University of Groningen

Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender

Kyriakides, Leonidas; Creemers, Bert P. M.; Charalambous, Evi

Published in:
School Effectiveness and School Improvement

DOI:
[10.1080/09243453.2018.1511603](https://doi.org/10.1080/09243453.2018.1511603)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kyriakides, L., Creemers, B. P. M., & Charalambous, E. (2019). Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender: implications for promoting equity . *School Effectiveness and School Improvement*, 30(3), 286-308.
<https://doi.org/10.1080/09243453.2018.1511603>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender: implications for promoting equity

Leonidas Kyriakides, Bert P.M. Creemers & Evi Charalambous

To cite this article: Leonidas Kyriakides, Bert P.M. Creemers & Evi Charalambous (2019) Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender: implications for promoting equity, *School Effectiveness and School Improvement*, 30:3, 286-308, DOI: [10.1080/09243453.2018.1511603](https://doi.org/10.1080/09243453.2018.1511603)

To link to this article: <https://doi.org/10.1080/09243453.2018.1511603>

Published online: 06 Sep 2018.

Submit your article to this journal [↗](#)

Article views: 286

View Crossmark data [↗](#)

Citing articles: 2 View citing articles [↗](#)

ARTICLE

Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender: implications for promoting equity

Leonidas Kyriakides^a, Bert P.M. Creemers^b and Evi Charalambous^a

^aDepartment of Education, University of Cyprus, Nicosia, Cyprus; ^bFaculty of Behavioural and Social Sciences, University of Groningen, Groningen, The Netherlands

ABSTRACT

This paper argues that research on differential teacher and school effectiveness can be used to explore the relation between quality and equity in education. The results of secondary analyses of 2 studies concerned with effectiveness in different phases of schooling (pre-primary and primary education) are presented. These secondary analyses reveal that the socioeconomic gap in student achievement in each of the 2 core subjects (mathematics and language) tends to be smaller in classrooms and schools that are more effective than others in promoting student learning outcomes (after controlling for prior achievement, socioeconomic status, and gender). Such a relation was not identified when searching for differential teacher and school effectiveness in terms of gender. Implications for research on promoting quality and equity are drawn.

KEYWORDS

Educational effectiveness research; quality and equity in education; socioeconomic background; gender; multilevel modelling

Introduction

It is generally expected in society that education should achieve high results in the various learning domains and subject areas. School failure was found to have a negative long-lasting impact on a child's life (Micklewright & Schnepf, 2007; Organisation for Economic Co-operation and Development [OECD], 2013), and the overall social and economic costs of school failure are estimated to be extremely high (Atkinson, 2015; Psacharopoulos, 2007). Thus, policymakers, researchers, and practitioners search for ways to ensure that specific learning objectives are achieved. In this context, international studies such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) evaluate the knowledge, skills, and abilities of different groups of students, and their results are treated as indicators of the effectiveness status of different educational systems. This means that the criteria for educational effectiveness should be at a level that is feasible to be obtained by individual students, classes, and schools with respect to specific learning objectives (excellence).

However, children are not all equal in regard to education failure. Evidence shows that children from the poorest homes are more likely to have worse school results and to drop out of school more frequently than children that come from better off families. Quantitative syntheses of educational studies revealed that the socioeconomic status (SES) of students has a relatively strong impact on student achievement (Sirin, 2005; White, 1982). Research on group composition effects should also be considered, especially since these studies reveal statistically additional effects over and above the effects of students' individual characteristics, when individual characteristics are aggregated at a higher level such as class or school (see Belfi, Haelermans, & De Fraine, 2016; Driessen, 2007; Hattie, 2002; Hornstra, Van der Veen, Peetsma, & Volman, 2015; Rjosk et al., 2014; Rumberger & Palardy, 2005; Van Ewijk & Sleegers, 2010). Thus, group composition effects may play an important role in the achievement gaps faced by students at risk (Verhaeghe, Vanlaar, Knipprath, De Fraine, & Van Damme, 2018). One can therefore look at the effectiveness of a teacher or school from a different angle, especially through investigating how far each school has managed to reduce the impact that student background factors have on student learning outcomes (Creemers & Kyriakides, 2015; Kelly, 2012). As an education system is a fundamental lever to make a society more equitable, achieving fair and inclusive systems is a challenge that countries/schools/teachers cannot afford to neglect not only for economic but also for social reasons. Regardless of political perspective, all agree that, in a democratic society, socioeconomic inequalities in educational outcomes should be minimal (Marks, Cresswell, & Ainley, 2006; OECD, 2013). Thus, one of the major objectives of education systems around the world is to understand the schooling processes that provide opportunities for all learners to succeed in school (Frempong, Reddy, & Kanjee, 2011) and thereby promote both quality and equity in education (Charalambous, Kyriakides, & Creemers, 2018). This objective has greater significance particularly in unequal societies that are striving for redistribution of opportunities. To achieve this objective, schools within these systems should not only help students achieve learning outcomes but they also need to function in a way that students' success in learning is not determined by their background characteristics.

However, there are different visions on how equity in education can be defined (Atkinson, 2015; Paquette, 1998) which can generally be divided into two broad categories: (a) a meritocratic view and (b) an egalitarian view. The meritocratic view is based on the assumption that "status" in society should be the reflection of one's own "merits", "talents", and "effort" (Gulson & Webb, 2012; McCoy & Major, 2007). In this view, differences in student learning outcomes are attributed to differences among students in terms of their cognitive abilities, talents, and amount of work dedicated to schooling (i.e., the effort they make in order to succeed). However, educational effectiveness research (EER) reveals that student learning outcomes are associated with various student factors including background characteristics such as SES, gender, and ethnicity (Scheerens, 2013). Even if students are given the same opportunities inside the schools, not all students will manage to develop their talents since different hidden mechanisms in the society, as well as differences in their home learning environment, are likely to affect their progress (Lim, 2013). Based on these critiques of the meritocratic view, the egalitarian vision has evolved as the mainstream view of equity, which implies that the main responsibility to achieve equity should be placed on society (Creemers &

Kyriakides, 2008; Van Damme & Bellens, 2017). In the case of education, national/state agencies and schools/teachers are expected to provide further support to those socially disadvantaged groups of students (based on their background characteristics such as SES, gender, and ethnicity) in order to ensure that differences in their learning outcomes are substantially reduced (Kelly, 2012; OECD, 2012). This implies that positive discrimination of different groups is not only legitimated to obtain equity in education but is also seen as a characteristic of effective education (Creemers & Kyriakides, 2008).

Equity in education could therefore be examined in two ways that are closely linked and can help us analyse the implications of school failure for schools/systems: equity as *fairness* and equity as *inclusion*. Specifically, school failure can be seen as twofold. On the one hand, school failure could be seen as the failure of an educational system, which is unable to provide an education of quality to all. In this case, overcoming school failure implies assuring inclusion: ensuring a basic minimum standard education for each and every student. The inclusion perspective has implications for designing effective national reform policies that minimize dropout rates and provide learning opportunities for all children. Second, school failure can be attributed to the fairness perspective, which is based on the fact that factors beyond those that students can control are associated with student learning outcomes. Fairness implies ensuring that personal and social circumstances should not be an obstacle to educational success, and inclusion implies ensuring a minimum standard of education for all (Field, Kuczera, & Pont, 2007).

In this paper, we focus on equity as fairness and argue that this element of equity should be taken into account in evaluating the impact that teachers and schools may have on promoting equity in education. This implies that the equity dimension of effectiveness in education demands that students' expected learning outcomes should depend only on their own effort and capacity, and not on considerations over which they have no influence (e.g., gender, ethnic origin, family's socioeconomic level). Thus, the teacher/school/system effectiveness status in terms of equity can be measured by looking at the extent to which differences in learning outcomes between groups of students with different background characteristics are reduced. At this point, it is important to mention that educational effectiveness studies reveal that factors associated with student achievement gains operate at different levels: student, class, school, and system (Muijs et al., 2014; Reynolds et al., 2014; Scheerens, 2016). Student-level factors are classified within EER into three categories: (a) sociocultural and economical background variables emerged from the sociological perspective of EER such as SES, gender, and ethnicity; (b) background variables emerged from the psychological perspective of EER such as motivation and thinking style; and (c) variables related to specific learning tasks associated with the learning outcomes used to measure effectiveness such as prior achievement, time on task, and opportunity to learn (see Creemers & Kyriakides, 2008; Scheerens, 2016). A distinction is also made among the student-level factors by referring to factors which are unlikely to change (e.g., gender, SES, ethnicity) and factors that may change over time (e.g., subject motivation, thinking styles). Factors that are unlikely to change are closely related with the fairness dimension of equity. The fairness dimension of equity implies that school stakeholders are expected to reduce the impact of this group of factors on student learning outcomes (Kyriakides, Creemers, & Charalambous, 2018).

By introducing two different dimensions of measuring effectiveness, a question that arises is the extent to which teachers/schools/systems can be effective in terms of not only promoting the learning outcomes of their students (quality) but also reducing the impact of student background factors (and especially SES) on student achievement (equity). Creemers and Kyriakides (2015) argue that a group of researchers in psychology, sociology, and economy of education treated these two dimensions of effectiveness as competing each other and supported different approaches on how to deal with the “cost” of promoting the one rather the other (Whitty, 2001). This can partly be attributed to the fact that these two dimensions were never explicitly defined and consequently there is not enough evidence investigating the relation between the two dimensions of effectiveness in classrooms, schools, and educational systems (Kelly, 2012; Kyriakides et al., 2018). This research question should be addressed by EER. Coming from the history of research on inequality in education, it was evident that EER would look at the educational outcomes of disadvantaged children in particular and search for equity in schools. This meant looking at the extent to which schools were able to compensate for initial differences in defined outcomes (Sammons, 2010). However, most effectiveness studies, while examining the magnitude of teacher and school effects, have paid very little attention to the extent to which teachers and schools perform consistently across different school groupings (Kyriakides, 2007). As a consequence, the concepts of teacher and school effectiveness have been developed in a generic way by drawing up a one-size-fits-all model (Campbell, Kyriakides, Muijs, & Robinson, 2004). Although greater effectiveness may somewhat improve the absolute performance of disadvantaged groups, critics of EER argue it will not improve their relative performance against more advantaged groups (Thrupp, 2001). This implies that researchers within the field of EER should examine further issues dealing with equity in education.

In this paper, it is argued that research into differential teacher and school effectiveness may provide a new perspective in the discussion about educational equality, and answers could be provided to the critics of EER who argue that EER has not given consideration to equity and justice. Fielding (1997) acknowledged the early work of EER as “a necessary corrective to an overly pessimistic, even deterministic, view of the influence of social and political factors on the efficacy of schools” (p. 141). Research into differential effectiveness may help in moving a step forward by raising issues regarding the development and effective implementation of policy on educational equality. If schools differ significantly in terms of their effectiveness for particular groups of students, issues concerning the extent to which specific factors are associated with school effectiveness in promoting the progress of specific groups of pupils can be examined (Kyriakides, 2007). In this respect, research on differential teacher treatment of students depending on their background characteristics (including SES and prior achievement) has been conducted. These studies seem to reveal that teachers in schools serving relatively high proportions of low-SES students emphasized basic skills more and engaged in more didactic teaching and less constructivist teaching approaches (e.g., Campbell, Kyriakides, Muijs, & Robinson, 2003; Stipek, 2004). In addition, studies on instructional quality and student background characteristics have revealed that classroom inequality within and between schools contributed substantially to achievement gaps developed during the early school years (e.g., Palardy, 2015; Palardy & Rumberger, 2008; Xue & Meisels, 2004). These studies revealed that the theoretical framework of EER

should be expanded by searching for factors associated with both the quality and equity dimension of effectiveness. The identification of these factors may also be useful for policymakers attempting to design and implement policies on equal opportunities (see Kyriakides et al., 2018).

In this context, the dynamic model of educational effectiveness (Creemers & Kyriakides, 2008) supports that both the quality and the equity dimension of educational effectiveness should be considered in establishing criteria for measuring effectiveness. Specifically, the dynamic model is multilevel in nature and refers to factors operating at four different levels: *student*, *classroom/teacher*, *school*, and *system*. At the classroom level, the model refers to the following eight factors which describe teachers' instructional role: orientation, structuring, questioning, teaching modelling, application, management of time, teacher role in making the classroom a learning environment, and classroom assessment. The model refers to skills associated with direct teaching and mastery learning (Joyce, Weil, & Calhoun, 2000) such as structuring and questioning. Factors included in the model such as orientation and teaching modelling are in line with theories of teaching associated with constructivism (Brekelmans, Slegers, & Fraser, 2000). Moreover, the collaboration technique is included under the overarching factor of teacher contribution to the establishment of classroom learning environment. Therefore, an integrated approach to quality of teaching is adopted. Since learning takes place primarily at the classroom level, factors situated at the school and system levels are expected to influence primarily the teaching practice and, through that, student learning outcomes. School-level factors are expected to influence the teaching-learning situation by developing and evaluating the school policy on teaching and the policy on creating the school learning environment. The system level refers to the influence of the educational system in a more formal way, especially by developing and evaluating the educational policy at the national/regional level. The teaching and learning situation is also influenced by the wider educational context in which students, teachers, and schools operate.

Moreover, the dynamic model is based on the assumption that each factor can be defined and measured using five dimensions: *frequency*, *focus*, *stage*, *quality*, and *differentiation* (Creemers & Kyriakides, 2006). First, the *frequency* dimension refers to the number of times that an activity associated with a factor is present in a system, school, or classroom. This is probably the easiest way to measure the effect of a factor on student achievement, and almost all studies used this dimension to define effectiveness factors (Kyriakides, 2008). Second, two aspects of the *focus* dimension are taken into account. The first one refers to the specificity of the activities, which can range from specific to general. The second aspect of this dimension addresses the purpose for which an activity takes place. An activity may be expected to achieve a single or multiple purposes. Third, the activities associated with a factor can be measured by taking into account the *stage* at which they take place. Factors need to take place over a long period of time to ensure that they have a continuous direct or indirect effect on student learning (Levin, 2005). Fourth, the *quality* dimension refers to the properties of the specific factor itself as discussed in the literature. For instance, teacher assessment can be measured by looking at the extent to which the formative rather than the summative purpose is served (Black & Wiliam, 2009; Marshall & Drummond, 2006). Finally, *differentiation* refers to the extent to which activities associated with a factor are

implemented the same way for all subjects involved with it. The use of different measurement dimensions reveals that looking at just the frequency dimension of an effectiveness factor does not help us identify those aspects of the functioning of a factor which are associated with student achievement. Thus, the five dimensions are not only important from a measurement perspective but also, and to a greater degree, from a theoretical point of view (Kyriakides, 2008).

The model does not only refer to factors associated with student achievement gains but also assumes that there is a close relation between the two dimensions of effectiveness (quality and equity). However, this assumption has never been made explicit, and the relation between the quality and the equity dimensions of effectiveness has not been examined. During the last 15 years, several longitudinal studies (e.g., Antoniou & Kyriakides, 2011; Azigwe, Kyriakides, Panayiotou, & Creemers, 2016; Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009; Panayiotou et al., 2014) and two meta-analyses of effectiveness studies (Kyriakides, Christoforou, & Charalambous, 2013; Kyriakides, Creemers, Antoniou, & Demetriou, 2010) provided support to the validity of the dynamic model by demonstrating relations between teacher and school factors included in the model with student achievement gains. This implies that factors of the dynamic model may be taken into account for promoting quality in education (see Creemers & Kyriakides, 2015) but this does not necessarily mean that these factors are relevant for promoting equity.

Research aims

In this paper, we reanalysed the data of two effectiveness studies (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009) and searched for differential teacher and school effectiveness in terms of SES and gender by conducting four separate multilevel analyses. Specifically, these two studies used the same research methods and instruments to measure the teacher and school factors of the dynamic model of educational effectiveness and searched for their effects on student learning outcomes in two core subjects (i.e., mathematics and language) at two different phases of schooling (i.e., pre-primary and primary education). In this way, drawing from the data of each study, two separate analyses searching for the effect of teacher and school factors on student achievement in two different subjects (i.e., mathematics and language) were conducted. The main findings of these two studies revealed that almost all teacher and school factors and their dimensions were associated with student achievement gains in two different subjects at two different phases of schooling. Both studies also demonstrated the added value of using the five dimensions included in the dynamic model to measure the teacher and school factors for explaining variation of student achievement in each subject (Creemers & Kyriakides, 2015; Kyriakides & Creemers, 2008). By taking into account the combination of frequency dimension with the other four dimensions of teacher-level factors, the explained variance on student achievement is increased. Moreover, there are factors which have no statistically significant effect on student achievement by measuring the effect of their frequency dimension, but they are associated with achievement when other dimensions are taken into account. This implies that previous studies concerned only with the frequency dimension might draw the wrong conclusions about the effect of a factor and consequently fail to explain

as much variance as possible at the teacher and school level. For example, the study conducted at pre-primary schools revealed that the frequency dimension of assessment did not have an effect on mathematics achievement but when measuring assessment by using the focus, quality, and differentiation dimensions, statistically significant effects on achievement gains in each subject were identified. Similarly, in the study conducted at primary schools, the frequency dimension of teaching modelling was not associated with student achievement on any outcome, but the quality dimension of this factor had an effect on achievement in both subjects. These findings reveal that emphasis should be given to other dimensions of effectiveness factors as well as to frequency, which has been used predominantly in all effectiveness studies in the past. Different dimensions to measure the functioning of teacher and school factors are used for theoretical reasons in our attempt to describe the complex nature of educational effectiveness.

However, in this paper, we move a step forward and search for differential teacher and school effectiveness in terms of two student background factors that are not likely to change (i.e., SES and gender) in order to explore the relation between the quality and equity dimension of effectiveness. We draw data from two effectiveness studies which collected data from two different phases of schooling and examine the extent to which four separate analyses (concerned with achievement in two different subjects and at two different phases of schooling) may generate similar results on the relation between quality and equity. Specifically, this paper investigates whether primary and/or pre-primary teachers/classrooms and schools which are more effective in terms of promoting student learning outcomes tend to have a smaller SES gap and/or smaller gender gap on student achievement. In this way, we explore the relation between quality and equity in different phases of schooling (i.e., pre-primary and primary education) and when achievement in the two core subjects of the curriculum (i.e., mathematics and language) is taken into account. However, we acknowledge a limitation concerned with the fact that four separate secondary analyses were conducted in order to find out whether similar results emerged in investigating the relation between quality and equity in education. It was therefore not possible to generate an indicator measuring the consistency in the results of these four analyses through using a multivariate technique in analysing our data.

Methods

This section presents the methods of each study conducted to test the validity of the dynamic model of educational effectiveness and shows that the two studies (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009) used similar methodologies in testing the validity of the model (for more information about the two studies, see Kyriakides & Creemers, 2008, 2009). Although we run four separate analyses to explore the relation between quality and equity, we argue that the results emerged from each of the four secondary analyses can be compared. The comparison may help us find out whether similar results about the relation between quality and equity in effectiveness at class and school level emerged from analysing separately the data of the two different studies.

Participants

Stratified sampling was used to select 80 pre-primary schools in Cyprus. All the students ($N = 2,812$) who attended all classes ($N = 141$) of the last year of pre-primary education of the school sample were chosen. The median age in the overall sample was 5.32 years with students ranging in age from 4.91 to 5.65 years at the time they entered the last year of pre-primary school. The total sample consisted of 1,518 (53.9%) girls and 1,294 (46.1%) boys. The chi-square test did not reveal any statistically significant difference between the research sample and the population in terms of students' sex ($X^2 = 0.84$, $df = 1$, $p = 0.42$) or type of school (i.e., public or private) ($X^2 = 0.35$, $df = 1$, $p = 0.55$). Although this study refers to other variables such as the SES of students and their achievement levels in different outcomes of schooling, there are no data about these characteristics of the Greek-Cypriot students who are in the last year of pre-primary education. Therefore, it was not possible to examine whether the sample was nationally representative in terms of any other characteristic than students' sex and type of school.

Stratified sampling (Cohen, Manion, & Morrison, 2000) was also used to select 52 Greek-Cypriot primary schools. All the Grade 6 students from all classes ($N = 108$) of the school sample were chosen. The chi-square test did not reveal any statistically significant difference between the sample of this study and the population in terms of students' sex ($X^2 = 0.69$, $df = 1$, $p = 0.41$). Moreover, the t test did not reveal any statistically significant difference between the research sample and the population in terms of the size of class ($t = 1.62$, $df = 111$, $p = 0.11$). Therefore, the sample of Study 2 was nationally representative in terms of students' sex and size of class.

Dependent variables

Measuring pre-primary student achievement in mathematics and language

Student skills in language and mathematics were measured at the beginning and at the end of school year 2005–2006. As far as the measurement of mathematics achievement is concerned, the performance test developed for the purposes of a study on baseline assessment in mathematics (Kyriakides, 2002) was administered in order to assess knowledge and skills in mathematics identified in the Cyprus curriculum. These are in line with the main findings of research into early mathematics development (Geary, 1994; Smith, 1997). Specifically, the students of our sample were asked by our research team to complete at least two practical criterion-referenced tasks (e.g., students were given a picture with two carpets and were asked to colour the widest) related to each objective in the teaching programme of mathematics for the last year of pre-primary education. The members of our research team used an empty room in which individual students completed the tasks of the performance test and observed how each student responded to the task. They also asked him/her to explain what he/she had done in order to decide whether a particular child had completed a task correctly. Thus, a loosely structured interview was used for the purpose of gaining insight into each child's thought. As far as the measurement of language achievement is concerned, the Emergent Literacy Baseline Assessment (ELBA) test (see Kyriakides & Kelly, 2003) was used. This test was designed to assess knowledge and skills in emergent literacy which are identified in the Cyprus

curriculum and are in line with the main findings of research into early language development. Information on the assessment development, the items, and the psychometric properties of the test is provided by Kyriakides and Kelly (2003). Students in our sample were also administered curriculum-based performance tests in mathematics and language when they were at the end of pre-primary school. In order to equate the tests used to measure mathematics and language achievement at different time periods, we used enough common items (i.e., approximately 15% of the number of items of each test) with representative content to be measured (Kolen & Brennan, 1995). The baseline and final tests were subject to control for reliability and validity (see Kyriakides & Creemers, 2009).

Measuring primary student achievement in mathematics and language

Data on achievement in mathematics and Greek language were collected by using external forms of assessment. Written tests were administered to our student sample when they were at the beginning of Grade 6 and when they were at the end of Grade 6. The construction of the tests was subject to control for reliability and validity (see Kyriakides & Creemers, 2008).

Creating interval scales to measure student achievement in each study

In each study, for the estimation of the achievement data in each of the two subjects, data scaling was conducted by applying the Partial Credit Model (Masters & Wright, 1997) as implemented in the software ConQuest (Wu, Adams, & Wilson, 1998). As estimation algorithm, an adaptation of the quadrature method described by Bock and Aitken (1981) was used. Thus, we used the weighted likelihood estimates, which emerged by using item response theory (IRT) (see Warm, 1989) in order to generate two different scores for the achievement of each student participating in each study (per subject) at the beginning and at the end of a school year.

Explanatory variables at student level

Aptitude

Aptitude refers to the degree to which a student is able to perform the next learning task. For the purpose of each study, it consists of prior knowledge of each subject (i.e., mathematics and language) and emerged from student responses to the tests administered to them at the beginning of their last year of pre-primary school or Year 6, respectively (i.e., baseline assessment).

Student background factors

Information was collected on two student background factors: sex (0 = boys, 1 = girls) and SES. Five SES variables were available: father's and mother's education level (i.e., graduate of a primary school, graduate of secondary school, or graduate of a college/university), the social status of father's job, the social status of mother's job, and the economic situation of the family. Following the classification of occupations used by the Ministry of Finance, it was possible to classify parents' occupation into three groups which have relatively similar sizes: occupations held by working class, occupations held by middle class, and occupations held by upper middle class. Representative parental occupations for the working

class are: farmer, truck driver, machine operator in a factory; for the middle class: police officer, teacher, bank officer; and for the upper middle class: doctor, lawyer, business executive. Relevant information for each child was taken from the school records. Then, standardized values of the above five variables were calculated, resulting in the SES indicator, which was an average of these five standardized values.

Methodology used for analysing the data of each study

Since at primary and pre-primary school level classroom teachers are responsible for delivering the curriculum of the two subjects mathematics and Greek language, the classroom and teacher level are identical. Separate multilevel analyses for each subject were conducted by using MLwiN (Rasbash, Steele, Browne, & Prosser, 2005), and the data were conceptualized as a three-level model, consisting of students at the first level, classrooms/teachers at the second level, and schools at the third level. The first step in the analysis was to determine the variance at each level without explanatory variables (empty model). This model contains random groups and random variation within groups. It can be expressed as a model where the dependent variable is the sum of a general mean (β_0), a random effect at the school level (V_{0k}), a random effect at the classroom level (U_{0jk}), and a random effect at the individual level (R_{ijk}). In Model 1, student background factors (i.e., SES, gender, and prior achievement) and their aggregated scores at classroom and school levels were added to the empty model. As was mentioned in the first part of the paper, our decision to consider the aggregated effects of background factors is based on the findings of studies investigating group composition effects which reveal that aggregated scores have in some countries even stronger effects on final achievement than the individual factors. For example, two meta-analyses of studies investigating the effect of SES on student achievement reveal that when researchers use aggregated measures of SES, they usually report much higher correlations between SES and academic achievement than when individual measures of SES are used (see Sirin, 2005; White, 1982). It is also important to note that variables measuring background factors were centred as Z scores with a mean of 0 and a standard deviation of 1. Gender was entered as a dummy variable with one of the two groups as baseline (i.e., boys = 0).

We then searched for differential effectiveness in terms of SES and gender by running random-slope models (see Snijders & Bosker, 2012). Below, we provide the equation of the random slope model (i.e., Model 2) which was used to find out whether the SES indicator and/or gender have any differential effects at the class and/or school levels on student achievement in mathematics.

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}(\text{SES})_{ijk} + \beta_{2jk}(\text{Gender})_{ijk} + \beta_3(\text{Prior Achievement in mathematics})_{ijk} + \beta_4(\text{Average Prior Achievement at class level})_{jk} + \beta_5(\text{Average SES at class level})_{jk} + \beta_6(\text{Percentage of Girls at class level})_{jk} + \beta_7(\text{Average Prior Achievement at school level})_k + \beta_8(\text{Average SES at school level})_k + \beta_9(\text{Percentage of Girls at school level})_k + R_{ijk}(\text{Level 1 model})$$

(1)

The intercept (β_{0jk}) as well as the slopes (β_{1jk} and β_{2jk}) are group dependent. These group-dependent coefficients can be split into an average coefficient and the group-dependent deviation:

$$\begin{aligned}\beta_{0jk} &= \beta_{0k} + u_{0jk} \text{ (Level 2 model for intercept)} \\ \beta_{1jk} &= \beta_{1k} + u_{1jk} \text{ (Level 2 model for slope of SES)} \\ \beta_{2jk} &= \beta_{2k} + u_{2jk} \text{ (Level 2 model for slope of gender)} \\ \beta_{0k} &= \beta_{000} + V_{0k} \text{ (Level 3 model for intercept)} \\ \beta_{1k} &= \beta_{100} + V_{1k} \text{ (Level 3 model for slope of SES)} \\ \beta_{2k} &= \beta_{200} + V_{2k} \text{ (Level 3 model for slope of gender)}\end{aligned}\quad (2)$$

Substitution now leads to Model 2, shown below:

$$\begin{aligned}Y_{ijk} &= \beta_0 + \beta_1(\text{SES})_{ijk} + \beta_2(\text{Gender})_{ijk} + \beta_3(\text{Prior Achievement in mathematics})_{ijk} + \\ &\quad \beta_4(\text{Average Prior Achievement at class level})_{jk} + \beta_5(\text{Average SES at class level})_{jk} + \\ &\quad \beta_6(\text{Percentage of Girls at class level})_{jk} + \beta_7(\text{Average Prior Achievement at school level})_k + \\ &\quad \beta_8(\text{Average SES at school level})_k + \beta_9(\text{Percentage of Girls at school level})_k + \\ &\quad V_{0k} + V_{1k} + V_{2k} + u_{0jk} + u_{1jk} + u_{2jk} + R_{ijk}\end{aligned}\quad (3)$$

The residual of the intercept at Level 2 can be treated as an indication of the teacher effect on the quality dimension of effectiveness especially since prior achievement and background factors are controlled and the teacher's "value-added" contributions can be estimated. These are typically referred to as the effectiveness scores of teachers (Creemers, Kyriakides, & Sammons, 2010). It should, however, be acknowledged that several student characteristics (e.g., ethnicity, motivation, time on task) were not controlled in this analysis and effectiveness studies conducted in different countries reveal that they are associated with student achievement gains (Creemers & Kyriakides, 2015). One could also claim that classroom contextual factors such as class size may be associated with student achievement. Therefore, these residual scores are likely to reflect several unmeasured factors (outside the control of the teacher) which are not controlled for in this analysis (Thomas, 2001). Similarly, the residual of the intercept at Level 3 (i.e., V_{0k}) was treated to represent the effect of school on quality at the school level, but the above limitation about the several factors not controlled in this analysis should also be considered. Nevertheless, researchers within the field of EER tend to control for initial student achievement and SES differences as well as their aggregated scores at classroom and school level in order to evaluate the effects of teachers and schools (Sammons, Toth, & Sylva, 2018).

The residuals of the slope at Level 2 concerned with the impact of SES (i.e., u_{1jk}) and the impact of gender (i.e., u_{2jk}) can be treated as indicators of the teacher effect on the equity dimension in terms of SES and gender, respectively. One could extend Model 2 by adding school characteristics (i.e., factors included in the dynamic model) to investigate the extent to which variation in the within-school slopes can be partly explained. For

example, Gustafsson, Nilsen, and Yang Hansen (2018) used this approach to search for the extent to which specific school characteristics may moderate the relation between SES and mathematics achievement, but no consistent results emerged in analysing TIMSS 2011 data of different countries. For this reason, we treated each residual of the slope (i.e., u_{1jk} and u_{2jk}) as indicators of equity (in terms of SES and gender, respectively) without attempting to predict any of them by adding school and/or teacher factors in the models and searching for moderating effects since such an approach may underestimate the teacher and school effect on promoting equity. By following this approach, one could argue that a negative value of u_{1jk} shows that the SES gap on final achievement is smaller, and a negative value of u_{2jk} reveals a small gender gap on final achievement at the class level. Similarly, the residuals of the slope at Level 3 (i.e., V_{1k} and V_{2k}) refer to the school effect on the SES gap and on the gender gap on final achievement, respectively. Therefore, one could search for the relations between the residual of the intercept and each of these two slope residuals at each level and identify the extent to which teachers and/or schools which are effective in terms of quality (residual of the intercept) tend to be effective in terms of equity when either the impact of SES is considered and/or the impact of gender is taken into account. One could even search for the relation between the residual of the slope concerned with the impact of SES (i.e., u_{1jk}) and the residual of the slope concerned with the impact of gender (e.g., u_{2jk}) at the class level. In case such a positive relation exists, one could claim that teachers who are able to reduce the SES gap on final achievement may also reduce the gender gap on final achievement.

Results

Tables 1 and 2 present the results of each multilevel analysis concerned with achievement in each subject separately per study. A comparison of the findings from each empty model reveals that the four separate analyses generate similar answers about the teacher and school effects at the end of pre-primary and at the end of primary school. Moreover, the results of the four separate analyses seem to reveal that school and teacher (classroom) effects were found to be approximately equally significant in the two cognitive outcomes of schooling (i.e., mathematics and language) used to measure effectiveness in the two phases of schooling. Specifically, the pre-primary study revealed that 76.1% of the variance in language is at the student level, 14.8% at the class level, and 9.1% at the school level. Similar results emerged from analysing student achievement in mathematics since 74.5% of the variance was found to be at the student level, 15.5% at the class level, and 10.0% at the school level (see Table 1). In the primary study, 72.5% of the variance in language is at the student level, 18.0% at the class level, and 9.5% at the school level. In regard to primary student achievement in mathematics, 74.1% of the variance was found to be at the student level, 16.8% at the class level, and 9.1% at the school level (see Table 2). Moreover, the four empty models reveal that in each analysis the variance at each level reaches statistical significance ($p < 0.05$), and this implies that multilevel analysis can be used to identify the explanatory variables which are associated with achievement in each subject (Goldstein, 2003).

In Model 1, the context variables at the student, classroom, and school levels were added to the empty model. The following observations arise by comparing the results

Table 1. Parameter estimates (and standard errors) for the analysis of emergent literacy and of mathematics achievement at the end of pre-primary school.

	Emergent Literacy			Mathematics Achievement		
	Model 0	Model 1	Model 2	Model 0	Model 1	Model 2
Fixed Part						
Intercept	−.49 (.09)*	−.43 (.08)*	.38 (.09)*	.46 (.05)*	.34 (.05)*	.34 (.05)*
Student Level						
Prior achievement		.32 (.05)*	.32 (.05)*		.41 (.11)*	.40 (.11)*
SES		.19 (.07)*	.19 (.07)*		.20 (.09)*	.22 (.09)*
Gender (0 = boys, 1 = girls)		.13 (.06)*	.12 (.06)*		.06 (.03)*	.07 (.03)*
Class Level						
Average prior achievement		.10 (.04)*	.10 (.04)*		.11 (.04)*	.11 (.04)*
Average SES		.06 (.04)	.05 (.04)		.08 (.06)	.08 (.05)
School Level						
Average prior achievement		.06 (.03)*	.06 (.03)*		.08 (.03)*	.08 (.03)*
Average SES		.03 (.03)	.03 (.03)		.05 (.04)	.05 (.03)
Random Part						
Level 1 (student) random effects:						
Intercept: $\text{Var}(R_{ijk}) = \sigma^2_{eo}$	93.91 (9.12)*	37.64 (8.85)*	27.09 (8.02)*	31.52 (3.21)*	12.27 (3.01)*	11.17 (3.01)*
Level 2 (class) random effects:						
Intercept: $\text{Var}(U_{0k}) = \sigma^2_{u0}$	18.26 (3.52)*	9.93 (3.22)*	11.21 (3.12)*	6.56 (2.12)*	4.36 (1.82)*	5.11 (1.72)*
SES: $\text{Var}(U_{1k}) = \sigma^2_{u1}$			6.63 (2.53)*			3.75 (1.23)*
Gender: $\text{Var}(U_{2k}) = \sigma^2_{u2}$			4.77 (1.31)*			2.77 (1.36)*
$\text{Cov}(U_{0k}, U_{1k}) = \sigma_{v01}$			−3.15 (1.56)*			−1.67 (0.77)*
$\text{Cov}(U_{1k}, U_{2k}) = \sigma_{v12}$			1.31 (0.71)			0.65 (0.52)
$\text{Cov}(U_{0k}, U_{2k}) = \sigma_{v02}$			−0.97 (0.69)			−0.49 (0.60)
Level 3 (school) random effects:						
Intercept: $\text{Var}(V_{0k}) = \sigma^2_{v0}$	11.23 (2.08)*	7.74 (1.98)*	7.67 (1.76)*	4.23 (1.21)*	3.03 (1.12)*	2.82 (0.38)*
SES: $\text{Var}(V_{1k}) = \sigma^2_{v1}$			4.23 (1.02)*			1.63 (0.52)*
Gender: $\text{Var}(V_{2k}) = \sigma^2_{v2}$			3.21 (0.92)*			1.47 (0.63)*
$\text{Cov}(V_{0k}, V_{1k}) = \sigma_{v01}$			−1.82 (0.79)*			−0.71 (0.34)*
$\text{Cov}(V_{1k}, V_{2k}) = \sigma_{v12}$			0.91 (0.76)			0.32 (0.54)
$\text{Cov}(V_{0k}, V_{2k}) = \sigma_{v02}$			−0.67 (0.59)			−0.27 (0.51)
Total Variance	123.4	55.31	N/A	42.31	19.66	N/A
Variance Explained		68.09	N/A		22.65	N/A
Significance Test						
χ^2	1415.6	899.6	787.3	1024.3	805.2	703.1
Reduction		516.0	112.3		219.1	102.1
Degrees of freedom		5 ⁺	10		5 ⁺	10
p value		.001	.001		.001	.001

Note: SES = socioeconomic status; *statistically significant effect at the .05 level; ⁺ models were estimated without the variables that did not have a statistically significant effect at the .05 level.

from the four separate analyses. First, each analysis shows that Model 1 explains approximately 50% of the total variance of student achievement in each outcome. For example, the study conducted in pre-primary education revealed that Model 1 was able to explain 54% of the *total variance* of student achievement (i.e., 22.65 out of 42.31) in mathematics and 57% of the total variance of achievement in emergent literacy (i.e., 68.09 out of 123.40). One can also see that most of the *explained variance* is situated at the student level. For example, 83% of the explained variance of achievement in emergent literacy was found to be situated at the student level (i.e., 56.27 out of 68.09). However, approximately 30% of the *total variance* remained unexplained at the student level. For example, 30.5% of the total variance of achievement in emergent literacy (i.e., 37.64 out of 123.4) remained unexplained at the student level. Second, the effect of all background factors at the student level (i.e., SES, prior knowledge, gender)

Table 2. Parameter estimates (and standard errors) for the analysis of Greek language and of mathematics achievement at the end of primary school.

	Greek Language			Mathematics Achievement		
	Model 0	Model 1	Model 2	Model 0	Model 1	Model 2
Fixed Part						
Intercept	.42 (.06)*	.31 (.05)*	.28 (.04)*	-.41 (.09)*	-.33 (.08)*	-.23 (.08)*
Student Level						
Prior achievement		.43 (.11)*	.42 (.10)*		.38 (.05)*	.37 (.04)*
SES		.19 (.09)*	.18 (.09)*		.19 (.07)*	.18 (.07)*
Gender (0 = boys, 1 = girls)		.12 (.03)*	.11 (.03)*		.11 (.04)*	.12 (.05)*
Class Level						
Average prior achievement		.10 (.04)*	.10 (.04)*		.10 (.04)*	.10 (.04)*
Average SES		.07 (.04)	.07 (.04)		.08 (.05)	.08 (.06)
School Level						
Average prior achievement		.07 (.03)*	.08 (.03)*		.06 (.03)*	.06 (.00)*
Average SES		.08 (.05)	.07 (.05)		.09 (.06)	.09 (.06)
Random Part						
Level 1 (student) random effects:						
Intercept: $\text{Var}(R_{ijk}) = \sigma^2_{eo}$	15.38 (0.81)*	6.15 (0.72)*	5.85 (0.66)*	24.23 (0.89)*	9.52 (0.88)*	7.19 (0.85)*
Level 2 (class) random effects:						
Intercept: $\text{Var}(U_{0k}) = \sigma^2_{uo}$	3.82 (0.22)*	3.16 (0.21)*	3.09 (0.22)*	5.49 (0.51)*	4.61(0.50)*	4.02 (0.44)*
SES: $\text{Var}(U_{1k}) = \sigma^2_{u1}$			1.81 (0.27)*			2.46 (0.46)*
Gender: $\text{Var}(U_{2k}) = \sigma^2_{u2}$			1.16 (0.24)*			2.82 (0.51)*
$\text{Cov}(U_{0k}, U_{1k}) = \sigma_{v01}$			-0.91 (0.25)*			-1.21 (0.45)*
$\text{Cov}(U_{1k}, U_{2k}) = \sigma_{v12}$			0.33 (0.26)			0.61 (0.48)
$\text{Cov}(U_{0k}, U_{2k}) = \sigma_{v02}$			-0.26 (0.24)			-0.51 (0.43)
Level 3 (school) random effects:						
Intercept: $\text{Var}(V_{0k}) = \sigma^2_{v0}$	2.01 (0.32)*	1.87 (0.28)*	1.70 (0.21)*	2.98 (0.58)*	2.85 (0.55)*	2.61 (0.43)*
SES: $\text{Var}(V_{1k}) = \sigma^2_{v1}$			1.11 (0.27)*			1.61 (0.38)*
Gender: $\text{Var}(V_{2k}) = \sigma^2_{v2}$			1.17 (0.25)*			1.34 (0.31)*
$\text{Cov}(V_{0k}, V_{1k}) = \sigma_{v01}$			-0.44 (0.21)*			-0.71 (0.33)*
$\text{Cov}(V_{1k}, V_{2k}) = \sigma_{v12}$			0.23 (0.23)			0.28 (0.31)
$\text{Cov}(V_{0k}, V_{2k}) = \sigma_{v02}$			-0.19 (0.25)			-0.21 (0.29)
Total Variance	21.21	11.18	N/A	32.70	16.98	N/A
Variance Explained		10.03	N/A		15.72	N/A
Significance Test						
χ^2	924.3	605.2	517.1	915.6	499.6	395.9
Reduction		319.1	88.1		416.0	103.7
Degrees of freedom		5	10		5	10
p value		.001	.001		.001	.001

Note: Each model was estimated without the variables that did not have a statistically significant effect at the .05 level; *statistically significant effect at the .05 level; SES = socioeconomic status.

was found to be statistically significant at the 0.05 level. Third, both studies reveal that prior knowledge (i.e., aptitude) has the strongest effect in predicting student achievement at the end of the school year in each subject. Prior knowledge was also found to be the only contextual variable which had a consistent effect on student achievement in each subject when aggregated either at the classroom or the school level. This finding is in line with the results of various effectiveness studies conducted in Cyprus (e.g., Antoniou & Kyriakides, 2013; Kyriakides, Creemers, Antoniou, Demetriou, & Charalambous, 2015). Finally, the deviance likelihood statistic shows a significant change between the empty model and Model 1 ($p < 0.001$), which justifies the selection of Model 1.

In the next step of the analysis, for each dependent variable, a random-slope model was run to find out whether SES and/or gender have any differential “effect” at the classroom and/or school level (see fourth and seventh columns of Tables 1 and 2). Both studies reveal that for each subject, all parameter estimates of the variances at the random part of Model 2 were statistically significant at level .05. Moreover, the likelihood

Table 3. Correlation coefficients between intercept and SES slope at the classroom and school level.

	Pre-Primary Study		Primary Study	
	Language	Mathematics	Language	Mathematics
School Level	$r = -0.32$ $N = 76$ $p = 0.004$	$r = -0.33$ $N = 76$ $p = 0.005$	$r = -0.32$ $N = 50$ $p = 0.023$	$r = -0.35$ $N = 50$ $p = 0.013$
Classroom Level	$r = -0.36$ $N = 141$ $p < 0.001$	$r = -0.38$ $N = 141$ $p < 0.001$	$r = -0.37$ $N = 108$ $p < 0.001$	$r = -0.38$ $N = 108$ $p < 0.001$

statistic (χ^2) shows a statistically significant change between Model 1 and Model 2 ($p < 0.001$) meaning that there is a differential effect of SES and gender at both the classroom and school level both in pre-primary and primary schools. By taking into account the random part at the school level, the correlation between the residual of the intercept (V_0) and the residual of the slope for SES (V_1) for each subject was estimated (see Table 3). Both studies reveal that there is a statistically significant relation at the 0.05 level in each subject. This implies that the achievement gap (in two different subjects) based on SES tends to be smaller both in pre-primary and primary schools which achieve better learning outcomes after controlling for contextual factors (including prior knowledge). By taking into account the random part of the model at the classroom level, it was also possible to estimate the correlation between the relevant residuals at the classroom level (i.e., U_0 and U_1). Both studies reveal that for each subject, these correlation coefficients were statistically significant at the .05 level (see Table 3). These findings show that the achievement gaps based on SES tend to be smaller in pre-primary and primary classrooms found to be more effective in terms of the quality dimension (i.e., the overall student achievement, after controlling for the effect of all contextual factors). However, the correlations between the residual of the intercept and the residual of the slope for gender for each subject at the school (i.e., the relation between V_0 and V_2) and at the classroom (i.e., U_0 and U_2) level were not statistically significant at the .05 level. This implies that the achievement gap (in two different subjects) based on gender does not tend to be smaller both in pre-primary and primary schools and classrooms which achieve better learning outcomes (after controlling for contextual factors). Similarly, no statistically significant relation between the residuals of the two slopes either at the school (i.e., V_1 and V_2) or the classroom (i.e., U_1 and U_2) level was found.

Discussion

This paper searches for differential teacher and school effectiveness in relation to SES and gender in an attempt to explore the relationship between quality and equity in education. To investigate this question, we made use of data on achievement in two different subjects (i.e., mathematics and language) of students of two different phases of schooling (i.e., pre-primary and primary education). In this way, we search for the extent to which similar results emerge from four separate analyses exploring the relation of quality and equity when achievement in two different learning outcomes and of two different age groups of students were collected in order to measure these two dimensions of teacher and school effectiveness. We found similar results from the four

separate analyses exploring the relationship of the two dimensions of effectiveness. Specifically, none of the two effectiveness studies presented in the results section (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009) reveals any negative relation between the two dimensions of effectiveness (i.e., quality and equity). Gaps in achievement based on background factors were not found to be bigger in pre-primary and primary classrooms and schools in Cyprus which could be considered more effective in terms of promoting student learning outcomes. These findings seem to provide empirical support to those who argue that promotion of equity may have a positive impact on quality (Kelly, 2012; Kyriakides et al., 2018). In addition, the secondary analyses of both studies reveal that the SES gap in student achievement in the two core subjects (i.e., mathematics and language) tends to be smaller in classrooms and schools which are more effective than others in promoting student learning outcomes. Thus, the secondary analyses of the two studies presented in this paper seem to reveal that the SES gap tends to be smaller in pre-primary and primary classrooms and schools in Cyprus which achieve better learning outcomes. It should however be acknowledged that, although we found that the SES gap is smaller in effective classrooms and schools, the secondary analyses of the two studies reported here do not reveal that the effect of SES on achievement disappears in these classrooms and schools. Another limitation of this study has to do with the approach used to measure teacher and school effects. One could claim that treating the residuals of the intercept at Levels 2 and 3 as indicators of teacher and school effectiveness in terms of quality may bias the estimation of the correlation between quality and equity due to the fact that variations in average achievements between classrooms and between schools can be attributed not only to teacher and school effectiveness but to several other factors (Thomas, 2001). In fact, there is a debate on the appropriateness of using this approach to evaluate effectiveness (see Creemers et al., 2010). However, this approach to measure effectiveness is used by most researchers in the field of educational effectiveness. It is also important to note that the secondary analyses of the two studies reported here collected data on teacher and school factors included in the dynamic model, and when measures of these factors were entered in the multilevel model a very large percentage (more than 75%) of variance in achievement at the teacher and school levels was explained (see Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2008, 2009).

These two studies seem to reveal a different picture when equity is examined in regard to the gender gap rather than SES gap. Although both studies revealed differential teacher and school effects in terms of the impact of gender on student achievement, the gender gap was not found to be smaller in schools and/or classrooms which were more effective in terms of promoting overall student achievement gains in each subject. This implies that in investigating the relation between the two dimensions of effectiveness, one should make explicit which of the background factors is taken into account in measuring equity. This argument is also supported by the fact that the covariance of the residuals of the two random slopes at each level, that is, $cov(u_{1jk}, u_{2jk})$ and $cov(v_{1k}, v_{2k})$, was not statistically significant at the .05 level. This result implies that in those classrooms and schools where the SES gap on achievement is smaller, the gender gap is not necessarily smaller. Therefore, the secondary analyses of these two studies seem to reveal that teachers and schools which could be considered effective when the equity dimension in terms of the SES gap is taken into account, do not tend to

be effective in terms of equity when the gender gap is taken into account. Although this finding is attributed below to the national and/or local policies on promoting equity in Cyprus, we argue for the importance of generating mechanisms measuring equity in a systematic way. In this way, we could provide a comprehensive feedback to schools and teachers to help them develop strategies and action plans addressing specific elements of equity. For example, teachers and schools which were found to be less effective when the gender gap was considered but not when the SES gap was taken into account, may have to develop different strategies and action plans to improve their own effectiveness status in terms of equity, rather than those who were found to be less effective in terms of equity when both SES and gender gaps were taken into account.

It should finally be stressed that the findings of the secondary analyses of the two effectiveness studies reported here reveal similar patterns in terms of the relationship between the two dimensions of effectiveness (i.e., quality and equity) in Cypriot pre-primary and primary schools when achievement in two different subjects is considered. This implies that investigating the relation between promoting student learning outcomes (quality) and reducing the impact of background factors on achievement (equity) might be less dependent on the learning outcome that is considered (i.e., mathematics or language) than the background factor that is taken into account in measuring equity. The fact that similar results emerged by measuring effectiveness in terms of two different subjects (i.e., mathematics and language) and in two different phases of schooling, seems to be in line with previous effectiveness studies conducted in Cyprus which revealed that teachers who are effective in promoting student achievement gains in a specific subject tend to be also effective in promoting achievement in another subject (e.g., Kyriakides, 2005; Kyriakides & Creemers, 2008). However, the reanalyses of these two effectiveness studies (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009) seem to reveal that there is no simple answer to the question about the relation between the two dimensions of effectiveness as they have been defined in this paper. In the pre-primary and primary schools of Cyprus, the relation between the two dimensions of effectiveness seems to be more dependent on the student background factor that is considered in searching for learning gaps between different groups of students (i.e., equity), and less on the learning outcome that is taken into account. When SES was considered, it was found that the SES gap tends to be smaller in those classrooms and schools which achieve better overall learning outcomes in each subject, but such a relation was not observed when the gender gap was considered.

Implications of findings for policy and practice and suggestions for further research can also be drawn. One could claim that in exploring the relation between promoting student learning outcomes (quality) and reducing the gap of background factors on student outcomes (i.e., equity) in Cyprus, policies on equal educational opportunities at the national and/or local level should be considered. Recent national policy guidelines promoting equity at the classroom and school levels stress mainly the impact that SES has on student achievement (Ministry of Education and Culture, 2015) and do not give special emphasis to the importance of addressing other student background factors in promoting equity (and especially gender). A similar national policy could be observed by analysing policy documents on equity produced by the Ministry of Education in the past when different political parties were in power (e.g., Ministry of Education and Culture, 2015). One could therefore argue that there is a need to develop a comprehensive policy

on promoting equal educational opportunities for different groups of students in pre-primary and primary schools of Cyprus addressing not only the impact of SES but also the impact of other background factors such as gender and ethnicity.

At the same time, these findings seem to reveal the importance of evaluating teacher and school effectiveness in terms of different aspects of the equity dimension rather than focusing only on equity in terms of SES. By establishing formative evaluation mechanisms, the importance of promoting equity in terms of not only SES but also the other background factors in Cypriot classrooms and schools may be made explicit. Similar findings may emerge in other countries where national and/or local policies on equity address specific background factors only. Thus, measuring teacher and school effectiveness in terms of the equity dimension by considering the various background factors that are not likely to change (e.g., SES, gender, ethnicity) may help educational systems not only to develop more comprehensive policies on equal opportunities, but also to evaluate the impact of their national policies on promoting equity.

Implications for EER are finally drawn. In the two studies by Creemers and Kyriakides (2010) and Kyriakides and Creemers (2009), data on student achievement (initial and final) during a single school year were collected. One cannot therefore search for changes and/or stability in the effectiveness status of teachers and schools in terms of quality and equity. In evaluating national and/or school reform policies on equal educational opportunities, data on student achievement during more than one school year could be collected to find out whether the effectiveness status of schools and/or teachers in terms of quality and/or equity has been improved. Suggestions for further research searching for stability and/or changes in the effectiveness status of schools in quality as well as in equity can also be provided. By collecting data on two or even more consecutive school years, the effectiveness status of schools in terms of quality and equity can be measured over time. It will then be possible to search for the extent to which teacher and school factors included in the dynamic model explain changes in the school effectiveness status in terms of each dimension of effectiveness. Such studies may also reveal whether similar changes in the effectiveness status of schools and/or teachers are observed when quality and equity in terms of different background factors are measured over time. These studies may also contribute to understanding the dynamic nature of effectiveness as well as to investigating not only the short-term but also the long-term effect of teachers and schools in promoting both quality and equity (Kyriakides et al., 2018).

In this paper, we reanalysed the data of two studies (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009) conducted in a single country and searched for the relation between the quality and the equity dimensions of effectiveness. Similar national studies and/or reanalyses of existing effectiveness studies are needed to find out the extent to which the reported findings can be generalized. The fact that we did not detect any negative relation between the quality and the equity dimension (irrespective of how it was measured) should be examined further by collecting and/or reanalysing data in different countries and phases of education. In this way, we can test further the arguments about possible negative relations between the two dimensions of effectiveness. At the same time, further studies conducted in different educational contexts are needed to find out whether the measurement of equity in terms of different student background factors is needed. Furthermore, international studies are needed to investigate differential

effectiveness at the system level in terms of various student background factors (e.g., SES, gender, and ethnicity). Such studies may also search for the extent to which national and/or school factors included in theoretical models of educational effectiveness research moderate the relation between SES and student achievement (see Gustafsson et al., 2018). In this way, we are not only able to search for factors that may reduce the relation between SES and achievement, but also to investigate in a more systematic way the relation between quality and equity at classroom, school, and system level.

Finally, we provide suggestions for further reanalyses of these two studies (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009) to measure the relation between quality and equity. It is acknowledged that different indicators for measuring equity at class and school level, such as the Gini-based Attainment Index, could have been employed. For example, Kelly (2012) analysed data from the National Pupil Database in England by applying the Index to more than 20,000 students with matched attainment records at two different assessment periods (i.e., Key Stage 2 and Key Stage 4) in two “statistical-neighbour” local authorities in England. They were then combined with existing contextual value-added measures to analyse school and local authority performance in terms of both quality and equity. By comparing the results of this approach with the methodology used in this paper to measure quality and equity, we can search for the impact that the use of a specific methodology may have on searching for the relationship between the two dimensions of effectiveness. However, we were not in a position to analyse the data of the two studies reported here (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009) by following the method proposed by Kelly (2012), since the use of the Gini-based Attainment Equity Index is appropriate when a specific threshold can emerge from the measurement of student achievement. Kelly (2015) provides an overview of the several indices that have the potential to measure school effectiveness (including the Gini-based Attainment Equity Index) in terms of equity, and the strengths and limitations of each index are identified. Further research on employing different methods to measure not only equity but also the relation of equity and quality is needed. In this context, this paper argues for the importance of conducting research in differential teacher and school effectiveness for investigating the relation between quality and equity. So far, researchers in the field of EER have made use of studies searching for differential effectiveness to identify whether specific factors can be treated as generic or whether they have differential effects (see Kyriakides, 2007). By reanalysing these two effectiveness studies (Creemers & Kyriakides, 2010; Kyriakides & Creemers, 2009), we argue that studies on differential teacher and school effectiveness can address issues about the relation between quality and equity and may contribute to designing school-based interventions aiming to promote both dimensions of effectiveness (Charalambous et al., 2018).

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Leonidas Kyriakides is Professor of Educational Research and Evaluation at the Department of Education of the University of Cyprus, Cyprus. His main research interests are in the area of school effectiveness and school improvement and especially in modelling the dynamic nature of

educational effectiveness and in using research to promote quality and equity in education. Leonidas acted as chair of the EARLI SIG on Educational Effectiveness and as chair of the AERA SIG on School Effectiveness and Improvement.

Bert (Hubertus Peter Maria) Creemers is Professor emeritus in educational sciences at the University of Groningen, The Netherlands. His main interest is educational quality and equity in terms of students' learning outcomes at classroom, school, and system level. Together with Leonidas Kyriakides, he works on the development and testing of the dynamic model on educational effectiveness and the approach based on this model for the improvement of education.

Evi Charalambous is a PhD student at the Department of Education of the University of Cyprus. She has participated in several international projects, and her main research interests are in the area of educational effectiveness research and school improvement for promoting quality and equity in education.

References

- Antoniou, P., & Kyriakides, L. (2011). The impact of a dynamic approach to professional development on teacher instruction and student learning: Results from an experimental study. *School Effectiveness and School Improvement*, 22(3), 291–311. doi:10.1080/09243453.2011.577078
- Antoniou, P., & Kyriakides, L. (2013). A dynamic integrated approach to teacher professional development: Impact and sustainability of the effects on improving teacher behaviour and student outcomes. *Teaching and Teacher Education*, 29, 1–12. doi:10.1016/j.tate.2012.08.001
- Atkinson, A. B. (2015). *Inequality: What can be done?* Cambridge, MA: Harvard University Press.
- 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. doi:10.1016/j.ijedudev.2016.07.004
- Belfi, B., Haelermans, C., & De Fraine, B. (2016). The long-term differential achievement effects of school socioeconomic composition in primary education: A propensity score matching approach. *British Journal of Educational Psychology*, 86(4), 501–525. doi:10.1111/bjep.12120
- Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. doi:10.1007/s11092-008-9068-5
- Bock, R. D., & Aitken, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46(4), 443–459. doi:10.1007/BF02293801
- Brekelmans, M., Slegers, P., & Fraser, B. (2000). Teaching for active learning. In R. J. Simons, J. van der Linden, & T. Duffy (Eds.), *New learning* (pp. 227–242). Dordrecht: Kluwer Academic.
- Campbell, R. J., Kyriakides, L., Muijs, R. D., & Robinson, W. (2003). Differential teacher effectiveness: Towards a model for research and teacher appraisal. *Oxford Review of Education*, 29(3), 347–362. doi:10.1080/03054980307440
- Campbell, R. J., Kyriakides, L., Muijs, R. D., & Robinson, W. (2004). Effective teaching and values: Some implications for research and teacher appraisal. *Oxford Review of Education*, 30(4), 451–465. doi:10.1080/0305498042000303955
- Charalambous, E., Kyriakides, L., & Creemers, B. P. M. (2018). Promoting quality and equity in socially disadvantaged schools: A group-randomisation study. *Studies in Educational Evaluation*, 57, 42–52. doi:10.1016/j.stueduc.2016.06.001
- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London: Routledge/Falmer.
- 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. doi: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*. Abingdon: Routledge.

- Creemers, B. P. M., & Kyriakides, L. (2010). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement*, 21(4), 409–427. doi:10.1080/09243453.2010.512795
- 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. doi:10.1080/09243453.2013.869233
- Creemers, B. P. M., Kyriakides, L., & Sammons, P. (2010). *Methodological advances in educational effectiveness research*. London: Routledge.
- Driessen, G. (2007). "Peer group" effecten op onderwijsprestaties: Een internationaal review van effecten, verklaringen en theoretische en methodologische aspecten ["Peer group" effects on academic achievement: An international review of effects, explanations, and theoretical and methodological issues]. Retrieved from <http://its.ruhosting.nl/publicaties/pdf/r1730.pdf>
- Field S., Kuczera, M., & Pont, B. (2007). *No more failures: Ten steps to equity in education*. Paris: OECD.
- Fielding, M. (1997). Beyond school effectiveness and school improvement: Lighting the slow fuse of possibility. In J. White & M. Barber (Eds.), *Perspectives on school effectiveness and school improvement* (pp. 137–160). London: Institute of Education.
- Frempong, G., Reddy, V., & Kanjee, A. (2011). Exploring equity and quality education in South Africa using multilevel models. *Compare*, 41(6), 819–835. doi:10.1080/03057925.2011.607488
- Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.
- Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). London: Edward Arnold.
- Gulson, K. N., & Webb, P. T. (2012). Education policy racialisations: Afrocentric schools, Islamic schools, and the new enunciations of equity. *Journal of Education Policy*, 27(6), 697–709. doi:10.1080/02680939.2012.672655
- Gustafsson, J.-E., Nilsen, T., & Yang Hansen, K. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation*, 57, 16–30. doi:10.1016/j.stueduc.2016.09.004
- Hattie, J. A. C. (2002). Class composition and peer effects. *International Journal of Educational Research*, 37(5), 449–481. doi:10.1016/S0883-0355(03)00015-6
- Hornstra, L., Van der Veen, I., Peetsma, T., & Volman, M. (2015). Does classroom composition make a difference: Effects on developments in motivation, sense of classroom belonging, and achievement in upper primary school. *School Effectiveness and School Improvement*, 26(2), 125–152. doi:10.1080/09243453.2014.887024
- Joyce, B. R., Weil, M., & Calhoun, E. (2000). *Models of teaching*. Boston, MA: Allyn & Bacon.
- Kelly, A. (2012). Measuring "equity" and "equitability" in school effectiveness research. *British Educational Research Journal*, 38(6), 977–1002. doi:10.1080/01411926.2011.605874
- Kelly, A. (2015). Measuring equity in educational effectiveness research: The properties and possibilities of quantitative indicators. *International Journal of Research & Method in Education*, 38(2), 115–136. doi:10.1080/1743727X.2014.914486
- Kolen, M. J., & Brennan, R. L. (1995). *Test equating: Methods and practices*. New York, NY: Springer.
- Kyriakides, L. (2002). A research-based model for the development of policy on baseline assessment. *British Educational Research Journal*, 28(6), 805–826. doi:10.1080/0141192022000019071
- Kyriakides, L. (2005). Extending the comprehensive model of educational effectiveness by an empirical investigation. *School Effectiveness and School Improvement*, 16(2), 103–152. doi:10.1080/09243450500113936
- Kyriakides, L. (2007). Generic and differentiated models of educational effectiveness: Implications for the improvement of educational practice. In T. Townsend (Ed.), *International handbook of school effectiveness and improvement* (pp. 41–56). Dordrecht: Springer.
- Kyriakides, L. (2008). Testing the validity of the comprehensive model of educational effectiveness: A step towards the development of a dynamic model of effectiveness. *School Effectiveness and School Improvement*, 19(4), 429–446. doi:10.1080/09243450802535208
- 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. doi:10.1016/j.tate.2013.07.010

- Kyriakides, L., & Creemers, B. P. M. (2008). A longitudinal study on the stability over time of school and teacher effects on student learning outcomes. *Oxford Review of Education*, 34(5), 521–545. doi:10.1080/03054980701782064
- Kyriakides, L., & Creemers, B. P. M. (2009). The effects of teacher factors on different outcomes: Two studies testing the validity of the dynamic model. *Effective Education*, 1(1), 61–85. doi:10.1080/19415530903043680
- Kyriakides, L., Creemers, B., 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. doi:10.1080/01411920903165603
- Kyriakides, L., Creemers, B. P. M., Antoniou, P., Demetriou, D., & Charalambous, C. (2015). The impact of school policy and stakeholders' actions on student learning: A longitudinal study. *Learning and Instruction*, 36, 113–124. doi:10.1016/j.learninstruc.2015.01.004
- Kyriakides, L., Creemers, B., & Charalambous, E. (2018). *Equity and quality dimensions in educational effectiveness*. Dordrecht: Springer.
- Kyriakides, L., & Kelly, K. L. (2003). The impact of engagement in large-scale assessment on teachers' professional development: The Emergent Literacy Baseline Assessment project. *Journal of Research in Childhood Education*, 18(1), 18–36. doi:10.1080/02568540309595020
- Levin, B. (2005). *Governing education*. Toronto: University of Toronto Press.
- Lim, L. (2013). Meritocracy, elitism, and egalitarianism: A preliminary and provisional assessment of Singapore's primary education review. *Asia Pacific Journal of Education*, 33(1), 1–14. doi:10.1080/02188791.2012.711294
- Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation*, 12(2), 105–128. doi:10.1080/13803610600587040
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with Assessment for Learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133–149. doi:10.1080/02671520600615638
- Masters, G. N., & Wright, B. D. (1997). The partial credit model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 101–121). New York, NY: Springer.
- McCoy, S. K., & Major, B. (2007). Priming meritocracy and the psychological justification of inequality. *Journal of Experimental Social Psychology*, 43(3), 341–351. doi:10.1016/j.jesp.2006.04.009
- Micklewright J., & Schnepf, S. V. (2007). Inequality of learning in industrialized countries. In S. P. Jenkins & J. Micklewright (Eds.), *Inequality and poverty re-examined* (pp. 129–145). Oxford, UK: Oxford University Press.
- Ministry of Education and Culture. (2015). *DRASE-school actions and strategies for social inclusion* (Circular YPP3191, 7. 1.10.6/9). Nicosia: Author.
- 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. doi:10.1080/09243453.2014.885451
- Organisation for Economic Co-operation and Development. (2012). *Equity and quality in education: Supporting disadvantaged students and schools*. Paris: Author.
- Organisation for Economic Co-operation and Development. (2013). *PISA 2012 results: Excellence through equity: Giving every student the chance to succeed (Volume II)*. Paris: Author.
- Palardy, G. J. (2015). Classroom-based inequalities and achievement gaps in first grade: The role of classroom context and access to qualified and effective teachers. *Teachers College Record*, 117(2), 1–48.
- Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in the first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis*, 30(2), 111–140. doi:10.3102/0162373708317680
- Panayiotou, A., Kyriakides, L., Creemers, B. P. M., McMahon, L., Vanlaar, G., Pfeifer, M., ... Bren, M. (2014). Teacher behavior and student outcomes: Results of a European study. *Educational Assessment, Evaluation and Accountability*, 26(1), 73–93. doi:10.1007/s11092-013-9182-x
- Paquette, J. (1998). Equity in educational policy: A priority in transformation or in trouble? *Journal of Education Policy*, 13(1), 41–61. doi:10.1080/0268093980130104

- Psacharopoulos, G. (2007). *The costs of school failure: A feasibility study* (European Expert Network on Economics of Education (EENEE) Analytical Report No. 2 prepared for the European Commission). Retrieved from <http://www.eenee.de/eeneeHome/EENEE/Analytical-Reports.html>
- Rasbash, J., Steele, F., Browne, W., & Prosser, B. (2005). *A user's guide to MLwin Version 2.0*. Bristol: University of Bristol.
- 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. doi:10.1080/09243453.2014.885450
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction*, 32, 63–72. doi:10.1016/j.learninstruc.2014.01.007
- Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record*, 107(9), 1999–2045.
- Sammons, P. (2010). Equity and educational effectiveness. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (3rd ed., Vol. 5, pp. 51–57). Oxford, UK: Elsevier.
- Sammons, P., Toth, K., & Sylva, K. (2018). The drivers of academic success for “bright” but disadvantaged students: A longitudinal study of AS and A-level outcomes in England. *Studies in Educational Evaluation*, 57, 31–41. doi:10.1016/j.stueduc.2017.10.004
- Scheerens, J. (2013). The use of theory in school effectiveness research revisited. *School Effectiveness and School Improvement*, 24(1), 1–38. doi:10.1080/09243453.2012.691100
- Scheerens, J. (2016). *Educational effectiveness and ineffectiveness: A critical review of the knowledge base*. Dordrecht: Springer.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75(3), 417–453. doi:10.3102/00346543075003417
- Smith, S. S. (1997). *Early childhood mathematics*. Needham Heights, MA: Allyn & Bacon.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: SAGE.
- Stipek, D. (2004). Teaching practices in kindergarten and first grade: Different strokes for different folks. *Early Childhood Research Quarterly*, 19(4), 548–568. doi:10.1016/j.ecresq.2004.10.010
- Thomas, S. (2001). Dimensions of secondary school effectiveness: Comparative analyses across regions. *School Effectiveness and School Improvement*, 12(3), 285–322. doi:10.1076/sesi.12.3.285.3448
- Thrupp, M. (2001). Recent school effectiveness counter-critiques: Problems and possibilities. *British Educational Research Journal*, 27(4), 443–457. doi:10.1080/01411920120071452
- Van Damme, J., & Bellens, K. (2017). Countries strive towards more quality and equity in education: Do they show success or failure? Evidence from TIMSS 2003 and 2011, for grade 4. In M. Rosén, K. Yang Hansen, & U. Wolff (Eds.), *Cognitive abilities and educational achievement: Measurement and determinant. A Festschrift in Honour of Jan-Eric Gustafsson* (pp. 127–148). Dordrecht: Springer.
- Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5(2), 134–150. doi:10.1016/j.edurev.2010.02.001
- Verhaeghe, J. P., Vanlaar, G., Knipprath, H., De Fraine, B., & Van Damme, J. (2018). Can group composition effects explain socioeconomic and ethnic achievement gaps in primary education? *Studies in Educational Evaluation*, 57, 6–15. doi:10.1016/j.stueduc.2017.07.006
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450. doi:10.1007/BF02294627
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3), 461–481. doi:10.1037/0033-2909.91.3.461
- Whitty, G. (2001). Education, social class and social exclusion. *Journal of Education Policy*, 16(4), 287–295. doi:10.1080/02680930110054308
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1998). *ConQuest: Generalized item response modelling software*. Melbourne: Australian Council for Educational Research (ACER).
- Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the Early Childhood Longitudinal Study – Kindergarten class of 1998–1999. *American Educational Research Journal*, 41(1), 191–229. doi:10.3102/00028312041001191