism is not necessary since these schools do not face a problem with this factor. On the other hand in schools, where student absenteeism occurs very often, a change in the school policy is likely to have an effect on reducing absenteeism and thereby increasing the quantity of teaching offered to students, which is then related to student outcomes. This implies that, depending on the situation of the school, changes in the school-factors included in the dynamic model may have an impact on student achievement. It is however important to acknowledge that, although some supportive material for the validity of the dynamic model have been provided, this assumption of the model needs further investigation. In this context, our report here refers to the results of a study which attempted to find out whether the impact of school factors depends on the situation that occurs in the school, implying that their effects are situational.

The dynamic model of educational effectiveness: an overview
The essential characteristics of the dynamic model
The dynamic model takes into account the fact that effectiveness studies conducted in several countries reveal that the influences on student achievement are multilevel (Teddlie & Reynolds, 2000). Therefore, the dynamic model is multilevel in nature and refers to four different levels shown in Figure 1. The teaching and learning situation is emphasised 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-level factors. It is expected that these factors influence the teaching-learning situation by developing and evaluating the school policy on teaching and the policy on creating a learning environment at the school. The final level refers to the influence of the educational system in a more formal way, especially through developing and evaluating the educational policy at the national/regional level. Also taken into account is that the teaching and learning situation is influenced by the wider educational context in which students, teachers, and schools are expected to operate. Factors, such as the values for learning of the society and the importance they attach to education, play an important role both in shaping teacher and student expectations.

The interrelations between the components of the model are also illustrated. In this way, the model assumes that factors at the school and context level have both direct and indirect effects on student achievement since they are able not only to influence student achievement directly but also to influence the teaching and learning situations. Therefore, teaching is emphasised and the description of the classroom level refers mainly to the behaviour of the teacher in the classroom and especially to his/her contribution in promoting learning at the classroom level. Moreover, defining factors at the classroom level is seen as a prerequisite for defining the school and the system level. Finally, the dynamic model is based on the assumption that, although there are different effectiveness factors, each factor can be defined and measured using five dimensions: f*requency, focus, stage, quality*, and
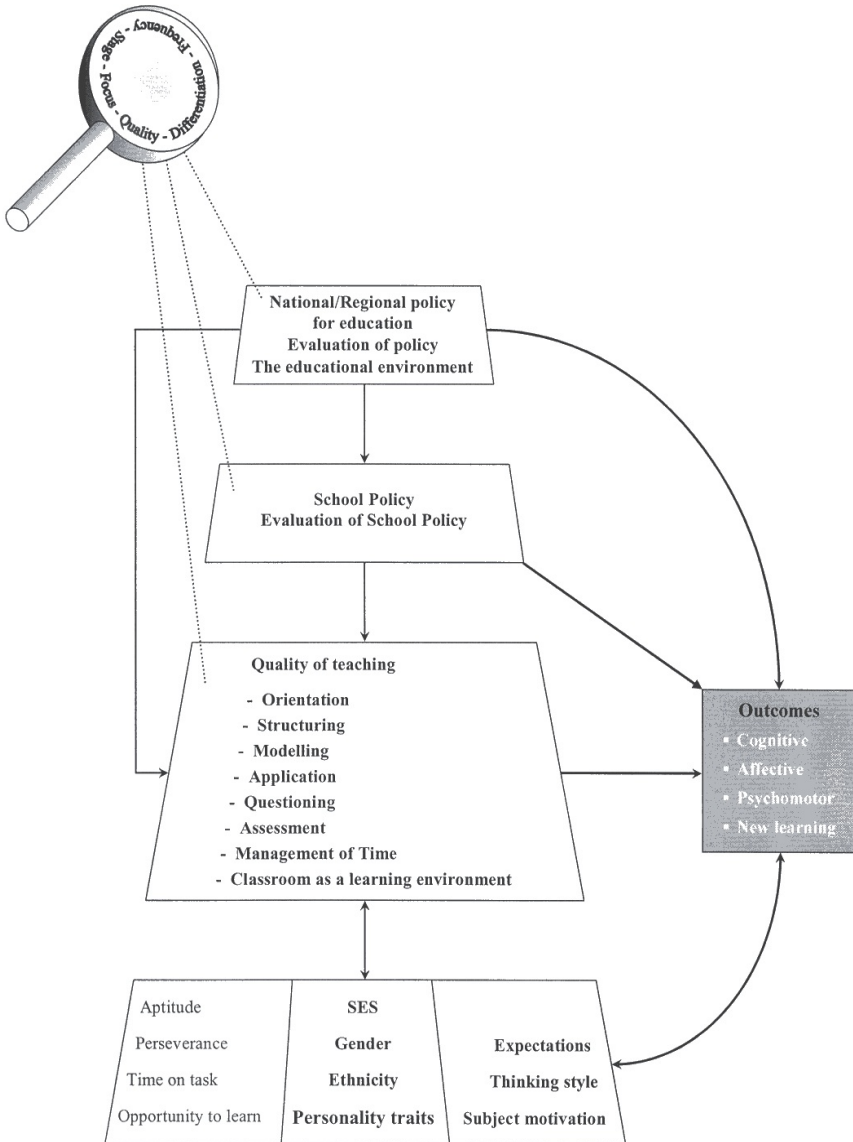
Figure 1 The dynamic model of educational effectiveness

differentiation. *Frequency* is a quantitative way to measure the functioning of each effectiveness factor. The other four dimensions examine qualitative characteristics of the functioning of the factors and help us describe the complex nature of effective teaching. A brief description of these four dimensions is given below. (For further information on the importance of these five dimensions and empirical support to using these dimensions see Kyriakides & Creemers, 2008.) Specifically, two aspects of the *focus* dimension are taken into account. The first one refers to the specificity of the activities associated with the functioning of the factor whereas the second one with the number of purposes for which an activity takes place. The *stage* at which tasks associated with a factor take place is also examined. It is expected that the factors need to take place over a long period of time to ensure that they have a continuous direct or indirect effect on student learning. The *quality* refers to properties of the specific factor itself, as these are discussed in the literature. Finally, *differentiation* refers to the extent to which activities associated with a factor are implemented in the same way for all the subjects involved with it (e.g. all the students, teachers, schools). It is expected that adaptation to specific needs of each subject or group of subjects will increase the successful implementation of a factor and ultimately maximize its effect on student learning outcomes.

Classroom factors of the dynamic model
Based on the main findings of teacher effectiveness research (e.g. Brophy & Good, 1986; Fraser *et al.*, 1987; Kyriakides, 2005; Muijs & Reynolds, 2001; Opdenakker & Van Damme, 2000; Rosenshine & Stevens, 1986; Seidel & Shavelson, 2007), the dynamic model refers to factors which describe teachers' instructional role and are associated with student outcomes. These factors refer to observable instructional behaviour of teachers in the classroom rather than on factors that may explain such behaviour (e.g. teacher beliefs and knowledge and interpersonal competences). The eight factors included in the model are as follows: orientation, structuring, questioning, teaching-modelling, applications, management of time, teacher role in making classroom a learning environment, and classroom assessment. These eight factors do not refer only to one approach of teaching such as structured or direct teaching (Joyce *et al.*, 2000) or to approaches associated with constructivism (Schoenfeld, 1998). An integrated approach in defining quality of teaching is adopted. For example, the dynamic model does not refer only to skills associated with direct teaching and mastery learning such as structuring and questioning but also to orientation and teaching modelling which are in line with theories of teaching associated with constructivism (Brekelmans, Sleegers & Fraser, 2000). Moreover, the collaboration technique is included under the overarching factor contribution of teacher to the establishment of classroom learning environment.

School factors of the dynamic model
The definition of the school level is based on the assumption that factors at

the school level are expected to have not only direct effects on student achievement but also indirect effects. School factors are expected to influence classroom-level factors, especially the teaching practice. This assumption is based on the fact that effectiveness studies show that the classroom level is more significant than the school and the system level (e.g. Kyriakides *et al.*, 2000; Teddlie & Reynolds, 2000) and that defining factors at the classroom level is seen as a prerequisite for defining the school and the system level (Creemers, 1994). Therefore, the dynamic model refers to factors at the school level which are related to the same key concepts of quantity of teaching, provision of learning opportunities, and quality of teaching which were used to define classroom-level factors. Thus, emphasis is given to the school policy for teaching which is expected to have an impact on these three concepts of teaching. The other aspect of school policy, which is taken into account by the model, is concerned with the school policy for creating a school learning environment. This can be seen as an attempt to define in a more specific way the climate of the school. In the literature, the school climate is defined very broadly as the total environment of the school (e.g. Stringfield 1994; Webster & Fisher, 2003). This makes it difficult to study specific factors of the school climate and examine their impact on student achievement (Creemers & Reezigt, 1999). Thus, the dynamic model refers to the school learning environment and not to the whole school climate as a broader concept. The school learning environment is an element of school climate that is seen as the most important predictor 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 order to achieve educational goals more effectively in conditions of uncertainty (Kyriakides *et al.*, 2002; Teddlie & Reynolds, 2000).

Guidelines are seen as one of the main indications of school policy and this is reflected in the way each school-level factor is defined. However, in using the term guidelines we refer to a range of documents, such as staff meeting minutes, announcements, and action plans, which make the policy of the school more concrete to teachers and other stakeholders. It should also be acknowledged that this factor does not imply that each school should simply develop formal documents to install the policy. The factors concerned with the school policy mainly refer to the actions taken by the school to help teachers and other stakeholders have a clear understanding of what is expected from them to do. Support offered to teachers and other stakeholders to implement the school policy is also an aspect of these two overarching factors.

Based on the assumption that the essence of a successful organization in the modern world is the search for improvement (Barber, 1986), we also examine the processes and the activities which take place in the school in order to improve the teaching practice and its learning environment. For this reason, the processes used to evaluate the school policy for teaching and the learning environment of the school are investigated. Thus, the following four

overarching factors at the school level are included in the model:
1. School policy for teaching and actions taken for improving teaching practice;
2. evaluation of school policy for teaching and of actions taken to improve teaching;
3. policy for creating a school learning environment and actions taken for improving the school learning environment; and
4. evaluation of the school learning environment.

Leadership is not considered as a school-level factor. This can be attributed to the fact that current meta-analyses of studies investigating the possible impact of the principal's leadership on student achievement (e.g. Kyriakides *et al.* (in press); Scheerens *et al.*, 2005) confirm earlier research findings on the limitations of the direct effects approach to linking leadership with student achievement (Witziers, Bosker & Kruger, 2003). Similar results are obtained from the few studies which were conducted in order to measure indirect effects of leadership on student achievement (Leithwood & Jantzi, 2006). Therefore, the model is not 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. This reveals one of the major characteristics of the model which is not focused on individuals as such but on the effects of the actions that take place at classroom/school/context levels. This holds for the students, teachers, principals and policy makers. Our decision is also consistent with the way classroom level factors are measured since, instead of measuring the teaching style of the teacher, we are focused on the actual behaviour of the teacher in the classroom (see Creemers & Kyriakides, 2006). Similarly, instead of measuring the leadership style of a principle we look at the impact of the end result of leadership (e.g. the development of school policy on teaching or the evaluation of school policy).

Finally, the dynamic model assumes that the impact of the school-level factors and the impact of the context-level factors have to be defined and measured in a different way from the impact of classroom-level factors. Policy on teaching and actions taken to improve teaching practice must be measured over time and in relation to the weaknesses that occur in a school/educational system. The assumption is that schools and educational systems which are able to identify their weaknesses and develop a policy on aspects associated with teaching and the learning environment of the school are also able to improve the functioning of classroom-level factors and their effectiveness status. Only changes in those factors for which schools face significant problems are expected to be associated with the improvement of school effectiveness. This characteristic of the proposed dynamic model reveals an essential difference in the nature of this model to all the current models of educational effectiveness.

Testing the validity of the model: findings and new research questions
Some supportive material for the validity of the proposed dynamic model has been provided. Specifically, a longitudinal study was designed to test the main assumptions of the model. Using data emerged during the first phase of the study it was possible to provide evidence supporting the validity of the model at the classroom level. It has also been shown that the proposed measurement framework can be used to describe each classroom level factor. The added value of using these five dimensions of the classroom-level factors to explain variation on student achievement has also been identified (Kyriakides & Creemers, 2008). Based on the data emerged during the second phase of the study, the importance of the four overarching school-level factors of the dynamic model has also been confirmed (Kyriakides & Creemers, 2007). In addition, a quantitative synthesis of the results of studies exploring the impact of school factors on student achievement has also been conducted. This meta-analysis provided some support to the validity of the model at the school level (Creemers & Kyriakides, 2008) but also revealed that there is no study that has investigated the extent to which the impact of school factors depends on the current situation of the school and especially on the type of problems/difficulties that the school is facing. Since this assumption of the dynamic model reveals one of its main differences from any other effectiveness model, we present the results of a longitudinal study investigating the validity of this essential characteristic of the dynamic model which reveals its dynamic nature.

Research aims
Therefore, in the study reported here we attempt to identify the conditions under which the two overarching school-level factors of the dynamic model — (a) school policy for teaching and actions taken to improve teaching, and (b) policy for the school learning environment (SLE) and actions taken to improve the SLE — explain changes in the effectiveness status of schools from one year to another. Specifically, we examine whether changes in the effectiveness status of schools can be attributed to changes in the functioning of school factors. We also try to find out whether the effect of school factors depends on the situation of the schools and the problems they are facing. These two assumptions are considered to be major elements of the dynamic model of educational effectiveness (Creemers & Kyriakides, 2008).

Methods
Participants
Stratified sampling was used to select 50 out of 191 Cypriot primary schools. All the Grade 5 students ($n = 2,542$) from each class ($n = 113$) of the school sample were chosen. The chi-square test did not reveal any significant difference between the research sample and the population in terms of students' sex ($\chi^2 = 0.84$, df = 1, $p = 0.42$). Moreover, the t test did not reveal any significant difference between the research sample and the population in terms of the size of class ($t = 1.21$, df = 107, $p = 0.22$). Although this study refers to

additional variables such as the socio-economic status of students and their achievement levels in different outcomes of schooling, there are no data about these characteristics of the Greek Cypriot students of year 5. Therefore, it was not possible to examine whether the sample was nationally representative in terms of any characteristic other than students' sex and the size of class. However, it can be claimed that a nationally representative sample of Cypriot year 5 students in terms of these two characteristics was drawn.

Variables
Output measures
Data on achievement in Mathematics, Greek Language and Religious Education were collected by using external forms of assessment. Written tests were administered to our student sample when they were at the beginning of Grade 5 (i.e. October 2004), at the end of Grade 5 (i.e. May 2005), and at the end of Grade 6 (i.e. May 2006). The construction of the tests was subject to controls for reliability and validity. Specifically, for each period of collecting achievement data, the Extended Logistic Model of Rasch (Andrich, 1988) was used to analyse the emerging data in each subject separately and four scales were created, which refer to student knowledge in mathematics, Greek language, and religious education, and also to student attitudes towards religious education. Analysis of the data revealed that each scale had relatively satisfactory psychometric properties (see Creemers & Kyriakides, 2008). Thus, for each student four different scores for his/her achievement at the beginning of year 5, the end of year 5, and the end of year 6, respectively, were generated, by calculating the relevant Rasch person estimate in each scale.

Student background factors
Information was collected on two student background factors: sex (0 = boys, 1 = girls), and socio-economic status (SES). Five SES variables were available: father's and mother's education level (i.e. graduate of a primary school, graduate of secondary school or graduate of a college/university), the social status of father's job, the social status of mother's job and the economic situation of the family. Following the classification of occupations used by the Ministry of Finance, it was possible to classify parents' occupations into three groups which have relatively similar sizes: occupations held by working class (33%), occupations held by middle class (37%) and occupations held by upper-middle class (30%). Representative parental occupations for the working class are: farmer, truck driver, machine operator in a factory; for the middle class are: police officer, teacher, bank officer; and for the upper-middle class are: doctor, lawyer, business executive. Relevant information for each child was taken from the school records. Then standardized values of the above five variables were calculated, resulting in the SES indicator.

Quality of teaching
The explanatory variables of the study which refer to the eight factors of the dynamic model dealing with teacher behaviour in the classroom (Creemers &

Kyriakides, 2006) were measured by both independent observers and students. Taking into account the way the five dimensions of each effectiveness factor are defined, one high-inference and two low-inference observation instruments were developed. These observation instruments generate data for all eight factors and their dimensions. Observations were carried out by six members of the research team who attended a series of seminars on how to use the three observation instruments. During the school year, the external observers visited each class nine times and observed three lessons per subject. For each scale of the three observation instruments, the alpha reliability coefficient was higher than 0.83, and the inter-rater reliability coefficient ρ^2 was higher than 0.81.

The eight factors and their dimensions were also measured by administering a questionnaire to students. Specifically, students were asked to indicate the extent to which their teacher behaves in a certain way in their classroom and a Likert scale was used to collect data. For example, an item concerned with the stage dimension of the structuring factor was asking students to indicate whether at the beginning of the lesson the teacher explains how the new lesson is related to previous ones, whereas another item was asking whether at the end of each lesson they spend some time in reviewing the main ideas of the lesson. Similarly, the following item was used to measure the differentiation dimension of the application factor: "the teacher of Mathematics assigns to some pupils different exercises than to the rest of the pupils". A Generalisability Study (Cronbach, Gleser, Nanda & Rajaratnam, 1972; Shavelson, Webb & Rowley, 1989) on the use of students' ratings was conducted. It was found that the data from almost all the questionnaire items could be used for measuring the quality of teaching of each teacher, in each subject separately (see Creemers & Kyriakides, 2008).

For each subject, separate CFA analyses for each effectiveness factor were conducted in order to identify the extent to which data emerged from different methods can be used to measure each factor in relation to the five dimensions of the dynamic model. The main results which emerged from using CFA approaches to analyse the multitrait multimethod matrix (MTMM) concerned with each classroom level factor of the dynamic model in relation to each subject are briefly presented below. (For more information see Creemers & Kyriakides, 2008.) Specifically, support to the construct validity of the five measurement dimensions of most effectiveness factors was provided. The few exceptions identified reveal the difficulty of defining the quality dimension. Moreover, the results of this study seem to reveal that the classroom as a learning environment cannot be treated as a single factor but as two interrelated factors in the learning environment concerning relations among students and relations between teacher and his/her students. Furthermore, the comparison of CFA models used to test each factor confirmed convergent and discriminant validity for the five dimensions. Convergent validity for most measures was demonstrated by the relatively high (i.e. higher than .60) standardized trait loadings, in comparison to the relatively lower (i.e. lower than .40) standardized method loadings. These findings support the use of multi-

method techniques for increasing measurement validity, construct validity, and, thus, stronger support for the validity of subsequent results.

School level factors of the dynamic model
The explanatory variables which refer to the four school-level factors of the dynamic model were measured by asking all the teachers of the school sample to complete a questionnaire. The questionnaire was designed in such a way that information about the five dimensions of the four factors could be collected. Of the 364 teachers approached 313 responded, a response rate of 86%. The chi-square test did not reveal any significant difference between the distribution of the teacher sample which indicates at which school each teacher works and the relevant distribution of the whole population of the teachers of the 50 schools of our sample (χ^2 = 57.12, df = 49, p < .38). Therefore, the sample was representative of the whole population in terms of how the teachers are distributed in each of these 50 schools.

Since it is expected that teachers within a school view the policy of their school and the evaluation mechanisms of their school similarly but differently from teachers in other schools, reliability was computed for each of the scales of the questionnaire by calculating multilevel λ (Snijders & Bosker, 1999) and Cronbach alpha for data aggregated to the school level. The value of Cronbach alpha represents consistency across items whereas multilevel λ represents consistency across groups of teachers. Since it was found that the reliability coefficients were high (around .80), it was decided to treat teacher responses to the questionnaire as indicators of the dimensions of each of the factors of their school. In order to test the construct validity of the school level factors, CFA approaches were used. It was found that the use of five measurement dimensions is appropriate for measuring school policy on teaching and the evaluation of school policy on teaching. However, in the case of the overarching factor concerned with the school as a learning environment, the items of the factor of the dynamic model concerned with the partnership policy were found to belong to two separate factors concerned with relation of the school with (a) parents and the wider community, and (b) the employers and supporting mechanisms offered by the ministry of education (e.g. inspectorate, pedagogical institute, advisory bodies). Moreover, we could not generate valid data on the factor concerned with the values towards learning. Finally, the measurement framework of the factors concerned with the evaluation of the school learning environment was supported but the focus dimension was not found to be related to any of the other four dimensions. For this reason, data associated with this dimension were not taken into account.

Results
Searching for changes in the effectiveness status of schools
Having established the reliability and the validity of the data, for each outcome, separate multilevel analysis for each outcome variable was conducted in order to examine the extent to which each dimension of each school and classroom factor is associated to student achievement. Based on the results

of each final multilevel model used to measure the effect of school and class-room factors on student achievement, the difference between the expected and the actual score for each school was plotted. The standard error of estimate for each school was also taken into account and is represented by the length of a vertical line. This line can be conceptualized as the range within which we are 95% confident that the "true" estimate of the school's residual lies (Goldstein, 2003). Thus, where this vertical line does not cross the horizontal zero line and is also situated below the zero line the school it represents is considered as one of the least effective schools of our sample. On the other hand, where this line does not cross the horizontal zero line and is situated above the zero line, the school it represents is characterized as one of the most effective schools. All the other schools which cannot be considered as either most or least effective schools are characterized as "typical".

Based on the results of the classifications concerning the effectiveness status of our school sample in mathematics, the extent to which each school can be considered equally effective during two consecutive school years was identified. This classification revealed that 41 out of 50 schools can be considered equally effective in mathematics across two consecutive school years. This implies that, to some extent, there is time stability in the short-term school effects in mathematics. However, we also identified four schools which managed to improve their effectiveness status whereas the effectiveness status of five other schools declined. Similar observations emerged from analysing changes in the effectiveness status of the schools during these two consecutive years in relation to the other three dependent variables (see Table 1). Since the numbers of schools where changes were observed are very small, the statistical power for conducting any further analysis to identify variables associated with this change is insufficient.

Table 1 Number of schools where different types of changes in their effectiveness status were observed during two consecutive school years per outcome measure

Groups of schools	Mathematics	Language	RE (Cognitive)	RE (Affective)
A. Stability				
Remain Typical*	35	32	37	38
Remain Least Effective	2	5	4	2
Remain Most Effective	4	4	3	2
B. Improvement				
From Least Effective to Typical	3	3	1	2
From Typical to Most Effective	1	2	2	2
C. Declining				
From Most Effective to Typical	2	3	2	2
From Typical to Least Effective	3	1	1	2

* This category refers to schools that were neither among the most nor among the least effective schools

Situational effects of school factors

In order to examine whether the effect size of each school factor depends on the situation and on the problems the schools are facing the following procedure was used. This procedure is based on the assumption of the dynamic model that school factors are expected to influence quality of teaching. Thus, based on standardized scores of the measures of quality of teaching that emerged from the collection of data during the first year of the study, the 50 schools were classified into three groups. The schools of the first group had overall measures of quality of teaching lower than –1, the measures of the second group were between –1 and 1, and the third group had measures higher than 1. For each group, we conducted separate multilevel analyses investigating the effect of school factors upon student achievement in each outcome, and the estimated effect sizes of each factor were compared. Tables 2–5 present the final multilevel model emerged from analysing the effect of explanatory variables upon each of the four outcomes per group of schools. The following observations arise from these tables.

In mathematics and Greek language, the estimates of the effects of most school factors emerged by the first group were found to be larger than those emerged by the other two groups. Specifically, in the case of Greek language almost all school factors had higher effect sizes for schools with low quality of teaching. This difference is more apparent in the case of policy for teaching (frequency and quality dimensions) and in evaluation of policy for teaching (frequency and stage). Moreover, in this group of schools, the final model is able to explain more variance than the analyses of the two other types of schools. In the case of mathematics, similar observations can be drawn since 11 out of 15 factors had higher effect sizes and also more than one dimension of the two overarching school factors associated with teaching had stronger effect sizes. In the case of RE, the effect sizes of school factors upon achievement were smaller and thereby differences in their effect sizes for the three types of schools are not so clear.

Discussion

The findings in this study appear to reveal that the school factors included in the dynamic model have situational effects. Implications of these findings for the improvement of practice can be drawn. Given that the school factors have situational effects, one could claim that there is no factor that can be considered as more important than the other. Schools should not only develop effective policies of teaching but also attempt to improve their learning environment. However, in order to develop action plans, which may have stronger impact on student outcomes, school stakeholders should first of all identify those factors and dimensions at classroom and school level that are not functioning as well as other factors and design improvement strategies that will address these specific factors. Therefore, top-down initiatives for introducing school improvement initiatives, irrespective of the situation of the school, are very likely to have little impact on student outcomes. This is a lesson which is also drawn from various studies attempting to evaluate national reform

Table 2 Parameter estimates and (standard errors) for the analyses of Greek language achievement in schools with low/typical/high quality of teaching*

Factors	Low quality teaching		Average quality teaching		High quality teaching	
	Model 0	Final mod.	Model 0	Final mod.	Model 0	Final mod.
Fixed part (Intercept)	−0.61(.08)	−0.22 (.08)	−0.41 (.08)	−0.27 (.08)	0.32 (.09)	0.19 (.08)
Student Level						
Prior knowledge		0.29 (.05)		0.21 (.05)		0.21 (.05)
Sex (boys = 0, girls = 1)		0.15 (.08)		0.12 (.06)		N.S.S.
SES		0.17 (.06)		0.15 (.05)		0.14 (.05)
Classroom Level: Context						
Average prior knowledge		0.10 (.05)		0.10 (.05)		0.10 (.05)
Average SES		N.S.S.**		N.S.S.*		N.S.S.*
Percentage of girls		N.S.S.		N.S.S.		N.S.S.
School Level						
Context						
Average SES		N.S.S.		N.S.S.		N.S.S.
Average prior knowledge		0.08 (.04)		N.S.S.		N.S.S.
Percentage of girls		N.S.S.		N.S.S.		N.S.S.
Frequency						
Policy for Teaching		0.15 (.05)		0.10 (.05)		0.08 (.05)
Evaluation policy for teaching		0.20 (.06)		0.12 (.05)		0.13 (.05)
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		N.S.S.		N.S.S.		N.S.S.
Relations with parents		N.S.S.		N.S.S.		N.S.S.
Relations with the centre		N.S.S.		N.S.S.		N.S.S.
Provision of resources		0.12 (.03)		0.10 (.03)		0.07 (.03)
Evaluation of the SLE		N.S.S.		N.S.S.		N.S.S.
Stage						
Policy for teaching		0.13 (.03)		0.07 (.02)		0.08 (.03)
Evaluation policy for teaching		0.15 (.03)		0.05 (.02)		0.06 (.02)
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		N.S.S.		N.S.S.		N.S.S.
Partnership		N.S.S.		N.S.S.		N.S.S.
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		0.06 (.02)		N.S.S.		N.S.S.
Focus						
Policy for Teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		0.06 (.02)		N.S.S.		N.S.S.
School learning environment		N.S.S.		N.S.S.		N.S.S.
Quality						
Policy for Teaching		0.22 (.05)		0.12 (.05)		0.13 (.06)
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		N.S.S.		N.S.S.		N.S.S.
Partnership		0.18 (.04)		0.08 (.04)		0.12 (.04)
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		0.12 (.02)		0.10 (.03)		0.09 (.03)

Table 2 Continued

Factors	Low quality teaching		Average quality teaching		High quality teaching	
	Model 0	Final mod.	Model 0	Final mod.	Model 0	Final mod.
Differentiation						
Policy for teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
School learning environment		N.S.S.		0.06 (.03)		0.07 (.03)
Evaluation of SLE		0.14 (.03)		0.12 (.03)		N.S.S.
Variance components						
School	10.0%	3.2%	10.7%	4.2%	10.7%	4.4%
Class	9.7%	8.3%	8.0%	7.4%	8.3%	8.0%
Student	80.3%	21.3%	81.3%	24.3%	81.0%	26.7%
Explained		67.2%		64.1%		60.9%
Significance test						
χ^2	815.6	300.2	756.8	330.4	801.6	397.2
Reduction		515.4		426.4		404.4
Degrees of freedom		16		14		12
p value		.001		.001		.001

* **Low:** Schools with score of quality of teaching lower than –1; **Average:** schools with score between –1 and 1; **High:** schools with scores higher than 1

** N.S.S. = No statistically significant effect at .05 level.

policies (Kyriakides *et al.*, 2006; Teddlie & Reynolds, 2000). Moreover, these findings reveal the importance of establishing school evaluation mechanisms in order to identify the factors that are likely to matter more for each school. This implies that support should be provided to the schools in order to establish such mechanisms that will guide their actions for improvement.

Implications of findings for the development of educational effectiveness research can also be drawn. First, an important finding of this study is concerned with the fact that school factors were found to have larger effects on student achievement in schools with low quality of teaching. These findings imply that in schools where quality of teaching is rather low, schools that give emphasis to the development of the policy on teaching and its evaluation are more likely to improve the quality of teaching and in this way have a strong influence on student outcomes. This implies that school factors should be defined in a different way from teacher factors in order to capture both the direct and indirect effects of these factors on student outcomes. Teacher factors are concerned with the behaviour of teachers in the classroom and their effects are expected to be the same in each class irrespective of the level of its students. On the other hand, school factors may matter more in schools with low quality of teaching, since in these schools improvement of quality of teaching is more important. This argument is also in line with the need to develop formative evaluation mechanisms at the school level which will help school stakeholders to identify priorities for improvement. For example, in

Table 3 Parameter estimates and (standard errors) for the analyses of mathematics achievement in schools with low/typical/high quality of teaching*

Factors	Low quality teaching		Average quality teaching		High quality teaching	
	Model 0	Final mod.	Model 0	Final mod.	Model 0	Final mod.
Fixed part (Intercept)	–0.15 (.05)	0.08 (.05)	0.12 (.04)	0.09 (.03)	0.22 (.03)	0.11 (.03)
Student Level						
Prior knowledge		0.32 (.10)		0.25 (.10)		0.20 (.10)
Sex (boys = 0, girls = 1)		–0.12 (.06)		–0.10 (.05)		N.S.S.
SES		0.23 (.10)		0.20 (.10)		0.10 (.05)
Classroom Level: Context						
Average prior knowledge		0.18 (.08)		0.12 (.03)		0.10 (.03)
Average SES		0.11 (.05)		0.09 (.04)		0.09 (.03)
Percentage of girls		–0.05 (.02)		N.S.S.		N.S.S.
School Level						
Context						
Average SES		N.S.S.		N.S.S.		N.S.S.
Average prior knowledge		0.11 (.05)		N.S.S.		N.S.S.
Percentage of girls		N.S.S.		N.S.S.		N.S.S.
Frequency						
Policy for Teaching		0.21 (.04)		0.11 (.04)		N.S.S.
Evaluation policy for teaching		0.19 (.04)		0.10 (.03)		0.11 (.03)
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		0.14 (.03)		N.S.S.		N.S.S.
Relations with parents		N.S.S.		N.S.S.		N.S.S.
Relations with the centre		N.S.S.		N.S.S.		N.S.S.
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		N.S.S.		N.S.S.		N.S.S.
Stage						
Policy for teaching		0.08 (.02)		0.08 (.02)		0.08 (.02)
Evaluation policy for teaching		0.15 (.04)		0.05 (.02)		0.05 (.02)
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		0.16 (.04)		0.12 (.04)		0.12 (.04)
Partnership		0.07 (.03)		0.07 (.03)		0.07 (.03)
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		0.12 (.03)		0.06 (.03)		0.06 (.03)
Focus						
Policy for Teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
School learning environment		N.S.S.		N.S.S.		N.S.S.
Quality						
Policy for Teaching		0.26 (.04)		0.16 (.04)		0.11 (.04)
Evaluation policy for teaching		0.15 (.03)		0.12 (.03)		0.10 (.03)
Student behaviour out of class		0.09 (.02)		0.06 (.02)		0.07 (.02)
Collaboration among teachers		0.08 (.02)		0.08 (.03)		0.09 (.03)
Partnership		0.18 (.04)		0.10 (.04)		0.11 (.03)
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		N.S.S.		N.S.S.		N.S.S.

Table 3   Continued

Factors	Low quality teaching		Average quality teaching		High quality teaching	
	Model 0	Final mod.	Model 0	Final mod.	Model 0	Final mod.
Differentiation						
Policy for teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
School learning environment		0.12 (.03)		N.S.S.		N.S.S.
Evaluation of SLE		0.18 (.02)		0.08 (.02)		0.12 (.04)
Variance components						
School	11.2%	4.0%	11.8%	4.3%	11.0%	4.0%
Class	8.8%	8.0%	9.0%	8.4%	8.8%	8.0%
Student	80.0%	23.2%	79.2%	27.0%	80.2%	28.2%
Explained		64.8%		60.3%		59.8%
Significance test						
χ^2	844.9	304.5	712.6	230.3	754.8	363.8
Reduction		540.4		482.3		391.0
Degrees of freedom		22		18		16
p value		.001		.001		.001

   *   **Low:** Schools with score of quality of teaching lower than −1; **Average:** schools with score between −1 and 1; **High:** schools with scores higher than 1

  ** N.S.S. = No statistically significant effect at .05 level.

schools where quality of teaching is appropriate, the development of a school policy for teaching may not have any strong effect on student outcomes but the establishment of a partnership policy may have stronger effects, especially if the schools do not manage to establish good relations with parents and the wider school community. However, to test this assumption further, we need longitudinal studies which will collect data on both the functioning of teacher and school factors and search for all types of situational effects of school factors and not only for differential effects in relation to the quality of teaching at classroom level. In this study, the data that were available could not help us address these causal relations especially since measures of school factors were only taken at the second phase of the study. The use of longitudinal designs where not only measures of student outcomes but also measures of the functioning of factors are repeatedly taken could test further this assumption and show whether changes in the functioning of school factors have any impact not only on changing the effectiveness status of schools but also the quality of teaching practice.

    Second, the dynamic model is based on the assumption that the effectiveness status of schools does not remain the same and that the functioning of school factors can explain changes in the effectiveness status of schools. Such an approach can be found in other disciplines such as developmental psychology, which attempts to identify factors that explain cognitive development of students. Conducting studies, which search for factors that explain changes in the effectiveness status of schools, may help us develop the dynamic model

Table 4 Parameter estimates and (standard errors) for the analyses of RE (Cognitive) achievement in schools with low/typical/high quality of teaching*

Factors	Low quality teaching		Average quality teaching		High quality teaching	
	Model 0	Final mod.	Model 0	Final mod.	Model 0	Final mod.
Fixed part (Intercept)	−0.85 (.22)	−0.28 (.08)	−0.15 (.12)	0.04 (.01)	0.15 (.04)	0.08 (.02)
Student Level						
Prior knowledge		0.17 (.05)		0.12 (.04)		0.15 (.05)
Sex (boys = 0, girls = 1)		0.08 (.03)		0.06 (.02)		0.08 (.03)
SES		0.11 (.04)		0.08 (.04)		0.11 (.04)
Classroom Level: Context						
Average prior knowledge		0.09 (.02)		0.05 (.02)		0.08 (.02)
Average SES		N.S.S.		N.S.S.		N.S.S.
Percentage of girls		N.S.S.		N.S.S.		N.S.S.
School Level						
Context						
Average SES		N.S.S.		N.S.S.		N.S.S.
Average prior knowledge		0.04 (.02)		0.04 (.02)		0.04 (.02)
Percentage of girls		N.S.S.		N.S.S.		N.S.S.
Frequency						
Policy for Teaching		0.08 (.03)		0.08 (.03)		0.09 (.03)
Evaluation policy for teaching		0.09 (.03)		0.06 (.03)		0.07 (.03)
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		N.S.S.		N.S.S.		N.S.S.
Relations with parents		0.14 (.05)		0.10 (.04)		0.12 (.05)
Relations with the centre		N.S.S.		N.S.S.		N.S.S.
Provision of resources		0.06 (.02)		0.06 (.02)		0.06 (.02)
Evaluation of the SLE		0.02 (.01)		N.S.S.		N.S.S.
Stage						
Policy for teaching		0.03 (.01)		0.03 (.01)		N.S.S.
Evaluation policy for teaching		0.08 (.02)		0.06 (.02)		0.09 (.03)
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		N.S.S.		N.S.S.		N.S.S.
Partnership		N.S.S.		N.S.S.		N.S.S.
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		0.06 (.02)		0.06 (.02)		0.05 (.02)
Focus						
Policy for Teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		0.06 (.02)		N.S.S.		N.S.S.
School learning environment		N.S.S.		N.S.S.		N.S.S.
Quality						
Policy for Teaching		0.07 (.02)		0.06 (.02)		0.05 (.02)
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		N.S.S.		N.S.S.		N.S.S.
Partnership		0.08 (.03)		0.07 (.03)		0.08 (.03)
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		0.06 (.02)		0.05 (.02)		0.04 (.02)

Table 4   Continued

Factors	Low quality teaching		Average quality teaching		High quality teaching	
	Model 0	Final mod.	Model 0	Final mod.	Model 0	Final mod.
Differentiation						
Policy for teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
School learning environment		0.07 (.02)		N.S.S.		0.05 (.02)
Evaluation of SLE		N.S.S.		N.S.S.		N.S.S.
Variance components						
School	8.2%	3.2%	8.7%	4.5%	8.8%	4.2%
Class	7.1%	6.8%	6.7%	6.1%	6.8%	5.8%
Student	84.7%	33.0%	84.6%	35.0%	84.4%	34.5%
Explained		57.0%		54.4%		55.5%
Significance test						
χ^2	504.9	204.5	543.8	260.4	545.9	250.5
Reduction		300.4		283.4		295.4
Degrees of freedom		18		15		15
p value		.001		.001		.001

 *   **Low:** Schools with score of quality of teaching lower than –1; **Average:** schools with score between –1 and 1; **High:** schools with scores higher than 1

 ** N.S.S. = No statistically significant effect at .05 level.

further so that it is not concerned only with the current situation of schools and teachers but attempts also to illustrate the actions that have to be taken in order to improve their effectiveness. However, this study has revealed that, in just a few schools, changes in their effectiveness status was observed during two consecutive years. Although this does not imply that schools can remain effective over a long period without taking any improvement actions, the fact that we could not identify many schools, which either had a statistically significant decline or improvement in their effectiveness, should be seen as an indication that we need longitudinal studies that last for a longer period to measure changes in the effectiveness status of schools. This implies that our conclusions in relation to the research question concerning the influence of school factors on changing the effectiveness status of schools should be qualified by considering some limitations of our study and especially the fact that we examined changes in the effectiveness status over only two years. The importance of conducting longitudinal studies which will last for more than three years is emphasised. Another issue which has to be considered is the sampling procedure. Given that a large number of schools were found to be typical during the two consecutive years and others were found to be stable, a larger sample is needed to find schools that change in their effectiveness status. In this way, the statistical power will be increased and will enable us to identify variables which may be associated with changes in the effectiveness status of schools. The other alternative is to use a purposive sampling approach and select schools based on the fact that changes in their context

Table 5 Parameter estimates and (standard errors) for the analyses of RE (Affective) achievement in schools with low/typical/high quality of teaching*

Factors	Low quality teaching		Average quality teaching		High quality teaching	
	Model 0	Final mod.	Model 0	Final mod.	Model 0	Final mod.
Fixed part (Intercept)	0.25 (.07)	0.08 (.02)	0.35 (.13)	0.18 (.02)	0.25 (.07)	0.08 (.02)
Student Level						
Prior knowledge		0.26 (.07)		0.16 (.07)		0.19 (.07)
Sex (boys = 0, girls = 1)		0.08 (.03)		N.S.S.		N.S.S.
SES		N.S.S.		N.S.S.		N.S.S.
Classroom Level: Context						
Average prior knowledge		0.12 (.04)		0.10 (.04)		0.12 (.04)
Average SES		N.S.S.		N.S.S.		N.S.S.
Percentage of girls		N.S.S.		N.S.S.		N.S.S.
School Level						
Context						
Average SES		N.S.S.		N.S.S.		N.S.S.
Average prior knowledge		0.07 (.02)		0.08 (.02)		0.07 (.02)
Percentage of girls		N.S.S.		N.S.S.		N.S.S.
Frequency						
Policy for Teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		0.10 (.03)		0.08 (.02)		0.09 (.03)
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		N.S.S.		N.S.S.		N.S.S.
Relations with parents		0.12 (.03)		0.07 (.02)		0.10 (.03)
Relations with the centre		N.S.S.		N.S.S.		N.S.S.
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		N.S.S.		N.S.S.		N.S.S.
Stage						
Policy for teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		0.10 (.03)		0.07 (.03)		0.08 (.02)
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		0.08 (.04)		0.06 (.02)		0.07 (.03)
Partnership		N.S.S.		N.S.S.		N.S.S.
Provision of resources		0.07 (.03)		0.07 (.03)		0.09 (.03)
Evaluation of the SLE		0.07 (.02)		0.08 (.02)		0.11 (.04)
Focus						
Policy for Teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
School learning environment		N.S.S.		N.S.S.		N.S.S.
Quality						
Policy for Teaching		0.07 (.02)		0.06 (.02)		N.S.S.
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
Student behaviour out of class		N.S.S.		N.S.S.		N.S.S.
Collaboration among teachers		N.S.S.		N.S.S.		N.S.S.
Partnership		N.S.S.		N.S.S.		N.S.S.
Provision of resources		N.S.S.		N.S.S.		N.S.S.
Evaluation of the SLE		0.08 (.02)		0.07 (.02)		0.08 (.02)

Table 5 Continued

Factors	Low quality teaching		Average quality teaching		High quality teaching	
	Model 0	Final mod.	Model 0	Final mod.	Model 0	Final mod.
Differentiation						
Policy for teaching		N.S.S.		N.S.S.		N.S.S.
Evaluation policy for teaching		N.S.S.		N.S.S.		N.S.S.
School learning environment		0.05 (.02)		0.04 (.02)		0.05 (.02)
Evaluation of SLE		N.S.S.		N.S.S.		N.S.S.
Variance components						
School	7.2%	3.2%	7.6%	3.8%	7.0%	3.9%
Class	5.8%	5.0%	5.0%	4.6%	5.9%	5.2%
Student	87.0%	41.0%	87.4%	43.0%	87.1%	43.1%
Explained		50.8%		48.6%		47.8%
Significance test						
χ^2	564.9	343.5	554.5	353.2	565.1	382.0
Reduction		221.4		201.3		183.1
Degrees of freedom		13		12		11
p value		.001		.001		.001

* **Low:** Schools with score of quality of teaching lower than –1; **Average:** schools with score between –1 and 1; **High:** schools with scores higher than 1

** N.S.S. = No statistically significant effect at .05 level.

take place. In this case, one may assume that in these schools changes in the functioning of classroom and school factors could also be identified especially since schools should respond to changes in their context in order to maintain their effectiveness status or even to improve it. Such studies may help not only to develop the dynamic model of effectiveness but also to establish stronger links between research on educational effectiveness and improvement of practice.

We should also consider the possibility of conducting case studies which will enable us to measure the functioning of school factors in specific contexts and identify the reasons for which changes in the effectiveness status of some schools can be observed. In such a case, the dynamic model could be used as a heuristic approach in designing case studies that will not only look at the functioning of the school in relation to changes in its context but also search for impact on learning and student outcomes. Furthermore, this approach could be seen as the starting point for action research projects to improve the effectiveness status of the schools using an evidence-based and theory-driven approach.

Finally, the fact that we found more time stability on school effects than is expected in the literature may be attributed to the functioning of the system level factors in Cyprus and especially to the fact that no systematic reform, either on the educational policy in respect to teaching (including time and opportunity to learn) or on improving the learning environment of the schools,

takes place. Moreover, the system factors concerned with evaluation are rather weak and evaluation data are not systematically collected by external agencies. This can be attributed to the political dimension of evaluation (Kyriakides & Demetriou, 2007) implying that changes in the evaluation policy affect power relations between the central level (i.e. Ministry of Education) and the periphery (i.e. schools and teachers) and may produce strong resistance to evaluation reform policies. Moreover, actions on improving the quality of education are not based on evidence. Therefore, the only pressure for changes in the effectiveness status of schools may come from changes in contextual factors such as the transfer of teachers from one school to another. It is important to note that in Cyprus teachers are appointed and transferred by a central committee and in some schools the composition of teaching staff may change dramatically at the beginning of a new school year. Therefore, comparative studies looking at the effect of system factors on school effectiveness status should be conducted, in order to clarify further the reasons for which more stability or change of effectiveness status in some countries than in others can be observed and also test the generic nature of the dynamic model in respect to the phenomenon of stability and change in educational effectiveness.

References

Andrich D 1988. A general form of Rasch's Extended Logistic Model for partial credit scoring. *Applied Measurement in Education*, 1:363-378.

Barber B 1986. Homework does not belong on the agenda for educational reform. *Educational Leadership*, 43:55-57.

Brekelmans M, Sleegers P & Fraser B 2000. Teaching for active learning. In: PRJ Simons, J van der Linden L & T Duffy (eds). *New learning*. Dordrecht: Kluwer Academic Publishers.

Brophy J & Good TL 1986. Teacher behaviour and student achievement. In: MC Wittrock (ed.). *Handbook of research on teaching.* New York: MacMillan.

Creemers BPM 1994. *The effective classroom.* London: Cassell.

Creemers BPM & Kyriakides L 2006. Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement*,17:347-366.

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

Creemers BPM & Reezigt GJ 1999. The concept of vision in educational effectiveness theory and research. *Learning Environments Research*, 2:107-135.

Cronbach LJ, Gleser GC, Nanda H & Rajaratnam N 1972. *The Dependability of Behavioral Measurements: Theory of Generalizability Scores and Profiles*. New York: Wiley.

Donaldson L 2001. *The Contingency Theory of Organizations: Foundations for Organisational Science.* Thousands Oaks, CA: Sage.

Fraser BJ, Walberg HJ, Welch WW & Hattie JA 1987. Syntheses of educational productivity research. *International Journal of Educational Research*, 11:145-252.

Goldstein H 2003. *Multilevel statistical models*, 3rd edn. London: Edward Arnold.

Joyce B, Weil M & Calhoun E 2000. *Models of teaching.* Boston: Allyn & Bacon.

Kyriakides L 2005. Extending the Comprehensive Model of Educational Effectiveness by an Empirical Investigation. *School Effectiveness and School Improvement*, 16:103-152.

Kyriakides L, Campbell RJ & Christofidou E 2002. Generating criteria for measuring teacher effectiveness through a self-evaluation approach: A complementary way of measuring teacher effectiveness. *School Effectiveness and School Improvement*, 13:291-325.

Kyriakides L, Campbell RJ & 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*, 11:501-529.

Kyriakides L, Charalambous C, Philippou G & Campbell RJ 2006. Illuminating reform evaluation studies through incorporating teacher effectiveness research: a case study in Mathematics. *School Effectiveness and School Improvement*, 17:3-32.

Kyriakides L & Creemers BPM 2007. Teacher and school factors explaining student achievement: Testing the Dynamic model of Educational Effectiveness. Paper presented at the *88th Annual Meeting of the American Educational Research Association*, Chicago, USA.

Kyriakides L & Creemers BPM 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*, 19:183-205.

Kyriakides L, Creemers BPM, Antoniou P & Demetriou D (in press). A synthesis of studies searching for school factors: Implications for theory and research. *British Educational Research Journal*.

Kyriakides L & Demetriou D 2007. Introducing a teacher evaluation system based on teacher effectiveness research: an investigation of stakeholders' perceptions. *Journal of Personnel Evaluation in Education*, 20:43-64.

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.

Mintzberg H 1979. *The structuring of organizations*. Englewood Cliffs, NJ: Prentice-Hall.

Muijs D & Reynolds D 2001. *Effective Teaching: evidence and practice*. London: Sage.

Opdenakker MC & Van Damme J 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*, 11:65-196.

Rosenshine B & Stevens R 1986. Teaching Functions. In: MC Wittrock (ed.). *Handbook of Research on Teaching*. New York: Macmillan.

Scheerens J 1992. *Effective Schooling: Research, Theory and Practice*. London: Cassell.

Scheerens J, Seidel T, Witziers B, Hendriks M & Doornekamp G 2005. *Positioning and validating the supervision framework*. Enschede/Kiel: University of Twente, Department of Educational Organisational and Management.

Schoenfeld AH 1998. Toward a theory of teaching in context. *Issues in Education*, 4:1-94.

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

Shavelson RJ, Webb NM & Rowley GL 1989. Generalizability theory. *American*

Psychologist, 44:922-932.

Slater RO & Teddlie C 1992. Toward a theory of school effectiveness and leadership. *School Effectiveness and School Improvement*, 3:247-257.

Snijders T & Bosker R 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modelling*. London: Sage.

Stringfield S 1994. A model of elementary school effects. In: D Reynolds, BPM Creemers, PS Nesselrodt, EC Schaffer, S Stringfield & S Teddlie (eds). *Advances in School Effectiveness Research and Practice*. Oxford: Pergamon Press.

Teddlie C & Reynolds D 2000. *The International Handbook of School Effectiveness Research*. London: Falmer Press.

Webster BJ & Fisher DL 2003. School–level environment and student outcomes in mathematics. *Learning Environment Research*, 6:309-326.

Witziers B, Bosker JR & Kruger LM 2003. Educational Leadership and Student Achievement: The Elusive Search for an Association. *Educational Administration Quarterly*, 39:398-425.

Authors

Bert Creemers is an Honorary Professor in Educational Sciences at the University of Groningen in the Netherlands. He is the founding editor of two journals, *School Effectiveness and School Improvement* and *Educational Research and Evaluation*. His main research interests are educational effectiveness, especially teacher effectiveness, and educational evaluation and improvement.

Leonidas Kyriakides is Associate Professor of Educational Research and Evaluation at the University of Cyprus. He is the book review editor of the journal *School Effectiveness and School Improvement*. His main research interests are modelling educational effectiveness, research methods in educational effectiveness, and educational evaluation and improvement.