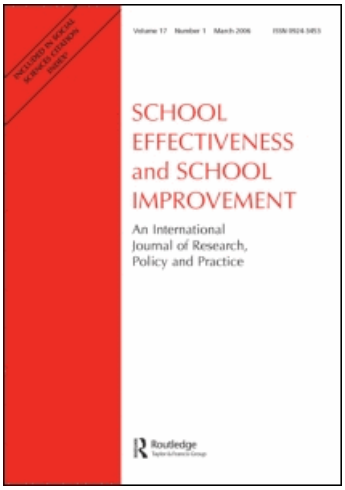

School Effectiveness and School Improvement
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t714592801

Classroom and school factors related to student achievement: what works for students?

Jelena Teodorović[a]
[a] Institute for Educational Research, Belgrade, Serbia

Online publication date: 13 May 2011

To cite this Article Teodorović, Jelena(2011) 'Classroom and school factors related to student achievement: what works for students?', School Effectiveness and School Improvement, 22: 2, 215 — 236
To link to this Article: DOI: 10.1080/09243453.2011.575650
URL: http://dx.doi.org/10.1080/09243453.2011.575650

School Effectiveness and School Improvement
Vol. 22, No. 2, June 2011, 215–236

Routledge
Taylor & Francis Group

Classroom and school factors related to student achievement: what works for students?

Jelena Teodorović*

Institute for Educational Research, Belgrade, Serbia

(*Received 1 October 2009; final version received 1 March 2011*)

This study aimed to identify the classroom and school characteristics that are associated with student achievement in mathematics and Serbian language in primary schools in Serbia. The study sample consisted of 119 public primary schools, 253 classrooms, and 4,857 third-grade students. Variables from 3 past research paradigms – input-output, effective schools, and instructional effectiveness – as well as key student background variables, were first organised in a conceptually integrated model of school effectiveness. Then, they were simultaneously examined as predictors of student achievement in a 3-level Hierarchical Linear Model (HLM). This research found little association between school-level variables and student achievement. Several classroom-level policy-malleable variables were found to have small positive associations with student achievement: clear and structured classroom instruction, emphasis on complex (as opposed to basic) skills, whole-class instruction (rather than individual or group work), teacher feedback, orderly climate, moderately frequent reinforcement of student effort, and – to a lesser extent – use of a variety of teaching tools.

Keywords: school effectiveness; input-output; effective schools; instructional effectiveness; Serbia; hierarchical linear modelling

Introduction

Since the pioneering efforts of Carroll (1963) and Bloom (1976), thousands of studies have sought to identify policy-malleable factors that impact student learning. Much of this research has suffered from the following problems: (a) researchers mostly focused on rather isolated sets of variables (those from input-output, effective schools, and instructional effectiveness paradigms; Scheerens, 2000), or, when different paradigms were integrated, they have not been guided by the conceptual model of school effectiveness; (b) statistically inappropriate methods were used to examine dependent levels (student, classroom, and school) in the hierarchical nature of schooling (Scheerens, 2000), or often one of the levels was omitted (Luyten, 2003; Opdenakker & Van Damme, 2000b); (c) teacher-related effects were difficult to determine as students are usually assigned to different teachers on an annual basis; and (d) countries that are neither fully industrialised nor fully developing have been mostly neglected in school effectiveness research.

*Email: jelenat@gwu.edu

ISSN 0924-3453 print/ISSN 1744-5124 online

DOI: 10.1080/09243453.2011.575650
http://www.informaworld.com

In the attempt to address these issues, this study examines a wide variety of variables in a conceptually integrated, three-level model of school effectiveness, utilising data on students who have spent 3 years with the same teacher in Serbia, a transitioning country in a region that has not been adequately explored in school effectiveness research (SER). Results from this study may pave the way toward re-thinking and reassessing of the priorities for investment, teaching methods, and school management in Serbia.

It is important to note that this study falls squarely in the SER realm, and as such it is somewhat limited for policymaking purposes. More school improvement research – research of intervention programs – is also needed to fully address the factors that impact student learning.

International research on school effectiveness

Input-output research

Throughout this paper, the term *input-output* refers to studies where quantifiable school-related characteristics are mostly taken as school inputs and student achievement test scores are mostly taken as a measure of school output (Hanushek, 1989). Student characteristics, which are often also treated as inputs, are addressed in a separate section. Large data sets were analysed, and the following variables were found to associate with student achievement in single studies or meta-analyses: teacher test scores, education, certification, experience and salary, per-student funding, and student–teacher ratio (Darling-Hammond, 2000; Ferguson, 1991; Goldhaber & Brewer, 1997; Hedges, Laine, & Greenwald, 1994; Payne & Biddle, 1999). On the other hand, Rivkin, Hanushek, and Kain (2000), the Programme for International Student Assessment (PISA) 2000 (Organisation for Economic Co-operation and Development [OECD], 2001), and Hanushek (1989, 1997) found that many of these input-output variables played a very small role in explaining student achievement.

There are several problems with both primary studies and meta-analyses: Output measures are not consistently controlled for student socioeconomic status (SES) or prior achievement, aggregated data are used often, and studies vary dramatically in the choice of variables included in regression analyses. Furthermore, existing measures of school inputs are likely inadequate to fully capture the true effects of schools (Hanushek, 1989), and most input-output studies assume a simple and direct relationship between inputs and outputs (Scheerens, 2000).

Findings of the input-output studies undertaken in developing countries are less controversial. The reviews of the studies that controlled student background (Farrell & Oliveira, 1993; Fuller & Clarke, 1994; Hanushek, 1994; Velez, Schiefelbein, & Valenzuela, 1993) showed that in many studies, the existing measures of school inputs, such as basic infrastructure, textbook availability, teacher education, per-student expenditure, and school facilities, consistently exhibited a significant impact on student achievement in developing countries. Heyneman and Loxley (1982) provided some estimates of school effects: In developing countries, schooling variables (analysed collectively) explained 2 to 3 times more achievement variance than in industrialised countries (after controlling for same student background factors).

Farrell and Oliveira (1993) explain the differences between the findings on student background and school inputs in industrialised and developing countries as

the industrialised countries being close to the limits of the perfectibility of the technology of schooling, so that even modest additional gains in achievement require costly and difficult educational efforts. Similarly, Scheerens (2000) attributes greater impact of school inputs in developing countries, compared to industrialised countries, to the greater variance in both school inputs and outputs.

Effective schools research

In the effective schools research, researchers mostly used surveys, case studies, and field studies to study the organisation of high-end outlier schools (Levin & Lockheed, 1991; Scheerens, 2000). Several prominent studies found the following effective schools variables to associate with student achievement: staff cohesion in academic and disciplinary matters, pleasant working environment, principal's leadership, high expectations for students, school goals, inter-staff relations, emphasis on academic achievement, encouragement and active engagement of parents, strong management team, and quality teaching at the school (Henderson & Mapp, 2002; Reynolds, Creemers, Stringfield, Teddlie, & Schaffer, 2002; Rutter, Maughan, Mortimore, & Ouston, 1979; Sammons et al., 1998; Walberg & Paik, 2000). However, some quantitative studies find little association between school environment and organisational variables and student achievement (Ellett et al., 1997; OECD, 2001; Van der Werf, 1997; Wang, Haertel, & Walberg, 1993).

Reasons behind discrepant findings are given by several authors: varied definition and operationalisation of effective schools factors between studies, a restricted range of variation in school organisation and content, failure to examine cultural factors, and omission of additional variables that impact student achievement.

Effective schools factors have been rarely examined in developing countries (Scheerens, 2000). Fuller and Clarke (1994) found that, in over 100 studies, effective schools factors were examined only three or four times. In Latin America and the Caribbean, the effective schools factors were rarely explored (Velez et al., 1993). In over 50 studies reviewed by Farrell and Oliveira (1993), four out of seven analyses found the quality of the principal to be significantly correlated with student achievement.

Instructional effectiveness research

The instructional effectiveness studies have mostly focused on teacher behaviours and practices in the classroom and were often based on experimental design (Scheerens, 2000; Walberg & Paik, 2000).

Several important reviews of instructional effectiveness research (Brophy & Good, 1986; Creemers, 1994; Scheerens, 2000; Walberg & Paik, 2000; Wang et al., 1993) singled out the following teaching behaviours as the ones most consistently associated with student achievement in industrialised countries: emphasising academic instruction; maximising efficient time on task; actively teaching (vs. allowing individual, unsupervised study by students); adjusting the difficulty and cognitive level of tasks and questions to the students; structuring, outlining, and reviewing lessons; questioning, testing, and providing homework; prompting and providing feedback; ensuring clear correspondence between covered material and tests (so-called opportunity to learn); monitoring for completion and accuracy in supervised independent seatwork and homework; teaching of learning strategies;

providing corrective instruction; preparing in advance; being flexible, clear, and enthusiastic; having high expectations; maintaining an orderly atmosphere; having quality academic and social interactions with students.

Several problems with instructional effectiveness research exist. First, some of the experimental research was done in a short period of time, with unknown results in long-term educational situations. Second, non-experimental studies on instructional effectiveness might artificially show large results if they omitted other theoretically implicated variables from consideration. Third, instructional effectiveness variables likely exert differential effects in different contexts (different student composition, grade levels, subject matter, etc.). Fourth, experimental research may be susceptible to the Hawthorne effect, that is, any innovation can artificially increase the effects of the experiment simply due to its novelty (Weiss, 1998), but the effect itself has been heavily disputed (Adair, Sharpe, & Huynh, 1989).

The available evidence suggests that some instructional effectiveness factors, although rarely explored, may play important roles in developing countries. For example, significant positive associations were found in several reviews between student achievement and instructional time, frequency of homework, homework practices, teacher's expectations of student performance, and teacher's time spent on class preparation (Farrell & Oliveira, 1993; Fuller & Clarke, 1994; Velez et al., 1993).

While consensus seems to be forming in industrialised countries that instructional effectiveness factors are important for student achievement, research on instructional effectiveness in both the industrialised and developing world would benefit from studies that simultaneously include other relevant student background and school effectiveness factors, use precisely specified and operationalised factors, and show a substantial range of variation in instructional practices.

Integrated school effectiveness research

In recent school effectiveness studies in industrialised countries, the differences between student scores on achievement tests were more attributable to the characteristics of individual students and their peer groups than to differences due to attending different classrooms and schools (Bosker & Witziers, as cited in Scheerens & Bosker, 1997; OECD, 2004; Opdenakker, Van Damme, De Fraine, Van Landeghem, & Onghena, 2002). Part of the unexplained, residual variance is likely due to the effect of unobserved student-level variables, and part is due to classroom-level and school-level variables that affect student achievement. Even though many studies suggest that this unexplained, residual variance is small in absolute terms (about 5–15%, OECD, 2004; Rowan, Correnti, & Miller, 2002; Scheerens & Bosker, 1997), there are ongoing efforts in the educational research community to identify specific classroom and school variables that would explain this part of the residual variance.

Results of the integrated studies on the effects of classroom-level factors in industrialised countries identified significant but small effects of these variables on student achievement: structured lessons, intellectually challenging teaching, a work-centred environment, limited focus within lessons, maximum communication between teachers and students, record keeping, parental involvement, and a positive climate at the classroom level (Mortimore, Sammons, Stoll, Lewis, & Ecob, 1988), positive feedback, emphasis of key lesson points, checking for student comprehension, frequent high-quality, academic-related questioning, motivating students, and showing high expectations (Reynolds et al., 2002), homework assignments, emphasis

on basic skills in early primary grades, introduction of advanced skills in middle primary grades, and ability grouping (D'Agostino, 2000), teacher attendance of a specialist literacy in-service course, teacher's high expectations and successfully matching instructional levels and student abilities (Hill & Rowe, 1998), and calm, learning-focused climate in the classroom (Opdenakker et al., 2002).

While these individual effects showed small association with achievement, Muijs and Reynolds (2000) found that the composite teaching quality variable, derived by summing the scores for classroom management, behaviour management, direct teaching, individual practice, interactive teaching, varied teaching, and classroom climate, explained between 60% and 100% of the unexplained between-classroom variance in student learning in various grades, after accounting for student prior achievement and other student background factors. (Unexplained between-classroom variance after including the controls comprised 7–24% of the total variance.) Even though this study did not control for classroom compositional effects, which could potentially reduce the effect of the composite variable (Opdenakker et al., 2002), and did not examine other classroom-level variables (e.g., teachers' preparation), or school-level variables, this finding implies that various effective teaching behaviours go together and that they have a large impact on student achievement. Rowan et al. (2002) also suggested that many small instructional effects need to be combined to produce a large effect on student learning.

Additionally, it is possible that identification of larger effects of teacher-related factors on student achievement remains elusive because in many countries, students annually change teachers, obscuring possible cumulative effects. Relatively strong and weak teachers will tend to cancel each other out, resulting in a weak overall relationship between teacher-related variables and student achievement. Alternatively, focusing on only one year may be too short a time to allow identification of large effects. Sanders and Rivers (1996) used a value-added method to determine that students with repeated exposure (three times in a row) to well-qualified teachers performed up to 50 percentile points better on mathematics tests than those with the same repeated exposure to poorly qualified teachers.

Integrated studies that examined school-level factors showed mixed effects on student achievement. The following school-level variables were positively associated with student achievement: head teacher leadership, teacher involvement, consistency among teachers, record keeping, parental involvement, and positive climate (Mortimore et al., 1988), social support and the shared mission of teachers, little decision-making, development, and planning activities at school (D'Agostino, 2000), teaching staff cooperation regarding teaching methods, pupil counselling, and an orderly learning environment (Opdenakker & Van Damme, 2000a). One exemplary study that controlled for individual student differences, classroom compositional, and classroom instructional variables, found that school-level variables – teacher attention to student differences and the amount of teacher consultation at school – negatively associated with student achievement (Opdenakker et al., 2002). On the other hand, Hill and Rowe (1998), PISA 2000 (OECD, 2001), and Webster and Fisher (2000) found that none of the various factors related to school climate, teacher autonomy and morale, and school resources associated with student achievement.

Integrated studies undertaken in developing countries (Dowd, 2001; Lockheed & Longford, 1989; Nyagura & Riddell, 1993; Willms & Somers, 2001) also showed that between-student variance was larger than between-classroom and between-school variances, but to a lesser extent than in industrialised countries. Some of the

classroom-level variables that were associated with student achievement are: amount of instructional time, the amount of teacher-supervised study, and the number of classroom math textbooks (Nyagura & Riddell, 1993), child-centred teaching and participatory teaching (Dowd, 2001), enriched mathematics curriculum and frequency of textbook use (Lockheed & Longford, 1989), no ability grouping and multigrade classrooms, regular testing and a positive classroom climate (Willms & Somers, 2001).

The same integrated studies showed that the role of school-level input-output variables varied among countries, while school-level effective schools variables were not sufficiently explored in developing countries. For example, in Zimbabwe, the percentage of trained teachers, the availability of textbooks, and pupil–teacher ratio explained 20% of the total variance in student English achievement, compared to 17% explained by student-level factors (Nyagura & Riddell, 1993). On the other hand, in Thailand and Latin America, input-output factors appeared to play a smaller role in student achievement than student-background factors (Lockheed & Longford, 1989) or both student- and classroom-level factors (Willms & Somers, 2001). In Zimbabwe, no effective schools variables were associated with achievement (Nyagura & Riddell, 1993), and in Malawi only the inclusion of community monitoring of teachers showed significant impact on student achievement (Dowd, 2001).

The relative importance of classroom-level and school-level factors varied between countries. Glewwe, Grosh, Jacoby, and Lockheed (1995) and Scheerens (1999) explained that this variation may be due to the possibility that classroom-level factors might begin to show their effects more strongly and school-level input-output factors less strongly when the educational system of a developing country begins to resemble educational systems of industrialised countries.

Most often, integrated studies have not followed a conceptual model in their choice of variables and hierarchical levels. Several SER researchers proposed comprehensive, conceptually integrated models on the basis of their extensive literature reviews and meta-analyses of school effectiveness research (Creemers, 1994; Heneveld & Craig, 1996; Scheerens, 1990; Wang et al., 1993). Kyriakides, Campbell, and Christofidou (2002), Reezigt, Guldemond, and Creemers (1999), and Van der Werf, Creemers, and Guldemond (2001) tested Creemers' model in Cyprus, The Netherlands, and Indonesia, respectively, and confirmed that student background factors account for a large portion of variance of the test scores. Classroom- and school-level variances were small, and classroom- or school-level variables (other than peer effect) that were associated with student achievement were either scant (Kyriakides et al., 2002; Reezigt et al., 1999) or not uniform in direction or subject (Van der Werf et al., 2001). More studies that examine a variety of variables on all three levels in conceptually integrated models need to be carried out.

In summary, the following conclusions emerge from school effectiveness research: (a) Student-level factors are very important in determining student achievement in industrialised countries and better-off developing countries, while their effect is less pronounced in poor developing countries; (b) classroom-level variables exhibit significant association with student achievement in industrialised and better-off developing countries, and less so in poor developing countries, with composite variables having a considerably larger effect than individual variables; and (c) school-level factors show the least consensus, with their likely effect sizes ranging from null to modest in industrialised and better-off developing countries, but school-level input-output variables are very important for poor developing countries.

Serbia – background

Serbia is a country that experienced decades of communist economic and educational policies only to transition toward democracy in a short, turbulent period. Serbian education features free and gender-equal access, adequate coverage of the student population, and the acceptable provision of basic school resources, basic teaching materials, and formally qualified teachers (International Bureau of Education, 2001). Yet, disadvantages still exist in rural areas and in some minority groups. Funding of education is unsatisfactory, as spending is inefficient (Levitas & Herczynski, 2006). Finally, the curriculum focuses on fact acquisition, teaching is teacher centred (United Nations Children's Fund [UNICEF], 2001), and mobility between different types of institutions in secondary and tertiary education is rigid.

When school inputs are concerned, Serbia appears to more resemble industrialised countries than developing countries, and as a result school inputs likely do not exhibit a strong relationship with student achievement. Very little attention is paid to effective schools factors in Serbia in both real life and research. For example, when educational leadership is concerned, Serbian principals (usually former teachers) do not undergo any compulsory training, as there are no undergraduate or graduate courses that teach administrative and management skills (UNICEF, 2001). In regards to instructional effectiveness, some seminars and pilot programs exist in practice (e.g., ''Active learning''). Research (mostly with nonrigorous design) found that more active forms of teaching (e.g., teachers using graphs and concepts or students preparing presentations, discussing, and writing research reports) improve student learning (Budic, 2000; Savic-Gutesa, 2000). This suggests that teaching-related variables may be important in Serbia, as is the case in industrialised and other mid-development countries. However, this is a speculation, and studies that examine relationships between student achievement and a wide variety of variables, and employ sophisticated methodological tools, are direly needed in Serbia.

Methods

Tests and questionnaires

In May 2004, the Serbian Institute for Education Quality and Evaluation (IEQE; in Serbian: Zavod za Vrednovanje Kvaliteta Obrazovanja i Vaspitanja – ZVKOV) undertook a national evaluative study of student achievement in primary education (see Baucal, Pavlovic-Babic, Gvozden, & Plut, 2006). The achievement tests in mathematics and Serbian language, developed by the IEQE, in consultation with the Educational Testing Service and the World Bank, were focused more on the application of knowledge than had previously been common in Serbia. Table 1 shows descriptive statistics for the two test scores of the student sample used in this study.

Three types of questionnaires were developed by this author in collaboration with the IEQE to collect information from students, teachers, and principals. The questionnaires were based on the conceptual model of school effectiveness,

Table 1. Descriptive statistics for dependent variables.

	N	Minimum	Maximum	*M*	*SD*
Mathematics IRT score	4857	32	937	501.1	95.753
Serbian language IRT score	4857	−33	1017	498.9	103.12

which was in turn created as a blend of four conceptual models proposed previously by SER researchers (Creemers, 1994; Heneveld & Craig, 1996; Scheerens, 1990; Wang et al., 1993). The conceptual model is described in Teodorović (in press).

Twenty-five student-level predictors representing seven out of eight larger constructs (Intelligence, Prior Achievement, Student Characteristics, SES, Parental Support, Time on Task, Opportunity to Learn, Home Environment) include the student's first-grade mark in the subject, number of children in the family, parental education, student's gender, student's age, developmental and family problems, Roma minority status, student's motivation for the subject, parental interest and involvement in the student's school work and life, student's reading, and so on. The relationship between these variables and student achievement is discussed in Teodorović (in press).

Classroom-level constructs in the model were: Teacher Characteristics, Classroom Instruction, Opportunity to Learn, Time on Task, Student Assessment and Feedback, Student and Teacher Social and Academic Interactions, and Classroom Climate. School-level constructs in the model consisted of School Demographics, School Infrastructure and Services, School Culture, and School Climate. Each of the constructs listed is represented by a set of variables. Detailed descriptions of the classroom- and school-level variables are given in Table 2 and Table 3, respectively.

Items on questionnaires were adopted and modified from other school effectiveness studies or collected from teachers and school records. They were mostly closed-ended 4-point Likert-type items. Most of the variables were constructed from items using Principal Component Analysis (PCA) and Principal Axis Factoring (PAF). At the classroom and school levels, continuous variables were normally distributed and standardised, except for several which were collapsed into dummy variables. More detailed discussion on the items and variables is offered in Teodorović (in press).

Even though 30 classroom-level predictors and 18 school-level predictors were constructed to represent the constructs, only 13 classroom-level variables and 5 school-level variables were used as final predictors of student achievement, in order to have an acceptable ratio of variables to sample size. This smaller set of classroom-level predictors was obtained by combining two approaches. The first approach was to include all 30 grand-mean-centred variables together in a classroom-level equation as predictors of student achievement (which was adjusted for the 25 student-level variables) and then eliminate insignificant variables in hierarchical linear modelling (HLM) iterations until all remaining ones were significant at or below the 0.20 level. The second approach first tested each of the 30 grand-mean-centred classroom-level variables individually in a classroom-level equation as a predictor of student achievement (that was adjusted for the 25 student-level variables), then ran together variables that had a p value equal to or lower than 0.20 in the HLM iterations, as was the case in the first approach. The four resulting models (two per subject) yielded similar results. Variables that were significant in at least two models were selected for the final model. The same approach was used for school-level variables.

Student-level variables, as well as classroom-level and school-level variables, were grand-mean centred in a three-level HLM model.

Sample

The sample of students in the study was created by randomly drawing two urban and two rural schools from each of the 25 administrative counties in Serbia. A total

Table 2. Descriptive statistics of classroom-level variables (unweighted).

Variable	Description
Teacher Characteristics	
Male teacher	Dichotomous, 1 = male teacher, 0 = female teacher. 18% of male teachers are in the sample.
Classroom Instruction	
Usefulness of teaching tools	Average of three 4-pt items in factor analysis in mathematics (teacher-rated usefulness in teaching of mathematical games, geometrical models, and meters, weights, and scales). Reliability of the scale is 0.72. Average of four 4-pt items in factor analysis in Serbian language (teacher-rated usefulness in teaching of language games, dictionaries, children's encyclopaedias, and various children's books). Reliability of the scale is 0.78.
Emphasis on basic over complex skills	Average of three items in factor analysis (teacher-estimated % of basic vs. complex problems/exercises assigned: during the class, for homework, and on tests). Separate variables for mathematics and Serbian language. Reliability of the scale for mathematics is 0.83. Reliability of the scale for Serbian language is 0.88.
Clarity and mastery teaching	Classroom-level aggregate of the average of seven student-reported 4-pt items in factor analysis (teacher reviews old lessons, teacher repeats the most important points of the lesson, teacher clearly and nicely explains things, teacher asks students whether they understood the lesson, teacher asks student question(s) related to the lesson, if something is unclear to the student teacher explains it again, and teacher comes over to student to check student's work). Reliability of the scale is 0.84.
Opportunity to Learn Time on Task	
Whole-class instruction	Teacher-reported % of class time spent on frontal lecturing and % of class time spent on whole-class discussion, added and converted to 10% increments. Separate variables for mathematics and Serbian language class time.
Student Assessment and Feedback	
Reliance on less direct assessment methods to assign a grade	Average of four 4-pt items in factor analysis (teacher-reported reliance on tests, homework, individual/group projects and student teaching, and seatwork to assign a grade). Separate variables for mathematics and Serbian language. Reliability of the scale for mathematics is 0.57. Reliability of the scale for Serbian language is 0.54.
Reliance on student social behaviour to assign a grade	Average of four 4-pt items in factor analysis (teacher-reported reliance on student effort, in-class activity, cooperation with others, and behaviour to assign a grade). Separate variables for mathematics and Serbian language. Reliability of the scale for mathematics is 0.70. Reliability of the scale for Serbian language is 0.65.
Teacher's feedback	Classroom-level aggregate of the average of two student-reported 4-pt items in factor analysis (teacher explains good and bad sides of student's answer during oral examinations, and teacher writes out good and bad sides of student's graded homework). Reliability of the scale is 0.64.
Frequency of grading homework	Classroom-level aggregate of one student-reported reverse-coded 4-pt item. Separate variables for mathematics and Serbian language.

(*continued*)

Table 2. (*Continued*).

Variable	Description
Student and Teacher Social and Academic Interactions	
Moderate vs. very frequent reinforcement of student effort	Dichotomous; collapsed average of two teacher-reported 4-pt items (praising student when the student knows an answer to the question and praising student when the student puts an effort into trying to answer the question). Reliability of the scale was 0.68. 1 = moderately frequent reinforcement; 0 = very frequent reinforcement. 26% of teachers praise students moderately vs. very frequently.
Classroom Climate	
Orderly climate	Classroom-level aggregate of the average of two student-reported reverse-coded 4-pt items (student does not pay attention to the lesson, and student argues with another student). Reliability of the scale is 0.58.
Control, Demographic, and Contextual variables	
Classrooms with over 20% students eliminated from HLM	Dichotomous, 1 = over 20% of students in the classroom were eliminated from HLM modelling; 0 = other. 7% of such classrooms are in the sample.
Average parental education	Classroom-level aggregate of the average of two teacher-reported items (student mother's and father's education).

The following variables were tested, but were not selected for the final HLM model for reasons described on page 222: teacher's degree, teacher satisfaction, years of teacher's experience, teacher's grade point average (GPA), time teacher spends in preparation for mathematics/Serbian language lesson, teacher self-efficacy beliefs, teacher's attendance of in-service training in last 3 years, similarity between mathematics/Serbian language national test and regular homework and tests, reliance on direct assessment methods (oral questioning and boardwork) to assign a mathematics/Serbian language grade, frequency of mathematics/Serbian language in-class grading, high expectations of students, classroom size, and classroom location (outpost rural, multigrade rural, regular rural, regular urban).

Table 3. Descriptive statistics of school-level variables (unweighted).

Variable	Description
School Demographics	
School size	Principal-reported item converted to 100 student increments.
School Infrastructure and Services	
Number of computers in school	Principal-reported item converted to 2 computer increments.
School Culture	
Time principal and teachers spend talking about teaching	Principal-reported % of time spent talking to teachers about teaching during informal discussions and % spent during formal meetings, added and converted to 10% increments.
School Climate	
School climate	School-level aggregate of the average of 10 teacher-reported 4-pt items in factor analysis (teachers' assessment of: efficacy of formal meetings in school, frequency of assessing student yearly progress in school, order and discipline in school, relationships between school staff, fairness in dividing teacher obligations, principal's interest in classroom functioning, principal's dedication to teacher's professional improvement, principal's availability to talk about teachers' concerns, principal's appreciation of teachers' work, and the

(*continued*)

Table 3. (*Continued*).

Variable	Description
	extent of teachers' support for principal's decisions). Reliability of the scale is 0.93.
Teacher decision-making in school	School-level aggregate of 10 teacher-reported items (whether or not teacher makes decisions on disciplining students, grade repetition and course failing, organisation of lessons, assessment methods, teaching methods, using school budget for teaching tools, hiring of teachers, firing of teachers, choice of textbooks, and professional training of teachers, added)

The following variables were tested, but were not selected for the final HLM model for reasons described on page 222: facilities in school, teaching-related materials and equipment, number of pedagogues, psychologists, or defectologists per 100 students, number of books for students in Grades 1–4 in library, high expectations, frequency of formal meetings in school, amount of principal's in-service training in last 3 years, principal's self-efficacy beliefs, principal's experience, principal's GPA, school climate, teacher attitudes, time during formal meetings in school spent discussing improvement of student achievement, and time during formal meetings in school spent discussing improvement of teaching.

sample of 119 public primary schools was selected, with the average number of students in Grades 1–4 (lower primary) of 339 (only four schools had more than 700). The average number of third-grade classrooms in school (including outpost and remote multigrade classrooms) was 4.7 per school. The average number of third-grade classrooms located *within* the school (81% of all classrooms) was 3.1 per school.

In each school, two third-grade classrooms (regular, outpost, or remote multigrade) were randomly chosen to participate in the study. A sample of 253 classrooms was selected with the average number of students in a classroom of 21.5.

All students present in the classroom during the day of testing responded to the Student questionnaire and took achievement tests in mathematics and Serbian language. There were 5,216 students in the drawn sample and 4,857 in the analysed sample, making for an excellent participation rate of 93.2%. All teachers of selected classrooms and all principals of selected schools responded to their questionnaires. For more detailed discussion of the sampling, see Teodorović (in press).

Findings

Random effects

Teodorović (in press) describes a fully unconditional model, Model 0, and two student-level models, Model 1A (includes 24 student background variables) and Model 1B (includes 24 student background variables and a student grade from the end of the first grade). In Model 2, 13 classroom-level variables were added to Model 1B, and then, in Model 3, five school-level variables were added to Model 2 (Table 4).

In Model 0, 86.7% of the variance in mathematics scores lies between the students, 5.4% lies between the classrooms, and 7.9% lies between the schools. In Serbian language, 83.9% of variance in test scores is due to the students, 6.0% is due to the classrooms, and 10.0% is due to the schools.

In Model 1B, 25 student-level variables explained slightly less than half of the student-level variance in mathematics and Serbian language (41.3% and 40.9% of the *total* variance, respectively). These 25 predictors explained very little of the

Table 4. Unexplained variances in fully unconditional model, student-level, classroom-level, and school-level models.

	Mathematics				Serbian language			
Unexplained variance	Model 0	Model 1B	Model 2	Model 3	Model 0	Model 1B	Model 2	Model 3
Student	0.871	0.456	0.456	0.456	0.844	0.432	0.432	0.432
(*SE*)	(0.018)	(0.010)	(0.009)	(0.009)	(0.018)	(0.009)	(0.009)	(0.009)
Classroom	0.054	0.052	0.040	0.041	0.061	0.053	0.045	0.048
(*SE*)	(0.013)	(0.010)	(0.008)	(0.008)	(0.014)	(0.010)	(0.009)	(0.009)
School	0.080	0.040	0.017	0.008	0.101	0.057	0.024	0.016
(*SE*)	(0.018)	(0.011)	(0.007)	(0.006)	(0.022)	(0.014)	(0.009)	(0.008)

classroom-level variance in mathematics and Serbian language (0.2% and 0.8% of the *total* variance, respectively) and about half of the school-level variance in mathematics and Serbian language (4.0% and 4.3% of the *total* variance, respectively).

Thirteen classroom-level variables explained about one sixth of the *unexplained* classroom-level variance in mathematics and Serbian language (1.2% and 0.8% of *total*, respectively) after controlling for student-level predictors. This showed that classrooms exhibited different achievement because they enrolled students with different characteristics, but even more so because they had different classroom characteristics. Thirteen classroom-level predictors also explained more than half of the *unexplained* school-level variance in mathematics and Serbian language (2.3% and 3.3% of the *total* variance, respectively) after controlling for student-level predictors. Therefore, school achievement varied in large part due to different characteristics of the students, but also, to a somewhat lesser degree, due to different characteristics of the classrooms.

Five school-level variables explained about half of the *unexplained* school-level variance in mathematics (0.9% of total) and a third of the *unexplained* school-level variance in Serbian language (0.8% of total) that was still unexplained after controlling for student- and classroom-level variables. Therefore, (in)effectiveness of schools can be attributed in very small part to different school traits.

Figure 1 and Figure 2 show a very similar final distribution of student-level, classroom-level, and school-level variance in mathematics and Serbian language.

Adding up percentages of the *total* variance that are explained by variables at each schooling level (Figure 1 and Figure 2), one calculates that student-level predictors explained around 45% of the *total* variance in student achievement, classroom-level predictors explained around 3.5–4% of the *total* variance, and school-level predictors explained around 1% of the *total* variance.

In summary, about 85% of the total variance in student achievement lies between the students, and only about 15% lies between classrooms and schools combined. Of these 15% of the total variance that are within realm of education policy, one third is explained by student-level variables used in this study, another third by classroom-level and school-level variables together (mostly classroom), and the last third is left unexplained.

The finding that only 15% of the variance in Serbia's third-grade achievement was between classrooms and schools combined is similar to the 19% average found

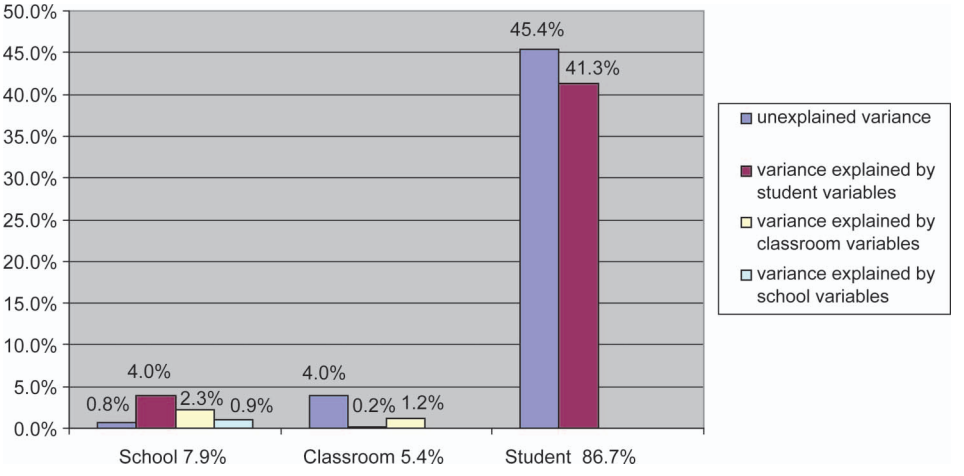

Figure 1. Mathematics, overview of the variance distribution.

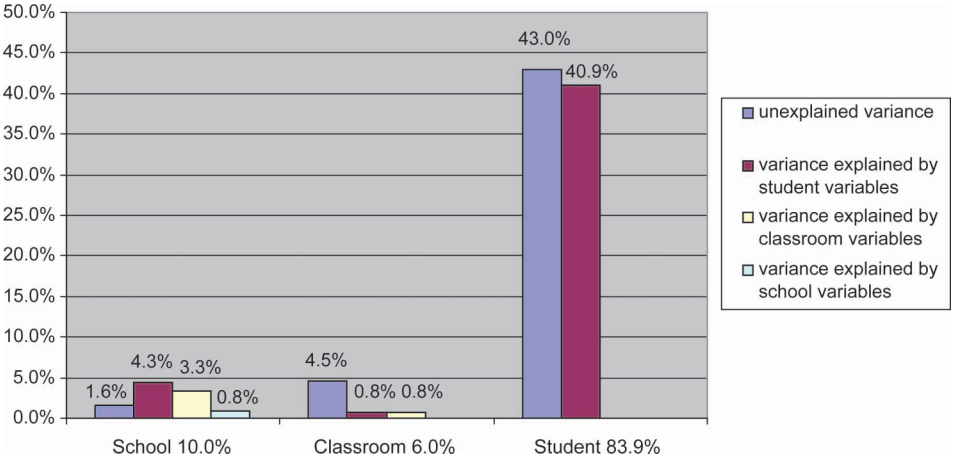

Figure 2. Serbian language, overview of the variance distribution.

between schools in a meta-analysis of over a hundred school effectiveness studies (Bosker & Witziers, as cited in Scheerens & Bosker, 1997).

In PISA 2003, the between-school variance in Serbia comprised 35.2% of the total variance in mathematics (OECD, 2004). The difference between this study and PISA 2003 is likely due to the following: (a) Serbian classrooms and schools in lower primary school grades are likely more uniform than classrooms and schools in already differentiated secondary education attended by the 15-year-old participants in PISA 2003; (b) unlike uniform and comprehensive pre-service training of lower primary teachers, pre-service training of upper primary and secondary school teachers is varied, unrelated to school curricula, and almost entirely comprised of subject-content courses (UNICEF, 2001); (c) the achievement tests used for this study directly targeted the nationally prescribed third-grade curriculum in Serbian primary schools, while the PISA test was not aligned with a myriad of curricula in

Serbian secondary education; and (d) this study explores all three schooling levels, while PISA 2003 focused only on the student and school levels, which omission may lead to an inappropriate distribution of variances and inaccurate estimates of variable coefficients.[1]

In this study, initial classroom-level and school-level variances in student achievement (5–6% and 8–10% of the total variance, respectively) were either similar to or slightly smaller than those identified in other studies (Kyriakides et al., 2002; Reezigt et al., 1999; Rowan et al., 2002). It should be noted that classroom-level variance in this study is partly an artefact of selecting only two classrooms from each school (26 schools in the sample were represented by only one classroom), which blurred the distinction between classroom-level and school-level variances, likely underestimating the former. Thus, it is probable that in reality more variance exists at the classroom level. Classroom-level and school-level variances adjusted for student-level variables (Model 1B, 9–10% and 7–11%, respectively) are in the range of those found in other studies (Kyriakides et al., 2002; Luyten, 2003; Reezigt et al., 1999).

Finally, this study found that classroom-level and school-level variables explain about 3.5–4% and 1% of the total variance, which is in the same range as in similar studies. Reezigt et al. (1999) showed that only 1–2% of the total variance was explained by policy-malleable classroom- and school-level variables together, and Kyriakides et al. (2002) found that 3.3% of the total variance was explained by classroom-level variables (including peer effect) and 2.4% of the total variance was explained by school-level variables (including peer effect).

It should be stressed, however, that the small impact of classrooms and schools may be somewhat misleading in this and other school effectiveness studies. This is so for the following two reasons: First, classrooms and schools simultaneously impact all of their students, so their effects are, in a sense, a multiple of the number of students in the classroom and the number of students in the school; second, newer analytical techniques, such as cross-classified random effects model (where the variance in student *growth* in achievement is decomposed), provide considerably larger estimates of teacher effects than variance decomposition models of student achievement that is adjusted for prior achievement (Rowan et al., 2002).

Despite the small amounts of variance attributed to classrooms and schools, this study was nonetheless able to identify the effects of variables representing almost all classroom-level constructs drawn from the theory (to be discussed below), in both mathematics and Serbian language. This is likely a consequence of utilising data from students who have spent all 3 years of their schooling with the same teacher and peers. Furthermore, policy-malleable variables (especially classroom ones) explained considerably more classroom-level and school-level variances than did the peer effect (analyses not shown).

Fixed effects

Classroom-level coefficients from both Model 2 and Model 3 and for both subjects are shown in Table 5. In both mathematics and Serbian language, in Model 2 and Model 3 the baseline achievement is slightly negative and significantly different from zero. Of the many classroom variables shown, no variable had a large association with achievement; however, there were small and statistically significant results.

Table 5. Standardised classroom-level gamma coefficients, Models 2 and 3.

	Mathematics		Serbian language	
	Model 2	Model 3	Model 2	Model 3
Intercept (*SE*)	**−0.038***	**−0.041****	**−0.036**	**−0.035***
	(0.020)	**(0.018)**	**(0.023)**	**(0.021)**
Variable	**Gamma coefficient (*SE*)**			
Teacher Characteristics				
Male teacher	−0.081	−0.055	−0.160**	−0.144**
	(0.054)	(0.054)	(0.068)	(0.069)
Classroom Instruction				
Usefulness of teaching tools	0.035	0.032	0.034	0.040
	(0.026)	(0.025)	(0.025)	(0.025)
Emphasis on basic over complex skills	−0.041*	−0.041**	−0.035	−0.040*
	(0.021)	(0.020)	(0.024)	(0.024)
Clarity and mastery teaching	0.048**	0.047**	0.030	0.030
	(0.024)	(0.023)	(0.024)	(0.024)
Time on Task				
Whole-class instruction	0.056**	0.077**	0.041*	0.047**
	(0.024)	(0.022)	(0.023)	(0.022)
Student Assessment and Feedback				
Reliance on less direct assessment methods to assign a grade	0.042*	0.063**	0.052**	0.059**
	(0.022)	(0.022)	(0.021)	(0.021)
Reliance on social behaviours to assign a grade	−0.032	−0.032	−0.045**	−0.041*
	(0.022)	(0.021)	(0.023)	(0.022)
Teacher's feedback	0.050**	0.042*	0.041	0.038
	(0.025)	(0.025)	(0.027)	(0.027)
Frequency of grading homework	−0.068**	−0.071**	−0.085**	−0.093**
	(0.020)	(0.020)	(0.020)	(0.020)
Student and Teacher Social and Academic Interactions				
Moderate vs. very frequent reinforcement of student effort	0.091**	0.094**	0.052	0.058
	(0.045)	(0.043)	(0.047)	(0.046)
Classroom Climate				
Orderly climate	0.048**	0.041*	0.052**	0.047**
	(0.021)	(0.022)	(0.024)	(0.024)
Control, Demographic, and Contextual variables				
Classrooms with over 20% students eliminated from sample	−0.204**	−0.202**	−0.193**	−0.196**
	(0.085)	(0.084)	(0.085)	(0.086)
Average parental education	0.060**	0.019	0.072**	0.062**
	(0.021)	(0.023)	(0.024)	(0.027)

*p value < 0.10.
**p value < 0.05.

Teacher Characteristics

Teacher gender – the only teacher characteristics variable remaining after preliminary screening out of weak predictors – proved to be a moderate predictor of student achievement in Serbian language. Having a male teacher was negatively associated with student achievement.

Classroom Instruction

Two variables in this construct were associated with student achievement: emphasis on basic versus complex skills (negative association with both subjects) and a

composite of clarity of presentation and mastery teaching, a variable that measures whether the teacher is a clear, organised lecturer who frequently reviews lessons, repeats key points, and checks student comprehension (positive association for mathematics). It should be stressed that clarity/mastery teaching is correlated to the teacher's feedback variable (correlation coefficient is 0.511), so it is argued that this variable is also important (albeit less obviously) in Serbian language. The variable that captured usefulness of various teaching tools to teachers (language games, dictionaries, encyclopaedias and various children's books in Serbian language, and meters, weights, scales, geometrical models, and mathematical games in mathematics) only indicated a trend of positive association with both Serbian language and mathematics achievement. The findings, taken together, resonate well with previous research (Brophy & Good, 1986; Creemers, 1994; Lockheed & Longford, 1989; Mortimore et al., 1988; Reynolds et al., 2002; Rowan et al., 2002; Scheerens, 2000; Walberg & Paik, 2000).

Time on Task

In both mathematics and Serbian language, more whole-class instruction (especially frontal lecturing and less discussion with the whole class) and less individual/group work (especially unsupervised) was significantly associated with higher achievement, confirming previous studies (Brophy & Good, 1986; Rowan et al., 2002). This may be because third graders need the teacher to organise their class time effectively, as they are too young to make themselves to do work during individual or group work or because Serbian teachers are not well trained to plan effective individual or group learning.

Student Assessment and Feedback

A teacher's more frequent feedback to students during oral examinations or on their graded homework was associated with higher scores on mathematics achievement test, as teacher's corrective feedback likely clarifies material for students, points out their strengths and weaknesses, and directs them to study better (Brophy & Good, 1986; Creemers, 1994; Reynolds et al., 2002). Again, given that feedback is considerably positively correlated with the clarity/mastery teaching variable, it is very likely that teacher's feedback is also important for Serbian language achievement.

The frequency of grading homework, as reported by students, showed a negative relationship with achievement. In both mathematics and Serbian language, a teacher who, in order to assign the grade, relied on assessment methods that let the students work on their own (tests, seatwork, homework and individual/group projects, and student teaching) had students with higher achievement. On the contrary, teacher's reliance on student's social behaviours during class (activity during lesson, effort, behaviour, and collaboration with others) in order to assign a grade exhibited a negative association with achievement in Serbian language and a similar but nonsignificant relationship in mathematics.

Student and Teacher Social and Academic Interactions

Teachers who praised their students moderately frequently as opposed to very frequently had, on average, students who were better achievers in mathematics.

The direction of the variable was the same in Serbian language, but its coefficient was not significant. As previous research showed, teachers should praise students moderately frequently and reserve the praise for specific and noteworthy effort, which is thought to boost students' intrinsic motivation to achieve and provide information to students about their competence and progress toward goals (Brophy & Good, 1986; Cotton, 1988).

Classroom Climate

An orderly, disciplined atmosphere where students are attentive during the lesson and where they do not argue or fight with other students was positively associated with student achievement in both subjects, as shown before (Creemers, 1994; Mortimore et al., 1988; OECD, 2001; Opdenakker et al., 2002; Wang et al., 1993; Willms & Somers, 2001).

Control, Demographic, and Contextual variables

In both mathematics and Serbian language, students from classrooms where more than 20% of students either missed achievement testing or were eliminated from the sample due to the lack of information scored considerably lower than the baseline.

Classrooms that had higher average parental education scored higher on the achievement testing in both mathematics and Serbian language in Model 2. In other words, the peer effect was present in Serbian classrooms: A student from the baseline classroom was outperformed on the test by an otherwise identical student from the classroom where the average education of students' parents was higher. (This was more prominent for Serbian language than for mathematics.) This effect disappeared for mathematics when school-level variables were added in Model 3 and was slightly reduced for Serbian language, likely because of the variable's correlation with certain school-level variables, such as school size.

Coefficients of school-level predictors in Model 3 are presented in Table 6 for both mathematics and Serbian language.

Table 6. Standardised school-level gamma coefficients, Model 3.

	Mathematics	**Serbian language**
Intercept	**−0.041****	**−0.035***
(*SE*)	**(0.018)**	**(0.021)**
Variable	**Gamma coefficient (*SE*)**	
School Demographics		
School size	0.080** (0.025)	0.002 (0.026)
School Infrastructure and Services		
Number of computers in school	0.022 (0.026)	0.040* (0.023)
School Culture		
Time principal and teachers spend talking about teaching	−0.031 (0.020)	−0.037** (0.017)
School Climate		
School climate	−0.060** (0.021)	−0.054** (0.024)
Teacher decision-making in school	0.036 (0.023)	0.045* (0.026)

*p value < 0.10.
**p value < 0.05.

School Demographics/School Infrastructure and Services

School size proved to be significant for mathematics but not for Serbian language: This variable was correlated to a degree with another variable present at the school level – the number of computers in school – which was significant only for Serbian language. It is therefore likely that these two variables indicate a type of school (larger and better equipped) that is related to higher student achievement rather than a particular school characteristic. Other school-level variables that were eliminated in preliminary HLM analyses – teaching materials and tools and number of books in the library – are also positively correlated to school size and number of computers in school, reinforcing this conclusion.

School Culture/School Climate

Better school climate showed a significant negative association with student achievement in both mathematics and Serbian language. Higher teacher involvement in decision-making activities in school showed a positive association with Serbian language achievement, while the variable indicating the amount of time that principal and teachers devote to talking about teaching showed a negative association with Serbian language achievement.

Implications for education policy

This study supports the current theory that behaviours and practices in *classrooms* are important to student achievement. However, no clear support is found for the importance of *school* variables that are indicated by theory. This may potentially be attributed to two reasons. First, students in this study spent 3 years with the same teacher and classroom peers, potentially making classroom environment more relevant for students than school environment. Second, school-level variables may be more evenly distributed throughout Serbia because the socialist policies of the past may have managed to equalise school-level variables (facilities, services, administrative decisions, etc.) more than classroom-level factors, which depend on less manageable human resources.

Even though this research identified specific variables that associate with student achievement, this study does not advocate that any individual variable be targeted for improvement. This is so because some of the variables are not easily manipulable by education policy (e.g., peer effect) and some have no clear theoretical justification (e.g., frequency of homework). Finally, several significant classroom-level variables that are policy malleable and mostly in agreement with prior research – clear and structured classroom instruction, greater emphasis on complex over basic skills, more whole-class instruction and less individual/group work, teacher feedback, orderly climate, moderate reinforcement of student effort, and possibly using a variety of teaching tools – are all under the purview of teachers. This indicates that these variables capture some measurable aspects of effective teaching and that they should be understood only as describing some of the principles of good teaching rather than as sole and indisputable prescriptions for good teaching.

Even though the standardised coefficients were generally very small (about 0.05 standard deviations on average), they are additive; so, in theory, if they could all be

simultaneously improved by one standard deviation, achievement of each student in the classroom could increase by as much as 0.30–0.35 standard deviations. Therefore, it is an amalgam of these teaching practices and behaviours and their interconnectedness that education policies should strive to improve. In other words, future reforms should introduce to classrooms in Serbia *active, clear, well-thought-out, and stimulating teaching that engages students intellectually and emotionally* and that is partly captured by the teacher practices and behaviours described above.

In order to accomplish such changes, the following activities should be considered: (a) identification of the most-promising and well-supported strategies for raising the effectiveness of Serbian primary education based on this and other school effectiveness and improvement research; (b) changes in pre- and in-service theoretical and practical coursework to reflect principles and practices of effective teaching; (c) in-depth qualitative research of the classrooms and schools that were shown in this study to have very high or very low achievement scores after adjusting for student background characteristics; and (d) a rigorous pilot study to test promising strategies in Serbian primary schools. Additionally, schools and higher administrative levels should be appropriately included to provide an adequate, learning-oriented climate and facilitate the acceptance, implementation, and modification of education policies aimed at improving student achievement.

Acknowledgements

This study would not have been possible without the enormous effort of the researchers and administrators at the Institute for Education Quality and Evaluation and the Serbian Ministry of Education. This paper was also supported by the Serbian Ministry of Science and Technological Development 2011–2014 projects ''From encouraging initiative, cooperation and creativity in education to new roles and identities in society'' (No. 179034) and ''Improving the quality and accessibility of education in modernization processes in Serbia'' (No. 47008). I am grateful to the Office of Fellowships and Graduate Student Support at the George Washington University for the University Policy Fellowship. I am immensely thankful to Dr Robert Croninger, Dr James Williams, Dr Iris Rotberg, Dr Pasi Sahlberg, and Dr Dylan Conger, who have provided invaluable suggestions for this study, graciously corrected my mistakes, and tremendously supported my efforts in this research. Finally, I will be forever indebted to my mentor, Dr Gregg Jackson, who wholeheartedly advised, taught, and supported me before, during, and after the course of this study.

Note

1. When classroom level was omitted from the fully unconditional model in this study, school-level variance was inflated in comparison to a three-level model (10–13% vs. 8–10%, respectively). In other words, between-school variances in PISA 2003 are probably overestimated.

Notes on contributor

Jelena Teodorović has spent numerous years living, studying, and working in the USA. Upon return to Serbia, she has participated as a researcher, consultant, and advisor in a number of educational evaluations, projects, and activities in the Institute for Educational Research, Serbian Ministry of Education, and several international agencies. She is a happy and tired mother of four.

References

Adair, J.G., Sharpe, D., & Huynh, C.-L. (1989). Hawthorne control procedures in educational experiments: A reconsideration of their use and effectiveness. *Review of Educational Research*, *59*, 215–228.

Baucal, A., Pavlovic-Babic, D., Gvozden, U., & Plut, D. (2006). *Obrazovna postignuca ucenika treceg razreda. Nacionalno testiranje 2004* [Academic achievements of third-grade students. National testing 2004]. Belgrade, Serbia: Zavod za Vrednovanje Kvaliteta Obrazovanja i Vaspitanja. Retrieved from http://www.ceo.edu.rs/images/stories/publikacije/ObrazovnaPostignucaUcenika.pdf

Bloom, B. (1976). *Human characteristics and school learning*. New York, NY: McGraw-Hill.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M.Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York, NY: Macmillan.

Budic, S. (2000). Uticaj individualizovane nastave na uspeh ucenika [The influence of individualised teaching on student achievement]. *Nastava i Vaspitanje*, *XLIX*(1–2), 217–223.

Carroll, J.B. (1963). A model of school learning. *Teachers College Record*, *64*, 723–733.

Cotton, K. (1988). *Instructional reinforcement*. Retrieved from http://educationnorthwest.org/webfm_send/542

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

D'Agostino, J.V. (2000). Instructional and school effects on students' longitudinal reading and mathematics achievements. *School Effectiveness and School Improvement*, *11*, 197–235.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, *8*(1), 1–52.

Dowd, A.J. (2001). Learning requires qualified teachers, good teaching and community support in Malawi. *Dissertation Abstracts International*, *62*(08). (UMI No. 3012916)

Ellett, C.D., Logan, C.S., Claudet, J.G., Loup, K.S., Johnson, B.L., & Chauvin, S.W. (1997). School learning environment, organizational structures and effectiveness: A synthesis of research in 443 schools. *International Journal of Educational Research*, *27*, 273–282.

Farrell, J.P., & Oliveira, J.B. (1993). *Teachers in developing countries: Improving effectiveness and managing costs*. Washington, DC: The World Bank.

Ferguson, R.F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation*, *28*, 465–498.

Fuller, B., & Clarke, P. (1994). Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy. *Review of Educational Research*, *64*, 119–157.

Glewwe, P., Grosh, M., Jacoby, H., & Lockheed, M. (1995). An eclectic approach to estimating the determinants of achievement in Jamaican primary education. *The World Bank Economic Review*, *9*, 231–258.

Goldhaber, D.D., & Brewer, D.J. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *The Journal of Human Resources*, *32*, 505–523.

Hanushek, E.A. (1989). The impact of differential expenditures on school performance. *Educational Researcher*, *18*(4), 45–51, 62.

Hanushek, E.A. (1994). Money might matter somewhere: A response to Hedges, Laine, and Greenwald. *Educational Researcher*, *23*(4), 5–8.

Hanushek, E.A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, *19*, 141–164.

Hedges, L.V., Laine, R., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Education Researcher*, *23*(3), 5–14.

Henderson, A.T., & Mapp, K.L. (2002). *A new wave of evidence – The impact of school, family and community connections on student achievement*. Austin, TX: Southwest Educational Development Laboratory.

Heneveld, W., & Craig, H. (1996). *Schools count: World Bank project designs and the quality of primary education in Sub-Saharan Africa* (World Bank Technical Paper No. 303). Washington, DC: The World Bank.

Heyneman, S.P., & Loxley, W.A. (1982). Influences on academic achievement across high and low income countries – A re-analysis of IEA data. *Sociology of Education*, *55*(1), 13–21.

Hill, P.W., & Rowe, K.J. (1998). Modeling student progress in studies of educational effectiveness. *School Effectiveness and School Improvement*, *9*, 310–333.
International Bureau of Education. (2001). *World data on education*. Geneva, Switzerland: International Bureau of Education.
Kyriakides, L., Campbell, R.J., & 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.
Levin, H.M., & Lockheed, M.E. (1991). *Effective schools in developing countries*. Washington, DC: The World Bank.
Levitas, A., & Herczynski, J. (2006). *The finance and management of primary education in Serbia: Findings and recommendations*. Unpublished report prepared for the Ministry of Education, Republic of Serbia.
Lockheed, M.E., & Longford, N.T. (1989). *A multi-level model of school effectiveness in a developing country* (Policy Research Working Paper No. WPS 242). Washington, DC: The World Bank.
Luyten, H. (2003). The size of school effects compared to teacher effects: An overview of the research literature. *School Effectiveness and School Improvement*, *14*, 31–51.
Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters: The junior years*. Somerset, UK: Open Books.
Muijs, D., & Reynolds, D. (2000). School effectiveness and teacher effectiveness in mathematics: Some preliminary findings from the evaluation of the Mathematics Enhancement Programme (Primary). *School Effectiveness and School Improvement*, *11*, 273–303.
Nyagura, L.M., & Riddell, A. (1993). *Primary school achievement in English and mathematics in Zimbabwe: A multi-level analysis* (Policy Research Working Paper No. WPS 1208). Washington, DC: The World Bank.
Opdenakker, M.-C., & Van Damme, J. (2000a). 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*, 165–196.
Opdenakker, M.-C., & Van Damme, J. (2000b). The importance of identifying levels in multilevel analysis: An illustration of the effects of ignoring the top or intermediate levels in school effectiveness research. *School Effectiveness and School Improvement*, *11*, 103–130.
Opdenakker, M.-C., Van Damme, J., De Fraine, B., Van Landeghem, G., & Onghena, P. (2002). The effect of schools and classes on mathematics achievement. *School Effectiveness and School Improvement*, *13*, 399–427.
Organisation for Economic Co-operation and Development. (2001). *Knowledge and skills for life: First results from PISA 2000*. Paris, France: Author.
Organisation for Economic Co-operation and Development. (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris, France: Author.
Payne, K.J., & Biddle, B.J. (1999). Poor school funding, child poverty, and mathematics achievement. *Educational Researcher*, *28*(6), 4–13.
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*, 193–216.
Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., & Schaffer, E. (2002). *World class schools: International perspectives on school effectiveness*. London, UK: Routledge Falmer.
Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2000). *Teachers, schools, and academic achievement*. Cambridge, MA: National Bureau of Economic Research.
Rowan, B., Correnti, R., & Miller, R.J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the *Prospects* study of elementary schools. *Teachers College Record*, *104*, 1525–1567.
Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (with Smith, A.). (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. London, UK: Open Books.
Sammons, P., Thomas, S., Mortimore, P., Walker, A., Cairns, R., & Bausor, J. (1998). Understanding differences in academic effectiveness: Practitioners' views. *School Effectiveness and School Improvement*, *9*, 286–309.
Sanders, W.L., & Rivers, J.C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.

Savic-Gutesa, D. (2000). Primena visefrontalnog oblika nastave i provera vaspitno-obrazovnih efekata koji se postizu njegovom primenom [Application of multifrontal teaching and evaluation of its effects]. *Nastava i Vaspitanje*, *XLIX* (1–2), 207–216.

Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, *1*, 61–80.

Scheerens, J. (1999). *School effectiveness in developed and developing countries: A review of the research evidence* (Human Development Network Report). Washington, DC: The World Bank.

Scheerens, J. (2000). *Improving school effectiveness* (Fundamentals of Educational Planning No. 68). Paris, France: UNESCO/International Institute for Educational Planning.

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

Teodorović, J. (in press). Student background factors influencing student achievement in Serbia. *Educational Studies*.

United Nations Children's Fund. (2001). *Comprehensive analysis of primary education in the Federal Republic of Yugoslavia*. Belgrade, Serbia and Montenegro: Author.

Van der Werf, G. (1997). Differences in school and instruction characteristics between high-, average-, and low-effective schools. *School Effectiveness and School Improvement*, *8*, 430–448.

Van der Werf, G., Creemers, B., & Guldemond, H. (2001). Improving parental involvement in primary education in Indonesia: Implementation, effects and costs. *School Effectiveness and School Improvement*, *12*, 447–466.

Velez, E., Schiefelbein, E., & Valenzuela, J. (1993). *Factors affecting achievement in primary education*. Washington, DC: The World Bank.

Walberg, H.J., & Paik, S.J. (2000). *Effective educational practices*. Geneva, Switzerland: International Bureau of Education.

Wang, M.C., Haertel, G.D., & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, *63*, 249–294.

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.

Weiss, C. (1998). *Evaluation* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

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