Teodorović, J.; Milin, V.; Bodroža, B.; Đerić, I. D.; Vujačić, M.; Jakšić, I. M.; Stanković, D.; Cankar, G.; Charalambous, C. Y.; Damme, J. V.; Kyriakides, L. Testing the Dynamic Model of Educational Effectiveness: The Impact of Teacher Factors on Interest and Achievement in Mathematics and Biology in Serbia. *School Effectiveness and School Improvement* 2021. https://doi.org/10.1080/09243453.2021.1942076.

Testing the dynamic model of educational effectiveness: the impact of teacher factors on interest and achievement in mathematics and biology in Serbia

SCHOOL EFFECTIVENESS AND SCHOOL IMPROVEMENT

J. TEODOROVIĆ ET AL.

Jelena Teodorović[a] Vladeta Milin[b] Bojana Bodroža[c] Ivana D. Đerić[d] Milja Vujačić[d] Ivana M. Jakšić[e] Dejan Stanković[f] Gašper Cankar[g] Charalambos Y. Charalambous[h] Jan Van Damme[i] Leonidas Kyriakides[h]

[a] Faculty of Education, University of Kragujevac, Jagodina, Serbia

[b] Faculty of Philosophy, University of Belgrade, Belgrade, Serbia

[c] Faculty of Philosophy, University of Novi Sad, Novi Sad, Serbia

[d] Institute for Educational Research, Belgrade, Serbia

[e] Faculty of Political Sciences, University of Belgrade, Belgrade, Serbia

[f] United Nations Children's Fund (UNICEF), Belgrade, Serbia

[g] National Examinations Centre, Ljubljana, Slovenia

[h] Department of Education, University of Cyprus, Nicosia, Cyprus

[i] Center of Educational Effectiveness and Evaluation, KU Leuven, Leuven, Belgium

CONTACT Jelena Teodorović jelenat@gwu.edu

History : received : 2020-5-6 accepted : 2021-6-7

ABSTRACT

Having in mind that student achievement and interest in subject are some of the most important educational goals, and that quality of teaching is the crucial schooling factor influencing them, we examined the contribution of teacher-level variables from the dynamic model of educational effectiveness to student achievement and interest in mathematics and biology. The representative sample included 5,476 students from 125 elementary schools in Serbia and 5,021 parents. Data on student and teacher variables were collected through student and parent questionnaires, while data on prior and current achievement were comprised from students' Trends in International Mathematics and Science Study (TIMSS) 2011 scores and national examination results, respectively. Data were analyzed using multilevel modeling. The results indicate that teacher factors from the dynamic model did not impact student achievement in mathematics and biology but influenced student interest in both subjects. We offer recommendations for educational policy and directions for further research.

KEYWORDS

- SStudent achievement
- interest in subject
- educational effectiveness
- teacher factors
- multilevel modeling

FUNDING

The work presented in this paper was funded by the European Commission's Comenius project "Improving Educational Effectiveness in Primary Schools" (IEEPS), 538992-LLP-1-2013-1-RS-COMENIUS-CMP, within the European Lifelong Learning Programme (LLP), subprogramme Comenius – Comenius Multilateral Projects. The European Commission support for the production of this publication does not constitute an endorsement of the contents, which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. This research was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia (Contracts No. 451-03-9/2021-14/-200140, 451-03-9/2021-14/200018 and 451-03-9/2021-14/200163).

Introduction

Since the 1966 Coleman report (Coleman et al., 1966), student achievement and educational progress have been the focus of thousands of educational effectiveness studies (Creemers & Kyriakides, 2008; Konstantopoulos, 2012; Scheerens, 2016). Achieving higher levels of student knowledge and skills has also become one of the strategic aims of *Europe 2020* (European Commission, 2010) and *Education and Training 2020* (Council of the European Union, 2009). Research in the field has been a conglomerate of versatile studies that could roughly be placed into three major research paradigms: (a) *input-output studies*, focusing on connections between schooling inputs and student outcomes; (b) *effective schools studies*, focusing on connections between school processes and context and student outcomes; and (c) *instructional effectiveness studies*, focusing on connections between teaching and learning processes and student outcomes (Scheerens, 2000). Attention has also been given to the connections between student outcomes and a particular set of educational inputs: socioeconomic and demographic, as well as psychological background characteristics of students and their families.

In the past 2 decades, more integrated studies have appeared that combine more research paradigms (e.g., Charalambous & Kyriakides, 2017; Scheerens, 2016). Some of those have been led by conceptual models of educational effectiveness that combine theory and empirical findings in an effort to paint an all-encompassing picture of ways in which various inputs interact and impact student achievement and/or progress (e.g., Dimosthenous et al., 2020; Praetorius & Charalambous, 2018). However, more empirical studies are needed to confirm the postulates of the conceptual models (Creemers & Kyriakides, 2012); their confirmation would serve both practitioners and researchers, as it would contribute to the generalizability of certain schooling factors.

While student cognitive outcomes (such as student achievement) have been the main focus of educational effectiveness studies, in recent decades it has also become increasingly important to examine student non-cognitive outcomes, such as student motivation, interest in subject, and well-being in school (Creemers, 1994; Opdenakker & Van Damme, 2000; Scheerens & Bosker, 1997; Van Landeghem et al., 2002). This is so because demands on education and goals of education have shifted, so now it is crucial for students to have broad, deep, and interconnected functional knowledge and skills, but also for them to become active citizens, possess digital competencies, speak more than one language, be entrepreneurial, and show a high level of interpersonal skills, among others (Council of the European Union, 2018).

In particular, interest in subject has been shown to play a very important role in learning and academic achievement. Data from a meta-analysis of over 150 studies that examined the relationship between interest and performance show that individual interest was correlated with both academic and laboratory performance (Schiefele et al., 1992). Although interest has usually been considered as an independent variable, an increasing number of studies have seen it as a dependent variable, trying to describe and/or explain the development of interest (Krapp, 1999). These different empirical approaches are related to the conceptualization of the interest that the researcher accepts (Krapp et al., 1992, as cited in Krapp, 1999).

Having in mind that interest is not only a characteristic of the person (as a personality trait), nor only a characteristic of the learning context, this concept should be understood as a specific relationship between the person and their environment, that is, as a "person–object relationship" (Krapp, 1999). Since interest depends on proximal situational factors, it does not refer to some general "academic interest"; rather, it should be seen as a preference for a specific subject (Schiefele et al., 1992). This is also in accordance with the methodological framework in the studies of interest in mathematics (Köller et al., 2001; Viljaranta et al., 2014), interest in science (Hulleman & Harackiewicz, 2009), and other subjects.

In this study, the impact of teacher factors on subject interest is investigated especially since the majority of effectiveness studies were concerned with promoting cognitive rather than affective student learning outcomes. More specifically, our goal was to examine teacher factors contributing to student achievement and interest in mathematics and biology using a well-known theoretical model of educational effectiveness[1]Right click to remove the hyperlink and edit the citation, which gives emphasis to the promotion of new learning goals including the development of positive attitudes towards learning (see Creemers & Kyriakides, 2006).

Theoretical perspectives

The conceptual model that served as a framework for this study was the dynamic model of educational effectiveness by Creemers and Kyriakides (2008). This model was chosen as it was the most recent comprehensive model that encompassed the most relevant theoretical and empirical findings in the field (Scheerens, 2016). The model takes into account a variety of variables known from prior research to impact student achievement (and/or other student outcomes), and analyzes them simultaneously on student, teacher, and/or classroom and school levels. Features of the model are also: (a) focus on practices and processes that could be perceived by students in classrooms and schools (e.g., questioning by teacher or teacher's structuring of lessons) versus characteristics of teachers and schools that may or may not manifest themselves (e.g., teacher working experience or per-student expenditure); (b) assignment of more weight to teacher than to school factors, because they are more proximal to student learning (Creemers & Kyriakides, 2008; Teddlie & Reynolds, 2000); and (c) use of a specific framework to measure the functioning of factors that refers to five different dimensions (i.e., frequency, focus, stage, quality, and differentiation).

The model postulates the following eight teacher-related practices which are considered key for teaching quality: (1) classroom as a learning environment (creating positive interactions in the classroom, order); (2) management of time (organizing an efficient learning environment and maximizing student engagement rates); (3) structuring (outlining the content, calling attention to main ideas, reviewing); (4) orientation (explaining and challenging students to identify why particular lesson activities are taking place); (5) application (providing needed practice through various tasks in class or at home); (6) teaching-modeling (teaching problem solving and other higher order thinking skills); (7) questioning (frequently asking different types of questions and responding to student questions); and (8) assessment of student learning outcomes (especially for formative purposes).[2]Right click to remove the hyperlink and edit the citation

More than 20 empirical studies conducted in different countries provided empirical support to the main assumptions of the model (for a review of these studies, see Kyriakides et al., 2021). Moreover, two quantitative syntheses of effectiveness studies revealed that the teacher-level factors (see Kyriakides et al., 2013) and the school-level factors (see Kyriakides et al., 2010) of the dynamic model are associated with a different type of student learning outcomes. However, it should be acknowledged that only one study (i.e., Kyriakides & Creemers, 2008) was concerned with the promotion of affective student learning outcomes (including subject motivation). Therefore, more studies investigating the impact of teacher factors on affective outcomes are needed especially since the model treats teacher factors as generic in nature meaning that they can have an effect on promoting different types of student learning outcomes.

Undertaking an educational effectiveness study that utilizes a dynamic model in the Serbian context may have two particular advantages. First, many educational effectiveness studies are undertaken in Anglo-Saxon countries and some other countries with a tradition in educational effectiveness research (e.g., the Netherlands, Belgium); therefore, a study placed in a country with a different cultural and historical background may test the transferability of the model and especially the generic nature of teacher practices. The second reason is related to identifying teacher and school practices associated with student learning outcomes (both cognitive and affective). The effects of individual teacher variables have been shown to be small but additive, meaning that good teaching entails many things done well together (Muijs & Reynolds, 2000; Teodorović, 2011). However, the study of teacher effects has been hindered by the lack of suitable data (Reynolds et al., 2014); if a relatively small impact of individual teacher variables is partly attributable to the short period over which their effect is examined (e.g., 1 year) or if studies have been done with younger students (in primary education) where there is less variability in teacher preparation and teaching practices, then the effects may be misleadingly small. Due to the characteristic of the Serbian education system – where students in lower secondary education (from fifth to eighth grade) usually stay with the same subject-matter teacher (i.e., they are usually exposed to only one mathematics teacher in this period) – we were in a position to potentially better detect teacher effects, as they can accumulate over a considerable time period but are not "cross-contaminated" by two or more teachers.

Research questions

In this context, the study aimed to address the following research questions:

- What teacher factors contribute to student achievement in mathematics and biology, after controlling for student background characteristics and prior achievement?
- What teacher factors contribute to student interest in mathematics and biology, after controlling for student background characteristics?

Materials and methods

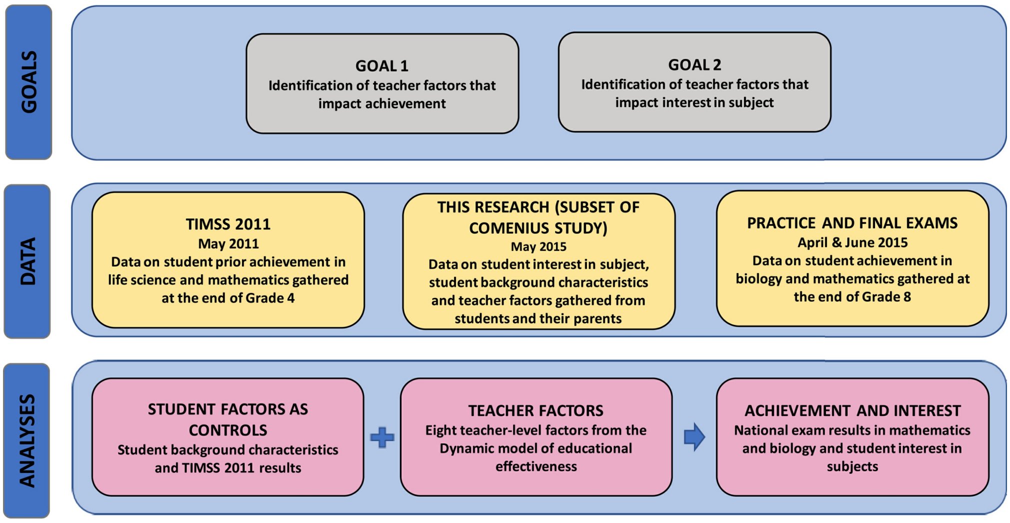
Study design

In order to determine how teaching affects achievement and interest in subject independently from student characteristics, data on various school organizational and teacher factors and a wide set of student background characteristics were collected. The schools' organizational data were provided by teachers, with measures of school variables being obtained by aggregating all teachers' responses. Similarly, the teaching data were provided by students, with measures of teacher factors being obtained by aggregating answers from all students who evaluated a particular teacher's practice. The data about student characteristics were provided by both students and their parents.

Students themselves reported on their interest in mathematics or biology. For the most part, half of the students in the class evaluated the teaching of the mathematics teacher and their own interest in mathematics, while the other half evaluated the teaching of the biology teacher and their own interest in biology.[3]Right click to remove the hyperlink and edit the citation As a measure of student achievement, the results from practice and final national exams that students took at the end of eighth grade were used. These exams included tests in mathematics and biology.[4]Right click to remove the hyperlink and edit the citation The results of the Trends in International Mathematics and Science Study (TIMSS) study in which these students participated in 2011 as fourth graders – specifically, for the mathematics and life science domain of TIMSS 2011 – were used as a measure of prior achievement. Prior achievement is an important predictor of the actual achievement especially since it was not possible to design the study so that the student progress (instead of achievement) could be measured. Analyzing the eighth-grade results while controlling for the fourth-grade results also allowed us to focus on the quality of lower secondary education, where student achievement may start to dwindle (Wijsman et al., 2016).

After collecting the abundance of data, we aimed to utilize multilevel modeling in order to appropriately disentangle and estimate the impact of student background and the impact of teacher practices on achievement and interest in mathematics and biology of lower secondary school students in Serbia (Goldstein, 2003). The basic concept and main steps of the study are presented in Figure 1.

Figure 1. Data sources, analyses, and aims of the study.



Note: Even though we also collected school-level data from the teachers, the figure shows only data that were utilized and analyses that were undertaken in the final, two-level models of student achievement and interest.

Sample

This study was conducted on a nationally representative sample consisting of 125 elementary schools[5]Right click to remove the hyperlink and edit the citation in Serbia which also participated in the TIMSS 2011 international testing. Out of 156 schools that participated in the TIMSS 2011 study, 27 were excluded from this study because fewer than 10 students from these schools participated in TIMSS 2011 research, which was inadequate for the design of our study. Also, four schools declined participation in our research. To a small degree, this disrupted the representativeness of the sample. In 2011, in each selected school usually two classes participated in TIMSS study.

In 115 schools in our sample, two eighth-grade classes were included in this study, while in 10 schools one eighth-grade class per school was included, totaling 240 eighth-grade classes. Out of 240 eighth-grade classes, 208 classes contained students that participated the in TIMSS 2011 study (133 of these were original classes that participated in TIMSS 2011, while from 38 schools which disbanded 53 original TIMSS 2011 classes we selected 78 re-formed classes – two per school – that contained the largest number of TIMSS 2011 students). The remainder of the 240 eighth-grade classes in our sample – 32 classes – did not participate in the TIMSS 2011 study, even though their school did (with only one TIMSS 2011 class), but were included in this study in order to increase representativeness of classes in each school.

The whole sample consisted of 5,423 students from 240 eighth-grade classes, with 3,329 students (61.4%) who also participated in the TIMSS 2011 study, and 2,094 (38.6%) students who did not participate in the TIMSS study. A total of 5,021 parents and 2,477 teachers participated in our research. In the achievement analyses reported in this study, we used only the sample of students who participated in the TIMSS 2011 study.[6]Right click to remove the hyperlink and edit the citation In the mathematics interest analysis, we included students who reported on their interest in mathematics (about half the sample), while in the biology interest analysis, we included students who reported on their interest in biology (another half of the sample).

Variables and instruments

Student variables and instruments

In this study we examined a vast number of student characteristics, recognized in the literature as significant for student achievement: student age, gender, preschool attendance, number of children in the family, single-parent families, number of household members, socioeconomic status (combined information on parental education, their profession, and certain aspects of family resources), parental expectation regarding their child`s education level, student`s spare time reading habits, impulsivity, conscientiousness, and parental involvement (Baucal, 2012; Creemers, 1994; Davis-Kean, 2005; Epstein et al., 2002; Gustafsson et al., 2013; Huntsinger et al., 2000; Marks et al., 2006; Martin & Mullis, 2013; Pavlović Babić & Baucal, 2013; Sirin, 2005; Stoffelsma, 2018; Teodorović, 2011, 2012; Teodorović et al., 2015; Vigil-Coleț & Morales-Vives, 2005; Vujić & Baronijan, 2013; Wagerman & Funder, 2007; Wilder, 2014). Items and scales that comprised student-level variables were taken – sometimes slightly adapted – from the Programme for International Student Assessment (PISA) and TIMSS questionnaires, as well as other studies (Demetriou & Kazi, 2001; Demetriou et al., 2003; Klieme et al., 2009; Köller et al., 2001; Opdenakker & Van Damme, 2000; Teodorović, 2011; Tsukayama et al., 2013). The questionnaires for measuring these variables were validated in a pilot study on 383 students. In the main study, data on these characteristics were gathered through student and parent questionnaires. Prior student achievement was measured according to the student results on the test taken in their fourth grade of primary school within the TIMSS 2011 study. More details on student-level variables are given below.

Gender

The sample included 51.20% girls and 48.20% boys, while 0.60% of students did not answer the question about gender.

Age

We included students from 13 to 15.5 years of age (all at the end of the eighth grade when research was undertaken), with an average value of 14.5 years.

Preschool education attendance.[7]Right click to remove the hyperlink and edit the citation

The students answered the question about preschool attendance with the following answers: "No" (14.60%); "Yes, for 1 year" (25%); or "Yes, more than 1 year" (57.30%).

Number of siblings

The question about the number of siblings was answered by writing down the relevant number. The median of the number of siblings in the included student sample is 1, the mean is 0.9, while the values range from 0 to 14.

Family incompleteness

On the basis of answers about family members they live with, a variable indicating the family incompleteness was constructed. One group consisted of students who live with both parents (97%), while the other group consisted of students living with one or no parents (3%).

SES

Calculating the index of socioeconomic status (SES) of the family was performed using an algorithm developed for the purposes of the PISA study. This SES index consists of three components: parental education, parental occupation, and wealth.[8]Right click to remove the hyperlink and edit the citation Because the SES index was calculated as the first principal component extracted from the three components of SES, the values vary from −2.80 to 2.25, with an average value of 0.

Parental involvement

This scale consisted of seven items and was used for measuring the extent of parental interest and active involvement in the child's school activities (e.g., "Parents help me understand the lesson or explain when something is not clear to me"). Items were measured on a 5-point Likert scale (1 = *never or almost never*; 2 = *rarely*; 3 = *sometimes*; 4 = *often*; 5 = *always or almost always*). Internal consistency of the scale was α = 0.77.

Reading habits

This scale consisted of four items that showed how much a student enjoys reading (e.g., "Reading is one of my favorite hobbies"). Items were measured on a 5-point Likert scale (1 = *I completely disagree*, 2 = *I mostly disagree*, 3 = *I neither agree nor disagree*, 4 = *I mostly agree*, 5 = *I completely agree*). The scale had acceptable internal consistency (α = 0.79).

Impulsivity

This variable was measured with the questionnaire *Domain-Specific Impulsivity* for school-age children (Tsukayama et al., 2013). This dimension of personality is about the impossibility to control impulses in behaviors that are characteristic of the school context. The scale consisted of eight items (e.g., "I interrupt other students when they speak"). Items were measured on a 5-point Likert scale (1 = *almost never*, 2 = *approximately once a month*, 3 = *two to three times a month*, 4 = *approximately once a week*, 5 = *at least once a day*). The scale showed good internal consistency (α = 0.82).

Conscientiousness

This personality trait was measured with a questionnaire consisting of five items (Demetriou & Kazi, 2001; Demetriou et al., 2003). This dimension of personality is about student's commitment, thoughtfulness, tidiness, and ability to plan and timely execute different tasks (e.g., "I am careful and committed in everything I do"). Items were measured on a 5-point Likert scale (1 = *does not apply to me at all*, 2 = *mostly does not apply to me*, 3 = *applies to me to a certain degree*, 4 = *mostly applies to me*, 5 = *completely applies to me*). Internal consistency of the scale was α = 0.75.

TIMSS achievement in the fourth grade

The data of the student achievement on the TIMSS 2011 test were taken from the international database. The statistical control of the student achievement in the fourth grade enabled us to focus on student achievement attained in the period from the fifth to the eighth grade. As a measure for prior achievement in biology, we considered the student's results on the questions referring to the life science domain included in the TIMSS 2011 study. The student's TIMSS 2011 score in mathematics was included as a measure for prior achievement in mathematics.

Teacher variables and instruments

Teacher-level variables were measured by the student questionnaire developed in order to measure the teacher factors of the dynamic model of educational effectiveness (Creemers & Kyriakides, 2008). The following variables were operationalized: classroom as a learning environment, management of time, structuring, orientation, application, teaching-modeling, questioning, and assessment. We were able to assess only the frequency dimension of these variables, and not the other four dimensions from the model (focus, stage, quality, and differentiation) due to study design and organizational and financial restraints.

Some items that comprised teacher-level variables were adapted from previous studies (Baumert et al., 2010; Creemers & Kyriakides, 2015; Fauth et al., 2014; Klieme, 2008; Teodorović, 2011), the DaQS database (https://daqs.fachportal-paedagogik.de/), PISA, and TIMSS, and some were constructed for the purpose of this study. Students indicated how often each teaching practice occurred during the class on a 4-point Likert scale (1 = *never or almost never*, 2 = *rarely*, 3 = *often*, 4

= *always or almost always*). The instrument was first validated in a pilot study with 683731 students[9]Right click to remove the hyperlink and edit the citation and then cross-validated on the data from the main study. Confirmatory factor analysis was carried out on the items belonging to the scale and factor scores were saved and used in the analyses. The factors that emerged were the same as theoretical factors. They are presented below in more detail.

Classroom as a learning environment

This refers to teacher's initiative and encouraging behavior responsible for creating a positive classroom climate (item example: "The teacher tells us why what we will study in class is important"; 9 items, α_{MAT} = 0.89, α_{BIO} = 0.84[10]Right click to remove the hyperlink and edit the citation).

Management of time

This variable consisted of three theoretical aspects: establishing order in the classroom as opposed to disorder (4 items), loss of class time (4 items), and teacher's management of activities in the class (4 items). Establishing order in class refers to teacher's capability to control the student discipline. Loss of class time refers to class time spent on activities that are not related to teaching or learning. Teacher's management of activities in the class refers to teacher's capability to establish clear rules of behavior in class that allow for undisturbed teaching. In this study, we used the general score of management of time, and the higher scores refer to better management of time in the classroom (item example: "The students don't pay attention in class and don't listen to what teacher is saying"; 12 items, α_{MAT} = 0.87, α_{BIO} = 0.87).

Structuring

This variable refers to the practice of creating structure within the lesson, that is, announcement of a particular question or problem within a lesson, announcement of transitioning to another topic, and emphasis and repetition of the most important ideas at the end of the class (item example: "At the beginning of the class, the teacher shortly gives us an overview of what we will do during the lesson"; 9 items, α_{MAT} = 0.78, α_{BIO} = 0.77).

Orientation

This variable was examined based on the frequency of the practice of highlighting the importance and purpose of particular teaching content within a wider context of student knowledge, everyday application, and scientific knowledge (item example: "When I do some activity in class [problems, exercise, experiment], I know why I'm doing it"; 5 items, α_{MAT} = 0.82, α_{BIO} = 0.81).

Application

This variable included the ways of encouraging application of already taught materials, as well as encouraging application of the teaching materials in different situations (item example: "The teacher asks us questions or assigns problems which require us to apply what we learned to similar situations"; 5 items, α_{MAT} = 0.83, α_{BIO} = 0.80).

Teaching-modeling

In our research, this variable referred to teaching students strategies for problem solving, that is, metacognitive strategies such as how to divide lessons into parts for easier learning, how to learn, how to spot important lesson parts, how to find and research different information sources, how to solve and approach problems in different ways, how to analyze tables, charts, and so forth (item example: "The teacher encourages us to find different ways we can use to solve the problems or learn the lesson"; 7 items, α_{MAT} = 0.88, α_{BIO} = 0.89).

Questioning

This variable referred to the frequency of questioning, rephrasing the questions the students did not understand, allowing students time to think about the answer, directing students with subquestions, asking them to explain their answer, and reviewing the extent to which the students understood the lesson (item example: "When we don't understand the question, the teacher asks it in a different way, so that it is more understandable"; 6 items, α_{MAT} = 0.88, α_{BIO} = 0.83).

Assessment

This variable referred to correction of students' answers and giving feedback about how the task should have been done, as well as feedback about specific aspects of students' work (item example: "The teacher tells me what I need to improve in my work in order to be better"; 6 items, α_{MAT} = 0.84, α_{BIO} = 0.82).

Student achievement and interest in subject: measuring scales and scores

The achievement in mathematics was calculated on the basis of the student's results on two tests: (1) the practice national exam in mathematics and (2) the final national exam in mathematics.[11]Right click to remove the hyperlink and edit the citation The analysis used both test results in order to enhance discriminativeness and validity. The two tests contained 20 tasks each, with possible task scores of 0, 0.5, or 1 for each task. While the student achievement on both exams is customarily expressed through the summation scores, for the purpose of this study we applied item response theory (IRT) analysis (i.e., Rasch analysis with partial credit model, two-parameter model) in calculating the scores, respecting the difficulty of tasks and assigning different weights to tasks on the basis of their difficulty. For ease of later interpretation, the IRT scores were transformed into a scale varying from 0 to 20 (as was the original scale of the test), with higher scores indicating higher achievement. ~~With this approach, we obtained an average student achievement in mathematics of 8.83 points (*SD* = 3.30), which means that the test was rather difficult.~~

The achievement in biology was measured through results of the practice exam (five questions) and the final exam (five questions).[12]Right click to remove the hyperlink and edit the citation IRT scores were also calculated and transformed into a scale varying from 0 to 5, with higher scores indicating higher achievement.

Student interest in mathematics and biology was measured through seven items (e.g., "I enjoy learning mathematics/biology"), four of which were taken and adapted from the PISA study, while the rest were author developed. As with the teacher factors, half of the students in the classroom reported on their interest in mathematics while the other half reported on their interest in biology. Internal consistency for both interest in mathematics and interest in biology was α = 0.74. The students responded on a 4-point Likert scale (1 = *I completely disagree* to 4 = *I completely agree*). The final scores were calculated as the average of all the items, with the higher number standing for higher interest for the subject.

Data collection

The research was conducted with help from school employees – the school coordinators – who were in charge of the research implementation in the school. The data collection lasted from April until June 2015. The students', parents', and teachers' data were collected during April and May 2015. The response rates were excellent: Out of 5,538 students in the planned sample, 5,423 students participated in the study (97.92%); out of 5,538 parents in the planned sample, 5,021 parents participated in the study (90.66%); and out of 2,500 teachers in the planned sample, 2,401 teachers participated in the study (96.04%). The data from the practice national exam organized at schools were collected in April 2015, while in June the students took the final exam. Linking the data from different phases and different sources (teachers, parents, and students) was enabled by a coding system. All research participants were guaranteed anonymity with questionnaires sent by school coordinators to the research team without personal information, only with a code, within closed envelopes. The student achievement in the TIMSS 2011 international testing was provided with the consent of the Ministry of Education, Science and Technological Development.

Data analysis

Due to the nested structure of the data (students within teachers within schools), multilevel analysis of the data was necessary to assess the effects of variables on student achievement and interest in subject. Therefore, we used a model with two levels – student nested within teacher level.[13]Right click to remove the hyperlink and edit the citation The teacher level was chosen instead of the school level, because prior research gives more weight to the teacher factors over school factors (Kyriakides et al., 2013) and because of the focus of this study on teaching quality and its effects on student learning.

Multilevel analysis was carried out stepwise, starting with an empty model (without explanatory variables) in which the variance of the outcome variable was partitioned into the student-level and teacher-level variances. For all the subsequent models, all continuous variables were grand-mean centered. Also, all student background factors and teacher factors in all multilevel analyses were transformed into z scores.

The next step was including student background variables in the model, to control for the student intake and avoid an overestimation of the effect of teacher-level variables. Then, we entered the teacher-level variables. Our preliminary analyses indicated moderate correlations between teacher-level variables, which often led to mutual exclusion of the predictors when they were entered and tested together in the multilevel model. Because of the possibility of multicollinearity, it was decided to add teacher-level variables in separate multilevel models to test how each one predicts the outcome variables beyond the student background variables already in the model. We used a random intercept model rather than a random slope model, meaning that the slopes for the teacher-level variables were not allowed to vary.

Finally, to examine how much variance of the outcome variable would be explained by all teacher factors together, but to avoid problems of multicollinearity, we decided to reduce the number of teacher factors by using exploratory factor analysis (EFA). The results of EFA indicated that all teacher factors converge into one higher order factor which we named *teaching quality* (see Appendix 1). This factor was used as a predictor in a separate multilevel model to examine the overall contribution of teaching to the student outcome variables beyond student background variables.

The described procedure was conducted separately for each of the four outcome variables (i.e., dependent variables of this study): achievement in mathematics, interest in mathematics, achievement in biology, and interest in biology.

Results

Descriptive statistics for all student- and teacher-level variables are given in Table 1. Correlations between the variables are presented in the supplementary material (see https://osf.io/dyq6s/).

Table 1. Descriptive statistics for student-level and teacher-level variables.

Variable	*N*	Min	Max	*M*	*SD*	Skewness	Kurtosis
Parental involvement	5,423	−5.43	1.16	0.00	1.00	−1.25	2.13
Reading habits	5,423	−2.04	1.72	0.00	1.00	0.01	−0.96
Impulsiveness	5,422	−1.31	3.42	0.00	1.00	0.87	0.28
Conscientiousness	5,423	−3.54	1.58	0.00	1.00	−0.48	−0.19
SES	5,413	−2.80	2.25	0.02	1.05	0.30	−0.73
TIMSS achievement in mathematics	3,329	182.11	772.94	522.04	80.86	−0.33	0.21
TIMSS achievement in biology	3,329	186.76	803.70	523.63	71.26	−0.29	0.25
Student achievement in mathematics	5,193	0.00	20.00	9.30	3.09	0.24	0.02
Student achievement in biology	5,193	0.00	5.00	3.69	0.84	−0.39	0.05
Interest in mathematics	2,527	1.00	4.00	2.70	0.73	−0.26	−0.58
Interest in biology	2,895	1.00	4.00	2.33	0.78	0.11	−0.76
M – Classroom as a learning environment	2,895	−3.01	1.55	0.00	1.00	−0.61	−0.21
M – Management of time	2,895	−3.20	1.69	0.00	1.00	−0.51	−0.24
M – Structuring	2,895	−3.28	1.96	0.00	1.00	−0.46	−0.04
M – Orientation	2,895	−2.66	1.60	0.00	1.00	−0.37	−0.53
M – Application	2,895	−2.47	1.68	0.00	1.00	−0.50	−0.12
M – Teaching-Modeling	2,895	−2.44	1.73	0.00	1.00	−0.44	−0.32
M – Questioning	2,895	−2.90	1.49	0.00	1.00	−0.56	−0.25
M – Assessment	2,895	−2.16	1.82	0.00	1.00	−0.25	−0.66

M – Teaching quality	2,895	−3.16	2.07	0.00	1.00	−0.45	−0.21
B – Classroom as a learning environment	2,526	−3.01	1.55	0.00	1.00	−0.59	−0.13
B – Management of time	2,527	−3.20	1.69	0.00	1.00	−0.49	−0.31
B – Structuring	2,526	−3.28	1.96	0.00	1.00	−0.51	−0.04
B – Orientation	2,527	−2.66	1.60	0.00	1.00	−0.51	−0.31
B – Application	2,527	−2.47	1.68	0.00	1.00	−0.19	−0.59
B – Teaching-Modeling	2,526	−2.44	1.73	0.00	1.00	−0.25	−0.70
B – Questioning	2,526	−2.90	1.49	0.00	1.00	−0.58	−0.09
B – Assessment	2,527	−2.16	1.82	0.00	1.00	−0.11	−0.83
B – Teaching quality	2,238	−3.32	2.07	0.00	1.00	−0.33	−0.41

Note: M = Mathematics; B = Biology.

Contribution of student characteristics and teacher factors to student achievement in mathematics

By testing an empty model, it was determined that 80.94% of the variance in mathematics achievement was situated at the student level, while 19.06% was due to the differences between teachers (Table 2). The following student and their family characteristics explained 45.94% of the total differences in the student achievement in mathematics (Table 2): *gender*, *socioeconomic status of the family*, *conscientiousness*, *impulsivity*, *reading habits*, and *prior achievement* (the achievement in mathematics in the TIMSS 2011 study).[14]Right click to remove the hyperlink and edit the citation When controlling the mentioned student characteristics, there remained 44.44% of differences at the student level that could not be explained by factors included in this study, and 9.62% of unexplained differences in student achievement at the teaching level (Table 2).

Table 2. Parameter estimates and standard errors for analyses of achievement in mathematics.

	Empty model			Basic model			Model 1			Model 2			Model 3			Model 4		
	B	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*
Intercept	8.67	0.14	***	8.61	0.11	***	8.61	0.11	***	8.61	0.11	***	8.61	0.11	***	8.61	0.11	***
STUDENT LEVEL																		
Gender (0 – male, 1 – female)				0.25	0.09	**	0.25	0.09	**	0.25	0.09	**	0.25	0.09	**	0.25	0.09	**
SES				0.64	0.05	***	0.64	0.05	***	0.64	0.05	***	0.64	0.05	***	0.64	0.05	***
TIMSS				1.81	0.05	***	1.81	0.05	***	1.81	0.05	***	1.81	0.05	***	1.81	0.05	***
Reading habits				0.24	0.05	***	0.24	0.05	***	0.24	0.05	***	0.23	0.05	***	0.24	0.05	***
Impulsivity				−0.16	0.05	**	−0.16	0.05	**	−0.16	0.05	**	−0.16	0.05	**	−0.16	0.05	**
Conscientiousness				0.21	0.05	***	0.21	0.05	***	0.21	0.05	***	0.21	0.05	***	0.21	0.05	***
TEACHING LEVEL																		
Learning environment							−0.04	0.09	n.s.									
Management of time										0.10	0.09	n.s.						
Structuring													0.04	0.09	n.s.			
Orientation																0.02	0.09	n.s.
Assessment																		
Teaching-modeling																		
Application																		
Questioning																		

TEACHING QUALITY																		
Unexplained variance (%)																		
Teaching level	19.06			9.62			9.61			9.51			9.59			9.61		
Student level	80.94			44.44			44.44			44.44			44.44			44.44		
Explained variance (%)																		
Total				45.94			45.95			46.05			45.97			45.95		
Introduced variables				45.94			0.01			0.11			0.03			0.01		
Chi-square	14415.43			12709.59			12709.44			12708.51			12709.43			12709.56		
ΔChi-square				1705.83			0.15			1.09			0.16			0.03		
Δ*df*				6			1			1			1			1		
p				***			n.s.			n.s.			n.s.			n.s.		

	Model 5			Model 6			Model 7			Model 8			Model 9		
	B	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*
Intercept	8.62	0.11	***	8.61	0.11	***	8.61	0.11	***	8.61	0.11	***	8.61	0.11	***
STUDENT LEVEL															
Gender (0 – male, 1 – female)	0.25	0.09	**	0.25	0.09	**	0.25	0.09	**	0.25	0.09	**	0.25	0.09	**
SES	0.64	0.05	***	0.64	0.05	***	0.64	0.05	***	0.64	0.05	***	0.64	0.05	***
TIMSS	1.81	0.05	***	1.81	0.05	***	1.81	0.05	***	1.81	0.05	***	1.81	0.05	***
Reading habits	0.24	0.05	***	0.24	0.05	***	0.24	0.05	***	0.24	0.05	***	0.24	0.05	***
Impulsivity	−0.16	0.05	**	−0.16	0.05	**	−0.16	0.05	**	−0.16	0.05	**	−0.16	0.05	**
Conscientiousness	0.21	0.05	***	0.21	0.05	***	0.21	0.05	***	0.21	0.05	***	0.21	0.05	***
TEACHING LEVEL															
Learning environment															
Management of time															
Structuring															
Orientation															
Assessment	−0.07	0.09	n.s.												
Teaching-modeling				0.01	0.09	n.s.									
Application							0.00	0.09	n.s.						
Questioning										0.00	0.09	n.s.			
TEACHING QUALITY													0.01	0.09	n.s.
Unexplained variance (%)															
Teaching level	9.59			9.61			9.62			9.62			9.61		
Student level	44.43			44.44			44.44			44.44			44.44		
Explained variance (%)															
Total	45.98			45.95			45.94			45.94			45.95		
Introduced variables	0.04			0.01			0.00			0.00			0.01		
Chi-square	12708.96			12709.58			12709.59			12709.59			12709.59		

ΔChi-square	0.64			0.01			0.00			0.00			0.00		
Δ*df*	1			1			1			1			1		
p	n.s.			n.s.			n.s.			n.s.			n.s.		

N = 2,829

**$p < .01$. ***$p < .001$.

However, when the teacher factors were introduced in the hierarchical linear model, none of them had a significant impact on the achievement in mathematics (Table 2). It should be reiterated that different aspects of teaching are mutually correlated to a certain degree, which is why each of the teacher factors was analyzed in a separate hierarchical linear model. We also tested a general factor *teaching quality* (see Appendix 1), which we constructed to surpass the multicollinearity of the individual teacher factors. However, even this general factor of teaching quality did not contribute significantly to student achievement in mathematics.

Contribution of student characteristics and teacher factors to student interest in mathematics

The analysis of the empty model for student interest in mathematics showed that 84.72% of the variance can be assigned to differences between students and 15.65% of the variance can be assigned to differences between teachers (Table 3). Among the various student characteristics, a statistically significant impact was shown by: *socioeconomic status of the family*, *reading habits*, *impulsivity*, *and conscientiousness* in said subject, with individual student characteristics explaining 9.91% of the total differences in student interest (Table 3). When these individual differences between students are controlled for, 76.11% of student interest in mathematics at the student level and 13.98% of student interest in mathematics at the teacher level are left unexplained (Table 3).

Table 3. Parameter estimates and standard errors for analyses of interest in mathematics.

	Empty model			Basic model			Model 1			Model 2			Model 3			Model 4		
	B	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*
Intercept	2.78	0.04	***	2.78	0.04	***	2.78	0.03	***	2.77	0.03	***	2.78	0.03	***	2.78	0.03	***
STUDENT LEVEL																		
SES				0.05	0.02	*	0.04	0.02	n.s.	0.05	0.02	*	0.04	0.02	*	0.05	0.02	**
Reading habits				0.11	0.02	***	0.12	0.02	***	0.11	0.02	***	0.11	0.02	***	0.11	0.02	***
Impulsivity				−0.18	0.02	***	−0.18	0.02	***	−0.18	0.02	***	−0.18	0.02	***	−0.18	0.02	***
Conscientiousness				0.12	0.02	***	0.12	0.02	***	0.12	0.02	***	0.12	0.02	***	0.11	0.02	***
TEACHING LEVEL																		
Learning environment							0.28	0.03	***									
Management of time										0.21	0.03	***						
Structuring													0.28	0.03	***			
Orientation																0.32	0.03	***
Assessment																		
Teaching-modeling																		
Application																		
Questioning																		
TEACHING QUALITY																		
Unexplained variance (%)																		
Teaching level	15.65			13.98			6.48			10.00			6.85			4.72		
Student level	84.72			76.11			76.11			76.11			76.11			76.02		
Explained variance (%)																		
Total				9.91			17.41			13.89			17.04			19.26		

Introduced variables				9.91			7.50			3.98			7.13			9.35		
Chi-square	7294.24			6985.72			6905.73			6948.17			6909.43			6875.08		
ΔChi-square				308.52			79.99			37.55			76.29			110.64		
Δdf				4			1			1			1			1		
p				***			***			***			***			***		

	Model 5			Model 6			Model 7			Model 8			Model 9		
	B	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*
Intercept	2.77	0.03	***	2.78	0.03	***	2.78	0.03	***	2.78	0.03	***	2.78	0.03	***
STUDENT LEVEL															
SES	0.05	0.02	*	0.05	0.02	*	0.04	0.02	*	0.04	0.02	*	0.04	0.02	*
Reading habits	0.11	0.02	***	0.12	0.02	***	0.12	0.02	***	0.11	0.02	***	0.11	0.02	***
Impulsivity	−0.18	0.02	***	−0.18	0.02	***	−0.18	0.02	***	−0.18	0.02	***	−0.18	0.02	***
Conscientiousness	0.11	0.02	***	0.12	0.02	***	0.12	0.02	***	0.12	0.02	***	0.12	0.02	***
TEACHING LEVEL															
Learning environment															
Management of time															
Structuring															
Orientation															
Assessment	0.30	0.03	***												
Teaching-modeling				0.30	0.03	***									
Application							0.27	0.03	***						
Questioning										0.30	0.03	***			
TEACHING QUALITY													0.31	0.03	***
Unexplained variance (%)															
Teaching level	5.65			5.37			6.85			5.83			5.19		
Student level	76.02			76.11			76.20			76.11			76.11		
Explained variance (%)															
Total	18.33			18.52			16.94			18.06			18.70		
Introduced variables	8.42			8.61			7.03			8.15			8.79		
Chi-square	6890.47			6890.23			6912.72			6894.73			6884.95		
ΔChi-square	95.26			95.49			73.00			90.99			100.77		
Δdf	1			1			1			1			1		
p	***			***			***			***			***		

N = 2,575

**p* < .05. ***p* < .01. ****p* < .001.

The analysis of the contribution of individual teacher factors to student interest in mathematics showed that all eight measured teaching-related factors have a statistically significant impact at the .05 level. These factors individually explain from 3.98% to 9.35% of the differences in interest in mathematics (Table 3). In order to determine the total contribution of intercorrelated teacher variables to student interest in mathematics, the impact of the general factor – *teaching quality* – was tested. This factor explains 8.79% of the differences in student interest out of 13.98% that can be assigned to the impact of teaching while 5.19% of teacher-level variance remains unexplained (Table 3).

Contribution of student characteristics and teacher factors to student achievement in biology

The analysis of the empty model for achievement in biology showed that 78.47% of the differences in student achievement in biology comes from differences between the students, while 21.53% of the differences is situated between their biology teachers (Table 4). Of the student characteristics, the significant contribution to achievement in biology can be attributed to: *gender, socioeconomic status, prior achievement* (the achievement in biology in the TIMSS 2011 study), *reading habits*, and *conscientiousness*, which together explained 10.48% of the total variance of student achievement (Table 4)**.** After levelling the individual differences between students, 69.83% of the differences at the student level was left unexplained and 19.69% of the differences in student achievement in biology at the teacher level was left unexplained (Table 4).

Table 4. Parameter estimates and standard errors for analyses of achievement in biology.

	Empty model			Basic model			Model 1			Model 2			Model 3			Model 4		
	B	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*
Intercept	3.69	0.04	***	3.64	0.04	***	3.64	0.04	***	3.64	0.04	***	3.64	0.04	***	3.64	0.04	***
STUDENT LEVEL																		
Gender (0 – male, 1 – female)				0.11	0.03	***	0.11	0.03	***	0.11	0.03	***	0.11	0.03	***	0.11	0.03	***
SES				0.10	0.02	***	0.10	0.02	***	0.10	0.02	***	0.10	0.02	***	0.10	0.02	***
TIMSS				0.17	0.02	***	0.17	0.02	***	0.17	0.02	***	0.17	0.02	***	0.17	0.02	***
Reading habits				0.07	0.02	***	0.07	0.02	***	0.07	0.02	***	0.07	0.02	***	0.07	0.02	***
Conscientiousness				0.06	0.02	***	0.06	0.02	***	0.06	0.02	***	0.06	0.02	***	0.06	0.02	***
TEACHING LEVEL																		
Learning environment							0.00	0.04	n.s.									
Management of time										0.02	0.04	n.s.						
Structuring													0.03	0.04	n.s.			
Orientation																0.03	0.04	n.s.
Assessment																		
Teaching-modeling																		
Application																		
Questioning																		
TEACHING QUALITY																		
Unexplained variance (%)																		
Teaching level	21.53			19.69			19.69			19.69			19.69			19.55		
Student level	78.47			69.83			69.83			69.83			69.83			69.83		
Explained variance (%)																		
Total				10.48			10.48			10.48			10.48			10.62		
Introduced variables				10.48			0.00			0.00			0.00			0.14		
Chi-square	6130.81			5827.81			5827.81			5827.55			5827.37			5826.96		
ΔChi-square				303.00			0.00			0.26			0.44			0.85		
Δdf				5			1			1			1			1		
p				***			n.s.			n.s.			n.s.			n.s.		

	Model 5			Model 6			Model 7			Model 8			Model 9		
	B	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*

Intercept	3.64	0.04	***	3.64	0.04	***	3.64	0.04	***	3.64	0.04	***	3.64	0.04	***
STUDENT LEVEL															
Gender (0 – male, 1 – female)	0.11	0.03	***	0.11	0.03	***	0.11	0.03	***	0.11	0.03	***	0.11	0.03	***
SES	0.10	0.02	***	0.10	0.02	***	0.10	0.02	***	0.10	0.02	***	0.10	0.02	***
TIMSS	0.17	0.02	***	0.17	0.02	***	0.17	0.02	***	0.17	0.02	***	0.17	0.02	***
Reading habits	0.07	0.02	***	0.07	0.02	***	0.07	0.02	***	0.07	0.02	***	0.07	0.02	***
Conscientiousness	0.06	0.02	***	0.06	0.02	***	0.06	0.02	***	0.06	0.02	***	0.06	0.02	***
TEACHING LEVEL															
Learning environment															
Management of time															
Structuring															
Orientation															
Assessment	0.01	0.04	n.s.												
Teaching-modeling				0.04	0.04	n.s.									
Application							0.05	0.04	n.s.						
Questioning										0.02	0.04	n.s.			
TEACHING QUALITY													0.03	0.04	n.s.
Unexplained variance (%)															
Teaching level	19.69			19.41			19.26			19.69			19.55		
Student level	69.83			69.83			69.83			69.83			69.83		
Explained variance (%)															
Total	10.48			10.76			10.91			10.48			10.62		
Introduced variables	0.00			0.28			0.42			0.00			0.14		
Chi-square	5827.66			5826.47			5825.73			5827.62			5827.23		
ΔChi-square	0.15			1.34			2.08			0.19			0.58		
Δ*df*	1			1			1			1			1		
p	n.s.			n.s.			n.s.			n.s.			n.s.		

N = 2,635

**p < .01. ***p < .001.

Out of eight examined teacher factors, none had a statistically significant impact on the achievement in biology (Table 4). The general factor – *teaching quality* – also did not have a statistically significant effect on student achievement in biology.

Contribution of student characteristics and teacher factors to student interest in biology

The analyses show that 87.41% of the differences between students in their interest in biology are individual differences between students within a class, while 12.59% of the differences in interest is linked to the fact that they attend different classes with different teachers (Table 5). Among the examined student characteristics, the following are statistically significant: *reading habits*, *impulsivity*, *conscientiousness*, and *parental involvement*. These characteristics of students and their family environments explain a total of 13.93% of the differences between the students in their interest in biology (Table 5). After the students were levelled by the relevant individual characteristics, 75.23% of unexplained variance was left at the student level and 10.84% of unexplained variance was left at the teacher level (Table 5).

Table 5. Parameter estimates and standard errors for analyses of interest in biology.

	Empty model			Basic model			Model 1			Model 2			Model 3			Model 4		
	B	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*
Intercept	3.26	0.04	***	3.27	0.03	***	3.27	0.02	***	3.27	0.03	***	3.27	0.03	***	3.27	0.02	***
STUDENT LEVEL																		

Reading habits				0.16	0.02	***	0.16	0.02	***	0.16	0.02	***	0.16	0.02	***	0.16	0.02	***
Parental involvement				0.06	0.02	**	0.05	0.02	**	0.06	0.02	**	0.05	0.02	**	0.05	0.02	**
Impulsivity				−0.14	0.02	***	−0.14	0.02	***	−0.13	0.02	***	−0.14	0.02	***	−0.14	0.02	***
Conscientiousness				0.16	0.02	***	0.15	0.02	***	0.16	0.02	***	0.15	0.02	***	0.15	0.02	***
TEACHING LEVEL																		
Learning environment							0.27	0.02	***									
Management of time										0.20	0.03	***						
Structuring													0.26	0.02	***			
Orientation																0.27	0.02	***
Assessment																		
Teaching-modeling																		
Application																		
Questioning																		
TEACHING QUALITY																		
Unexplained variance (%)																		
Teaching level	12.59			10.84			3.20			6.81			3.61			3.10		
Student level	87.41			75.23			75.23			75.23			75.23			75.13		
Explained variance (%)																		
Total				13.93			21.57			17.96			21.16			21.78		
Introduced variables				13.93			7.64			4.02			7.22			7.84		
Chi-square	6224.63			5885.74			5795.92			5845.01			5801.61			5791.49		
ΔChi-square				338.89			89.82			40.73			84.14			94.26		
Δ*df*				4			1			1			1			1		
p				***			***			***			***			***		

	Model 5			Model 6			Model 7			Model 8			Model 9		
	B	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*	*B*	*SE*	*p*
Intercept	3.27	0.02	***	3.27	0.02	***	3.27	0.03	***	3.27	0.03	***	3.27	0.02	***
STUDENT LEVEL															
Reading habits	0.16	0.02	***	0.16	0.02	***	0.16	0.02	***	0.16	0.02	***	0.16	0.02	***
Parental involvement	0.05	0.02	**	0.05	0.02	**	0.05	0.02	**	0.05	0.02	**	0.05	0.02	**
Impulsivity	−0.14	0.02	***	−0.14	0.02	***	−0.14	0.02	***	−0.14	0.02	***	−0.14	0.02	***
Conscientiousness	0.15	0.02	***	0.15	0.02	***	0.16	0.02	***	0.15	0.02	***	0.15	0.02	***
TEACHING LEVEL															
Learning environment															
Management of time															
Structuring															
Orientation															

Assessment	0.27	0.02	***												
Teaching-modeling				0.27	0.02	***									
Application							0.23	0.03	***						
Questioning										0.26	0.03	***			
TEACHING QUALITY													0.28	0.02	***
Unexplained variance (%)															
Teaching level	3.30			3.10			4.95			3.72			2.58		
Student level	75.13			75.23			75.23			75.23			75.13		
Explained variance (%)															
Total	21.57			21.67			19.81			21.05			22.29		
Introduced variables	7.64			7.74			5.88			7.12			8.36		
Chi-square	5794.38			5792.91			5822.05			5803.47			5782.72		
ΔChi-square	91.36			92.84			63.69			82.28			103.02		
Δ*df*	1			1			1			1			1		
p	***			***			***			***			***		

N = 2,270.

$^{**}p < .01$. $^{***}p < .001$.

The analyses of examined teacher factors showed that out of eight factors measured in this study, all factors significantly predict students' interest in biology (Table 5). The general factor, *teaching quality*, which encompasses all examined teacher factors, is responsible for 8.36% of the differences in student interest in biology, while only 2.58% of the teaching-related differences are left unexplained (Table 5).

Additional analyses

Considering that in the models for achievement we controlled for prior achievement, while in the models for interest in subjects we could not control for prior interest, we carried out additional multilevel analyses. In the models with achievement, we removed prior achievement from the basic model to make it comparable with the model for interest in subject. In the model with interest in subject as an outcome variable, we added current achievement[15]Right click to remove the hyperlink and edit the citation as a student-level control in the basic model. After these changes, we repeated the analyses with teacher variables in separate models as was described. The results of these analyses are presented in the supplementary material (https://osf.io/dyq6s/). However, even with these changes in the basic models for achievement and interest in subjects, the results regarding the teacher factors remained the same – none of the teacher factors were statistically significant predictors of achievement in mathematics or biology, while all teacher factors were statistically significant predictors of interest in both subjects.

Results summary

In our study, for achievement, 80.94% of the variance in mathematics and 78.47% of the variance in biology lie at the student level; for interest in a subject, 84.72% of the variance in mathematics and 87.41% of the variance in biology are situated amongst the students.

After the control of student-level factors, no teacher factor was found to be significant for achievement in mathematics nor for achievement in biology.

The largest impact on interest in mathematics (arbitrarily taken to be above 5% of the explained variance) comes from the following teacher-level variables (in descending order): (1) *orientation* (9.35% explained variance); (2) *teaching-modeling* (8.61%); (3) *assessment* (8.42%); (4) *questioning* (8.15%); (5) *classroom as a learning environment* (7.50%); (6) *structuring* (7.13%); and (7) *application* (7.03%). Importantly, the same seven teacher-level variables exhibit the largest effect on interest in biology, albeit in a somewhat different order: (1) *orientation* (7.84%); (2) *teaching-modeling* (7.74%); (3&4) *classroom as a learning environment* (7.64%); (3&4) *assessment* (7.64%); (5) *structuring* (7.22%); (6) *questioning* (7.12%); and (7) *application* (5.88%). In both subjects, *management of time* explains less variance (3.98% in mathematics and 4.02% in biology).

The overall factor measuring teaching quality explains 8.97% of the variance in interest in mathematics, leaving 5.19% of unexplained variance at the teacher level, and 8.36% of the variance in interest in biology, leaving only 2.58% of unexplained variance at the teacher level. Considering that the teacher-level variables in our research explain around 63% of the available variance in interest in mathematics (8.79% out of 13.98% at the teacher level) and around 77% of the available variance in interest in biology (8.36% out of 10.84% at the teacher level), our study strongly supports the dynamic model of educational effectiveness for explaining variation in student affective learning outcomes in each subject (i.e., interest in mathematics and biology).

Discussion

The prevalence of student-level variance in studies like ours should come as no surprise, since Serbia, like many other middle- and high-income countries, has a relatively homogeneous and sufficient provision of school infrastructure, teaching materials, and a qualified teacher workforce, which account for smaller variability at levels beyond the student level (Scheerens, 1990, 1999). Also, such findings are common in educational effectiveness studies that focus on status rather than progress of students (Scheerens, 2016). Focusing on student gains (value-added models) allows one to get a better estimate of teacher and school effects; those studies show considerably larger effects of teaching and schooling (Scheerens, 2016) but require utilization of more sophisticated measures of dependent variables. Nonetheless, even in status studies, small teacher-level and/or school-level variances translate into effects of great practical relevance[16]Right click to remove the hyperlink and edit the citation (Scheerens, 2016). The results of our study in terms of variance decomposition fit well with two other studies of educational effectiveness in Serbia. A somewhat larger student-level variance was reported for third-grade achievement, where 86.70% of the variance in mathematics and 83.90% of the variance in the Serbian language were found at the student level (Teodorović, 2011), and for fourth-grade achievement, where 85.2% of the variance in mathematics achievement was found at the student level (Jakšić et al., 2017). In comparison to studies in some other countries, student-level variance in achievement in this study is larger (e.g., Kyriakides & Luyten, 2009, as cited in Scheerens, 2016; Opdenakker & Van Damme, 2000). However, two things should be noted: (1) Student-level variance is smaller and school-level and/or teacher-level variance is larger in studies where the student sample comes from schools with differentiated tracks, and (2) lower secondary education in Serbia resembles primary education more than upper secondary education, as students mostly attend the same comprehensive school from Grade 1, for a large part take the same classes with the same group of 30 students from Grade 1, and mostly have the same set of subject teachers for Grades 5 to 8, as well as follow the same, national curriculum.

In terms of differences in variance decomposition between achievement and interest models, teacher-level variance was bigger on cognitive than on affective outcomes in both mathematics and biology. This is line with other studies (see Creemers & Kyriakides, 2008; Kyriakides et al., 2021) and probably reflects more time invested by teachers and schools in pursuing cognitive outcomes as well as a larger influence of other social organizations on affective outcomes (Creemers & Kyriakides, 2008).

Several student characteristics showed a significant impact on achievement and/or interest in subject, fitting very well with current literature (Köller et al., 2001; Schiefele et al., 1992). Two student-level variables significantly predicted all four outcome variables, in expected directions: reading habits and conscientiousness. SES was a significant factor in all outcomes, but interest in biology and impulsivity was a significant factor in all outcomes but biology achievement. Gender[17]Right click to remove the hyperlink and edit the citation and prior achievement appear to play an important role in both mathematics and biology achievement. Finally, parental involvement showed an effect on interest in biology. Student-level variables from this study are discussed in more detail in another paper.

When it comes to teacher factors, we found no evidence of impact of teacher-level factors from the dynamic model of educational effectiveness on either mathematics or biology achievement. This study, therefore, cannot support the dynamic model of educational effectiveness for mathematics and biology achievement and is at odds with a large body of evidence on the impact of these factors on student achievement (Antoniou et al., 2011; Brophy & Good, 1986; Creemers, 1994; Klieme, 2012; Kyriakides et al., 2009; Mortimore et al., 1988; Muijs et al., 2014; Muijs & Reynolds, 2000, 2010; Scheerens, 2000; Scheerens & Bosker, 1997; Teodorović, 2011, 2012; Walberg & Paik, 2000). However, it should be acknowledged that most studies which were in the position to detect effects of teacher factors on cognitive and meta-cognitive learning outcomes collected data on quality of teaching through observations. In addition, this study seems to reveal that there was a significant contribution of almost all teacher-level factors on student interest in mathematics and biology, all in the expected direction. These results seem to provide support for the impact of teacher factors on promoting affective learning outcomes and are significant for testing the generic nature of the factors especially since most of the studies testing the validity of the model were able to demonstrate effects of teacher factors on cognitive rather than affective learning outcomes (see Kyriakides et al., 2021).

Importantly, all teacher factors from the dynamic model of educational effectiveness are present in the findings, supporting an integrated approach to teaching rather than focusing on “new learning” (e.g., teaching-modeling or questioning) or on direct teaching (e.g., structuring or application). In other words, students become most interested in mathematics and biology if their teachers frequently teach them how to learn more easily, approach problems in different ways, and find and analyze different information; if their teachers frequently ask them questions, rephrasing, redirecting, and inviting them to explain their answers; and if their teachers frequently correct their work and give them feedback. As expected, orientation – helping students identify why the topic and lesson activities are important and where they fit in real life – generates students’ interest in subject. Similarly, when teachers encourage and support students, creating a positive classroom climate, students show more interest in a subject. Good structuring and application practices help students feel interested in the subject, likely because the subject is clear to them. The fact that these variables are moderately correlated further adds to our conclusion that effective teacher factors go hand in hand and that good teachers usually know how to do many things well in order to interest their students in the subject. These findings resonate well with previous literature.

Findings regarding the lack of impact of teacher-level variables on cognitive learning outcomes can be interpreted in different ways.

First, the measure of achievement used in this study (a test consisting of practice and final national exams in the subject) might depend on a number of factors outside of regular teaching – that is, on the preparation of students for the final test by means of private tuition, preparatory classes at school/municipality, or homework (these are all frequent in Serbia) – so the effects of the examined teacher factors could hardly be applied in that situation. Unlike the final exam, interest is to a much smaller degree subject to outer conditions and pressures on students (interest is not subject to evaluation through grades, nor do students prepare for it); therefore, in this case the effects of examined teacher factors manifest themselves more easily.

Second, the practice exam and the final exam might not adequately represent the majority of what is taught and learned during the fifth to eighth grades of elementary school and are not under the influence of the examined teacher factors. For example, a teacher might spend several biology lessons explaining the steps of photosynthesis and its importance for life on the planet, utilizing many of the effective teacher factors, while on the final national exam in biology photosynthesis is skipped in lieu of some more obscure and less relevant topic. In that case, the examined variables might not show a correlation with this measure of student achievement. On the other hand, the interest was not measured with any external tests, and it represents student opinions, so it is feasible that it is correlated with effective teacher factors.[18]Right click to remove the hyperlink and edit the citation

Third, even if practice and final tests did cover the relevant fifth- to eighth-grade material, it is entirely possible that the format of the test (closed items) and potential inclusion of easily testable, fact-recalling knowledge narrow the capabilities of capturing higher cognitive processes such as analysis, synthesis, and so forth, which may be more prone to depend on teacher factors from the dynamic model of educational effectiveness. In other words, while the final national exam might be a valid tool for student certification and selection processes – although this can be debatable too – it might be less useful for evaluating teacher factors.

Fourth, in this study we examined only one dimension of variables – frequency – and not the other four (focus, stage, quality, and differentiation). Studies testing the validity of the dynamic model revealed the benefit of using different dimensions to measure the functioning of teacher factors. Some factors were not found to have an effect when only the frequency dimension was considered, but had an effect on cognitive learning outcomes when dimensions measuring qualitative characteristics of the factors were considered (see Kyriakides et al., 2018). It may be that, although Serbian teachers differ among themselves in the frequency with which they employ the examined teacher factors, it is not this dimension of variables that makes a difference for promoting cognitive learning outcomes. On the other hand, it is quite possible that the frequency dimension of variables is sufficient for variables to exert their influence on student interest in subject, a finding that seems to be in line with the study searching for effects of teacher factors on affective outcomes (see Kyriakides & Creemers, 2008). In studies that examined both achievement and interest, the effect of variables differs, so this option may be at play (Klieme et al., 2009).

Fifth, questionnaires might not be the adequate instrument for measuring teacher practices. For example, in a study conducted in Ghana, data obtained from questionnaires could not explain the variance in student achievement gains in mathematics, while data acquired through observations could reveal teacher factors that were associated with achievement gains in that subject (Azigwe et al., 2016). In a recent study investigating the impact of teacher factors on student learning in primary schools of the Maldives, the effects of only few teacher factors were identified when data emerged from student questionnaires, whereas observation data revealed effects of all teacher factors on cognitive learning outcomes (Musthafa, 2020).

Sixth, perhaps teacher-level variance of student achievement in each subject is too small for the effects of teacher factors to be discovered. However, this does not seem too likely, since the impact of many similar teacher variables was identified in a study of third-grade achievement in Serbia where even less variance was situated at the class and school levels (Teodorović, 2011). We think that a combination of these options explains the results of the study.

Regardless of these limitations, the significant contribution of teacher factors to student interest in both subject matters should not be downplayed and can have important implications of our study, some for improvement of policy and practice and some for further research. It is to these implications that we now turn.

Improvement of policy and practice

With regards to teachers, our results do not find the impact of teacher-level variables from the dynamic model of educational effectiveness on student achievement in mathematics and biology, but garner solid support for the effect of almost all variables in the model on student interest in mathematics and biology. If we know that intrinsic subject motivation and subject interest are important for student achievement (Deci & Ryan, 1985, 2002; Köller et al., 2001; Schiefele et al., 1992; Wigfield & Cambria, 2010), and if we consider the limitations of this study with regards to achievement (potential problems with the relation between measures of achievement and explanatory variables), we deem it important to improve the examined teacher factors in schools. This could be done in numerous ways.

First, initial education and in-service training of teachers need to be aligned more with effective teacher factors. Not only should curricula and content encompass factors from the dynamic model of educational effectiveness and other similar research, but also the delivery and incentive systems should be designed in ways that have been proven to promote adoption and application of these factors in real-life settings (see Creemers et al., 2013). For example, professional development should be tailored to a specific school's needs and capacities and/or a specific teacher's needs and capacities, not uniformly doled out *en masse* to schools regardless of how competent their teaching force is, what population of students they teach, or what the school development plans are (cf. Creemers et al., 2013; King & Newman, 2004). Also, professional development should be long lasting, continual, and iterative (Lieberman & Pointer Mace, 2008). In-school professional development which entails needs assessment and joint teacher work is an especially effective improvement method (Creemers & Kyriakides, 2012). Similarly, licensing, induction, and mentoring, as well as teacher appraisal and career progression need to be tightly linked to teachers' attainment of effective teacher factors. All of this needs to be preceded with consensus among teacher practitioners and policymakers that improvement of teaching should be at the center of education reforms.

Second, standards of teacher competencies that are implemented in many countries and regions including Serbia (Singapore, Australia, Ontario, etc.) should heavily rely on findings of educational effectiveness studies which identify the most important teacher factors that impact student achievement and interest. This would focus teachers' and policymakers' attention on practices that lead to results. Standards that are aligned with educational effectiveness findings could serve as a blueprint for all aspects of teachers' professional lives that were mentioned above: initial education, licensing, induction, teacher appraisal, in-service training, and career progression. Considering that teachers differ in their mastery of effectiveness factors (Kyriakides et al., 2009), it appears that standards of teacher competencies should be differentiated, for example, developed for several classes of teachers, from novice teachers to master teachers (Teodorović et al., 2019).

Third, measures of student achievement should not be high-stakes tests designed for other purposes for which students prepare heavily and which possibly override some of the teacher influences and/or which cannot capture finer, more cognitively demanding processes. Perhaps better candidates for outcome variables would be results on international tests or annual tests with sufficient coverage of subject material that carry less weight for students.

Finally, studies that encompass classroom observation and interviews/focus groups with students and teachers (especially of high-achieving schools after controlling for student-level variables) should give a richer picture of good teaching.

Further research

Our study also opened an array of very important questions that should be answered by future research: What other teacher-level variables predict achievement in mathematics and biology (cognitive learning outcomes), since variables from the dynamic model of educational effectiveness did not show their impact? Teachers' implementation of curriculum and choice and depth of content appear to be important factors (Klieme, 2019). Are other dimensions of teacher-level variables rather than frequency responsible for their effects on achievement? To what extent may data on the functioning of teacher factors emerged from external observations help us explain variation in cognitive student learning outcomes? Finally, it is important to search for the effect of teacher-level factors on promoting not only quality (i.e., student achievement gains) but also equity in education (i.e., the reduction in the impact of student background factors that are unlikely to change such as SES, gender, and ethnicity on student achievement in cognitive and/or affective learning outcomes). This is an issue that needs more attention from educational effectiveness research especially since almost all studies testing the validity of theoretical models of EER were only concerned with the quality dimension of effectiveness (see Kyriakides et al., 2018).

The present study made a first step at examining the contribution of teacher factors from the dynamic model of educational effectiveness to Serbian student achievement and interest in mathematics and biology. The inclusion of two different subject matters, two different domains of student learning (cognitive and affective), and the fact that it was conducted in a country in which there is no strong tradition in educational effectiveness research underline both the possibilities and the potential of educational effectiveness research worldwide to help us understand what contributes to student learning (broadly defined) and under what circumstances.

Notes

1. This study was a part of the wider project "Improving Educational Effectiveness in Primary Schools" (IEEPS) which had three broad aims: (1) determining how school and teaching influence student achievement in mathematics and sciences, considering the diversity of student population regarding individual student characteristics; (2) delivering individual feedback to schools about student achievement and school's value added; (3) designing and implementing professional development training programs in order to help teachers improve their teaching.

2. The dynamic model of educational effectiveness also contains four school-level factors, but they were not included in this research.

3. Due to time constraints in schools, we could not ask each student to fill out the questionnaires for both mathematics and biology. We estimated that we needed at least 10 students per subject per class to obtain a valid estimate of teacher factors in that subject in that class. So, if an eighth-grade class had 20 or more students, then odd-numbered students (from their alphabetical list) filled out the questionnaire for one subject, while the even-numbered students filled out the questionnaire for another subject. If an eighth-grade class had fewer than 20 students, then an entire class filled out the questionnaire only for one subject, the choice of which was made according to additional criteria. In the end, there were 189 classes with 20 or more students, 34 classes with fewer than 20 students who filled out the mathematics questionnaire, and 17 classes with fewer than 20 students who filled out the biology questionnaire.

4. National exams include tests in mathematics, Serbian language, biology, history, geography, physics, and chemistry. Mathematics and biology were chosen as outcome variables because they are taught from the fifth to the eighth grade (physics and chemistry are taught for fewer years) and because they had an adequate prior achievement measure – TIMSS 2011 results (which Serbian language, history and human geography did not have).

5. Elementary schools in Serbia consist of primary school (Grades 1–4, covering ages 7–10) and lower secondary school (Grades 5–8, covering ages 11–14). Just before starting school, first-grade students are assigned into three to five heterogeneous classes within which they usually remain for the rest of their elementary school period. (During 8 years of elementary school, about a quarter of classes get disbanded and re-formed with students from other same-grade classes in the school, for various reasons.) For example, a large urban school with about 150 students per grade will have five classes with 30 students each. Each class has only one classroom teacher for the first 4 years of elementary school. In fifth grade, classes get their sets of subject teachers, which, in general, remain with them until the end of eighth grade. (However, there is some teacher turnover and change.) Sets of teachers overlap, so, for example, one mathematics teacher may teach eighth-grade Classes 2, 3, and 5, and another eighth-grade Classes 1 and 4, while one biology teacher may teach eighth-grade Classes 1, 3, and 5 and another eighth-grade Classes 2 and 4.

6. Some of the analyses were based on smaller samples, in order to respect the specificities of different variables and meet the aims of the planned analyses.

7. In Serbia, preschool includes both kindergarten (mandatory year before entering elementary school, from the age of 6 to the age of 7, on average) and all the time before kindergarten that is spent in early childhood education and care (before the age of 6, on average).

8. Instead of taking into account general wealth which refers to ownership of objects that can be considered a measure of a lesser or greater luxury (a car, a television set, a washing machine, a dishwasher, etc.), in this study the students only reported the resources relevant for learning and school achievement: a desk, own room, number of books available, a computer available for learning, educational software, internet, classic literature, poetry collections, artwork, dictionaries, and other books useful for learning.

9. In the pilot study, a preliminary version of the questionnaire was created based on the theoretical content of the dynamic model factors. The structural validity of every subscale, that is, the dynamic model factor, was analyzed by exploratory factor analysis, and, when needed, the scale was modified by eliminating problematic items until we reached a satisfactory solution.

10. Since about half of the students in the sample responded to questions about their biology teacher and half responded to questions about their mathematics teacher, we calculated two Cronbach's alpha values for each scale.

11. The practice exam has the same format and number of items as the final exam. It is given to students in April, about 2 months before they take their final exam, just so that they can assess what they are prepared for and what else they need to study. We used both in order to increase the validity of our dependent variable.

12. In comparison to mathematics, there are considerably fewer items for biology in both practice and final exams, because biology items are part of the 20-item test encompassing biology, geography, history, physics, and chemistry.

13. Preliminary analyses concluded that we should not have a three-level model including students, teachers, and schools, as there was little variability of teachers within schools. (In 115 schools we had two classes per school, and in 10 schools we had only one class per school.) Having a two-level model also means that there is probably some school-level variance included in the teacher-level variance.

14. Analyses of the teacher factors' impact on achievement in mathematics and biology, as well as the interest in subject were conducted on a significantly lower number of students than the number that was included in the study itself. The difference in the sample size appeared because, in the achievement models, only students who participated in the TIMSS 2011 study were included, while in the interest models only students who were asked to assess their interest in the subject were included in the analyses. To a smaller extent, missing data also influenced the size of the analytic sample. The exact number of examinees for each analysis is presented in the legend provided below the tables with results of the analysis.

15. In the analyses with interest in subject as an outcome, we could not control for prior achievement (TIMSS 2011 score) because, due to the design of our study, the sample would be reduced too much and multilevel analyses would not be possible. Therefore, we opted for the less optimal option – to control for the current achievement.

16. It has been suggested that effect sizes of about r = 0.15–0.20 translate into the learning gain of 1 school year.

17. Girls perform better than boys on both mathematics achievement and biology achievement.

18. There is also a possibility that associations between teacher factors and interest were due to the fact that both were self-reported data; however, all the associations have expected directions, and it is unlikely that students are familiar with all effective teacher factors in order to report them in such a way. For example, it is not expected that students who show more interest in biology report with bias that their biology teacher announces transitioning to another topic in the lesson more frequently.

Acknowledgements

We thank Vesna Petrović, Irena Golubović Ilić, and Aleksandra Mihajlović for their contribution to the creation of some questionnaire items and piloting of questionnaires.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Jelena Teodorović is an associate professor at the Faculty of Education, University of Kragujevac, Jagodina, Serbia. Her areas of professional interest are: education policy, educational leadership, educational effectiveness, and school improvement. She has published articles in national and international journals and has led two international projects: TEMPUS project "Master Program in Educational Leadership" and Comenius project "Improving the Educational Effectiveness of Primary Schools".

Vladeta Milin is an assistant professor at the Department for Pedagogy and Andragogy, Faculty of Philosophy, University of Belgrade. His main research field is didactics, and the research areas are quality of teaching, teaching methods, theories and models of teaching, teachers' professional development, etc. He has participated in several scientific international and national projects in the field of education and has developed and conducted various teacher professional programs.

Bojana Bodroža, PhD, is an assistant professor at the Department of Psychology, Faculty of Philosophy, University of Novi Sad. Her primary field of interest is social psychology, that is, self & identity, experimental social psychology, cyberpsychology & social networking. She has published articles in international and national journals and has participated in numerous scientific conferences.

Ivana D. Đerić, PhD, is a research associate at the Institute for Educational Research, Belgrade, Serbia. The fields of her research are professional learning of practitioners, student motivation and autonomy, as well as school effectiveness. She was engaged as a National Research Coordinator for Serbia in TIMSS 2019.

Milja Vujačić, PhD, is a senior research associate at the Institute for Educational Research, Belgrade, Serbia. The fields of her research are: inclusive education, teacher professional development, cooperative learning, school effectiveness. She has published over 30 scientific papers and participated at numerous conferences both in our country and abroad.

Ivana M. Jakšić, PhD, is a teaching assistant at the Faculty of Political Sciences, University of Belgrade. Her research interests include the nature, content, and different means of measurement of social perceptions which can promote or undermine social justice, psychological mechanisms through which social perceptions influence behavior, and gender stereotyping.

Dejan Stanković is a graduate psychologist with an MA degree in educational psychology and is currently working as an Education Officer in the UNICEF country office in Serbia. His main areas of research interest are education policy, school improvement, and teacher professional development. He used to work at the Institute for Educational Research, Belgrade, Serbia, where he participated in several large-scale research projects – "TIMSS 2011", "Master Program in Educational Leadership", and "Improving the Educational Effectiveness of Primary Schools".

Gašper Cankar is a researcher and methodology expert working on problems of knowledge assessment. He works in a National Examinations Centre (NEC, Slovenia) methodologically supporting general and vocational matura examinations and national assessments in Grades 6 and 9. Some of the topics he works on include error of rating process, value-added models of knowledge, interplay between socioeconomic background and educational outcomes, and Simpson's paradox.

Charalambos Y. Charalambous is an associate professor in the Department of Education at the University of Cyprus, specializing in educational research and evaluation. His research interests include teaching effectiveness with a particular focus on understanding the work of teaching and measuring teaching quality. He is currently the coordinator of the Special Interest Group 18 of Educational Effectiveness and Improvement of the European Association for Research on Learning and Instruction.

Jan Van Damme is professor emeritus at the Center of Educational Effectiveness and Evaluation at the KU Leuven, Belgium. The focus of his research was: following students and explaining their cognitive and non-cognitive development and trajectory, with the intention to study the functioning of classes, teachers, schools, and even systems. In recent years Jan has focused on the educational subsystems within Flanders, especially the Catholic schools and the public systems.

Leonidas Kyriakides is Professor of Educational Research and Evaluation at the Department of Education of the University of Cyprus. His main research interests are in the area of educational effectiveness and improvement and especially in modeling the dynamic nature of educational effectiveness and in using research to promote quality and equity in education. Leonidas is a member of the editorial board of various international journals with a referee system and the author of more than 130 papers, 10 books, and 100 chapters in books.

ORCID

Jelena Teodorović http://orcid.org/0000-0003-4700-2567

Vladeta Milin http://orcid.org/0000-0002-6939-1201

Milja Vujačić http://orcid.org/0000-0003-1318-338X

Leonidas Kyriakides http://orcid.org/0000-0002-7859-5126

References

Antoniou, P., Kyriakides, L., & Creemers, B. P. M. (2011). Investigating the effectiveness of a dynamic integrated approach to teacher professional development. *Center for Educational Policy Studies Journal*, *1*(1), 13–41.

Azigwe, J. B., Kyriakides, L., Panayiotou, A., & Creemers, B. P. M. (2016). The impact of effective teaching characteristics in promoting student achievement in Ghana. *International Journal of Educational Development*, *51*, 51–61. https://doi.org/10.1016/j.ijedudev.2016.07.004

Baucal, A. (2012). Uticaj socio-ekonomskog statusa učenika na obrazovno postignuće: Direktni i indirektni uticaji [Socioeconomic status and its impact on educational achievements: Direct and indirect impact]. *Primenjena Psihologija*, *5*(1), 5–24. https://doi.org/10.19090/pp.2012.1.5-24

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180. https://doi.org/10.3102/0002831209345157

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

Charalambous, C. Y., & Kyriakides, E. (2017). Working at the nexus of generic and content-specific teaching practices: An exploratory study based on TIMSS secondary analyses. *The Elementary School Journal*, *117*(3), 423–454. https://doi.org/10.1086/690221

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. U.S. Government Printing Office.

Council of the European Union. (2009). Council conclusions of 12 May 2009 on a strategic framework for European cooperation in education and training ("ET 2020"). *Official Journal of the European Union*, C 119, 28.5.2009.

Council of the European Union. (2018). Council recommendation of 22 May 2018 on key competences for lifelong learning. *Official Journal of the European Union*, C 189/1, 4.6.2018.

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

Creemers, B. P. M., & 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*(3), 347–366. https://doi.org/10.1080/09243450600697242

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

Creemers, B. P. M., & Kyriakides, L. (2012). *Improving quality in education: Dynamic approaches to school improvement*. Routledge.

Creemers, B. P. M., & Kyriakides, L. (2015). Developing, testing, and using theoretical models of educational effectiveness for promoting quality in education. *School Effectiveness and School Improvement*, *26*(1), 102–119. https://doi.org/10.1080/09243453.2013.869233

Creemers, B. P. M., Kyriakides, L., & Antoniou, P. (2013). *Teacher professional development for improving quality of teaching*. Springer.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, *19*(2), 294–304. https://doi.org/10.1037/0893-3200.19.2.294

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.

Deci, E. L., & Ryan, R. M. (Eds.). (2002). *Handbook of self-determination research*. University of Rochester Press.

Demetriou, A., & Kazi, S. (2001). *Unity and modularity in the mind and the self: Studies on the relationships between self-awareness, personality, and intellectual development from childhood to adolescence*. Routledge.

Demetriou, A., Kyriakides, L., & Avraamidou, C. (2003). The missing link in the relations between intelligence and personality. *Journal of Research in Personality*, *37*(6), 547–581. https://doi.org/10.1016/S0092-6566(03)00048-5

Dimosthenous, A., Kyriakides, L., & Panayiotou, A. (2020). Short- and long-term effects of the home learning environment and teachers on student achievement in mathematics: A longitudinal study. *School Effectiveness and School Improvement*, *31*(1), 50–79. ttps://doi.org/10.1080/09243453.2019.1642212

Epstein, J. B., Sanders, M. G., Simon, B. S., Salinas, K. C., Jansorn, N. R., & Van Voorhis, F. L. (2002). Family, school, and community partnerships. In **M. H. Bornstein** (Ed.), *Handbook of parenting: Vol. 5. Practical issues in parenting* (2nd ed., pp. 407–437). Erlbaum.

European Commission. (2010). *Europe 2020: A strategy for smart, sustainable and inclusive growth* (Communication from the Commission, Brussels, 3.3.2010).

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, *29*, 1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001

Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). Edward Arnold.

Gustafsson, J.-E., Yang Hansen, K., & Rosén, M. (2013). Effects of home background on student achievement in reading, mathematics, and science at the fourth grade. In **M. O. Martin & I. V. S. Mullis** (Eds.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade – Implications for early learning* (pp. 181–287). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, *326*(5958), 1410–1412. https://doi.org/10.1126/science.1177067

Huntsinger, C. S., Jose, P. E., Larson, S. L., Krieg, D. B., & Shaligram, C. (2000). Mathematics, vocabulary and reading development in Chinese American and European American children over the primary school years. *Journal of Educational Psychology*, *92*(4), 745–760. https://doi.org/10.1037/0022-0663.92.4.745

Jakšić, I., Marušić-Jablanović, M., & Gutvajn, N. (2017). Činioci postignuća učenika iz Srbije u oblasti matematike [Factors of mathematics achievement of students in Serbia]. In **U. M. Marušić-Jablanović, N. Gutvajn, & I. Jakšić** (Eds.). *TIMSS 2015 u Srbiji* (pp. 67–94). Institut za pedagoška istraživanja.

King, B., & Newmann, F. (2004). Key link: Successful professional development must consider school capacity. *Journal of Staff Development*, *25*(1), 26–30.

Klieme, E. (Ed.). (2008). *Teaching and skills acquisition in German and English: Results of DESI study*. Beltz.

Klieme, E. (2012, August 23). *Qualities and effects of teaching: Integrating findings across subjects and cultures* [Paper presentation]. EARLI SIG Educational Effectiveness Conference. Zurich, Switzerland.

Klieme, E. (2019, August 12–16). *Research on teaching: Content matter, practices, quality and effectiveness* [Paper presentation]. European Association for Research on Learning and Instruction (EARLI), Aachen, Germany.

Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In **T. Janik & T. Seider** (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Waxmann.

Köller, O., Baumert, J., & Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education*, *32*(5), 448–470. https://doi.org/10.2307/749801

Konstantopoulos, S. (2012). Teacher effects: Past, present and future. In **S. Kelly** (Ed.), *Assessing teacher quality: Understanding teacher effects on instruction and achievement* (pp. 33–48). Teachers College Press.

Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education*, *14*(1), 23–40. https://doi.org/10.1007/BF03173109

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education*, *36*, 143–152. https://doi.org/10.1016/j.tate.2013.07.010

Kyriakides, L., & Creemers, B. P. M. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: A study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, *19*(2), 183–205. https://doi.org/10.1080/09243450802047873

Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, *25*(1), 12–23. https://doi.org/10.1016/j.tate.2008.06.001

Kyriakides, L., Creemers, B. P. M., Antoniou, P., & Demetriou, D. (2010). A synthesis of studies searching for school factors: Implications for theory and research. *British Educational Research Journal*, *36*(5), 807–830. https://doi.org/10.1080/01411920903165603

Kyriakides, L., Creemers, B., & Charalambous, E. (2018). *Policy implications of research in education: Vol. 8. Equity and quality dimensions in educational effectiveness*. Springer.

Kyriakides, L., Creemers, B. P. M., Panayiotou, A., & Charalambous, E. (2021). *Quality and equity in education: Revisiting theory and research on educational effectiveness and improvement*. Routledge.

Lieberman, A., & Pointer Mace, D. H. (2008). Teacher learning: The key to education reform. *Journal of Teacher Education*, *59*(3), 226–234. https://doi.org/10.1177/0022487108317020

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Education Research and Evaluation*, *12*(2), 105–128. https://doi.org/10.1080/13803610600587040

Martin, M. O., & Mullis, I. V. S. (Eds.). (2013). *TIMSS and PIRLS 2011: Relationships among reading, mathematics and science achievement at the fourth grade – Implications for early learning*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).

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

Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art – Teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, *25*(2), 231–256. https://doi.org/10.1080/09243453.2014.885451

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*(3), 273–303. https://doi.org/10.1076/0924-3453(200009)11:3;1-G;FT273

Muijs, D., & Reynolds, D. (2010). *Effective teaching: Research and practice* (3rd ed.). Paul Chapman.

Musthafa, H. S. (2020). *A longitudinal study on the impact of instructional quality on student learning in primary schools of Maldives* [Unpublished doctoral dissertation]. University of Cyprus.

Opdenakker, M.-C., & 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*(2), 165–196. https://doi.org/10.1076/0924-3453(200006)11:2;1-Q;FT165

Pavlović Babić, D., & Baucal, A. (2013). *Podrži me, inspiriše me. PISA 2012 u Srbiji: prvi rezultati* [Support me, inspire me. PISA 2012 in Serbia: First results]. Institut za psihologiju i Centar za primenjenu psihologiju.

Praetorius, A.-K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: Looking back and looking forward. *ZDM Mathematics Education*, *50*(3), 535–553. https://doi.org/10.1007/s11858-018-0946-0

Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): A state-of-the-art review. *School Effectiveness and School Improvement*, *25*(2), 197–230. https://doi.org/10.1080/09243453.2014.885450

Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, *1*(1), 61–80. https://doi.org/10.1080/0924345900010106

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

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

Scheerens, J. (2016). *Educational effectiveness and ineffectiveness: A critical review of the knowledge base*. Springer.

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

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In **K. A. Renninger, S. Hidi, & A. Krapp** (Eds.), *The role of interest in learning and development* (pp. 183–212). Lawrence Erlbaum Associates.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417–453. https://doi.org/10.3102/00346543075003417

Stoffelsma, L. (2018). Short-term gains, long-term losses? A diary study on literacy practices in Ghana. *Journal of Research in Reading*, *41*(S1), S66–S84. https://doi.org/10.1111/1467-9817.12136

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

Teodorović, J. (2011). Classroom and school factors related to student achievement: What works for students? *School Effectiveness and School Improvement*, *22*(2), 215–236. https://doi.org/10.1080/09243453.2011.575650

Teodorović, J. (2012). Student background factors influencing student achievement in Serbia. *Educational Studies*, *38*(1), 89–110. https://doi.org/10.1080/03055698.2011.567027

Teodorović, J., Bodroža, B., & Stanković, D. (2015). Porodični resursi i kvalitetna nastava kao faktori postignuća učenika iz matematike i prirodnih nauka: Analiza TIMSS 2011 u Srbiji [Family resources and quality teaching as factors of student achievement in mathematics and science: Analysis of TIMSS 2011 in Serbia]. In **U. J. Radišić & N. Buđevac** (Eds.), *Sekundarne analize istraživačkih nalaza u svetlu novih politika u obrazovanju* (pp. 173–185). Beograd: Ministarstvo prosvete, nauke i tehnološkog razvoja.

Teodorović, J., Milin, V., & Stanković, D. (2019). Standardi kompetencija za nastavnike u Srbiji: Poređenje sa odabranim zemljama [Standards of teacher competencies in Serbia: Comparative analysis with selected countries]. *Zbornik Instituta za pedagoška istraživanja*, *51*(2), 613–653. https://doi.org/10.2298/ZIPI1902614T

Tsukayama, E., Duckworth, A. L., & Kim, B. (2013). Domain-specific impulsivity in school-age children. *Developmental Science*, *16*(6), 879–893. https://doi.org/10.1111/desc.12067

Van Landeghem, G., Van Damme, J., Opdenakker, M.-C., De Fraine, B., & Onghena, P. (2002). The effect of schools and classes on noncognitive outcomes. *School Effectiveness and School Improvement*, *13*(4), 429–451. https://doi.org/10.1076/sesi.13.4.429.10284

Vigil-Coleț, A., & Morales-Vives, F. (2005). How impulsivity is related to intelligence and academic achievement. *The Spanish Journal of Psychology*, *8*(2), 199–204. https://doi.org/10.1017/S1138741600005072

Viljaranta, J., Tolvanen, A., Aunola, K., & Nurmi, J. E. (2014). The developmental dynamics between interest, self-concept of ability, and academic performance. *Scandinavian Journal of Educational Research*, *58*(6), 734–756. https://doi.org/10.1080/00313831.2014.904419

Vujić, S., & Baronijan, H. (2013). Odnosi između pohađanja predškolskog obrazovanja i školskog uspeha učenika i učenica i mogućnosti unapređenja predškolskog obrazovanja u Srbiji [Relationship between pre-school attendance and academic success of students and possibilities of improving pre-school education in Serbia]. *Psihološka istraživanja*, *16*(2), 105–140. https://doi.org/10.5937/PsIstra1302105V

Wagerman, S. A., & Funder, D. C. (2007). Acquaintance reports of personality and academic achievement: A case for conscientiousness. *Journal of Research in*

Personality, *41*(1), 221–229. https://doi.org/10.1016/j.jrp.2006.03.001

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

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, *30*(1), 1–35. https://doi.org/10.1016/j.dr.2009.12.001

Wijsman, L. A., Warrens, M. J., Saab, N., van Driel, J. H., & Westenberg, P. M. (2016). Declining trends in student performance in lower secondary education. *European Journal of Psychology of Education*, *31*(4), 595–612. https://doi.org/10.1007/s10212-015-0277-2

Wilder, S. (2014). Effects of parental involvement on academic achievement: A meta-synthesis. *Educational Review*, *66*(3), 377–397. https://doi.org/10.1080/00131911.2013.780009

Appendix 1. Exploratory factor analysis of dimensions from the dynamic model of educational effectiveness for mathematics and biology teacher factors

	Component	
Teacher factors	Mathematics	Biology
Teaching-modeling	.90	.88
Questioning	.87	.85
Structuring	.86	.85
Application	.85	.83
Assessment quality	.82	.83
Orientation	.82	.83
Classroom as a learning environment	.82	.81
Management of time	.63	.58

Note: Exploratory factor analysis was done using principal components analysis. The extracted component explained 67.76% of the variance of the variables for mathematics and 66.11% for biology.