Using the dynamic approach to school improvement to promote quality and equity in education

Kyriakides, Leonidas; Charalambous, Evi ; Creemers, Bert P. M.; Antoniou, Panayiotis; Devine, Dympna; Papastylianou, Dona; Fahie, Declan

Published in:
Educational assessment evaluation and accountability

DOI:
10.1007/s11092-018-9289-1

Educational Assessment, Evaluation and Accountability (2019) 31:121–149
https://doi.org/10.1007/s11092-018-9289-1

Using the dynamic approach to school improvement to promote quality and equity in education: a European study

Leonidas Kyriakides[1] · Evi Charalambous[1] · Bert P. M. Creemers[2] · Panayiotis Antoniou[3] · Dympna Devine[4] · Dona Papastylianou[5] · Declan Fahie[4]

Received: 6 April 2018 / Accepted: 15 October 2018 / Published online: 30 October 2018

Abstract

This paper presents the results of a European study investigating the extent to which the Dynamic Approach to School Improvement (DASI) can help schools situated in socially disadvantaged areas to improve their effectiveness. At the beginning of the school year 2015–2016, a sample of 72 primary schools in four European countries (Cyprus, England, Greece and Ireland) was randomly allocated into the experimental and control groups. A questionnaire measuring the functioning of school factors related with the school learning environment, school policy for teaching and school evaluation was administered to all teachers of the school sample ($n = 762$). A battery of mathematics tests and a questionnaire measuring students' socioeconomic status (SES) were administered to all students of grades 4–6 of the school sample ($n = 5560$). The experimental group made use of DASI to develop improvement strategies and action plans. Feedback was provided to the control group regarding their students' achievement and the functioning of school factors in their school. In each country, DASI had an effect on promoting student learning outcomes. For the control group of each country, the total effect of SES on student achievement at the end of the intervention was bigger than the effect of SES at the beginning of the intervention. No increase in the effect of SES was identified in the schools of the experimental group. Implications of findings for establishing a theory-driven and evidence-based approach to improve the quality and the equity dimensions of school effectiveness are discussed and suggestions for future studies are provided.

Keywords Educational effectiveness research · School improvement · Equity in education · Group randomisation study · Multilevel modelling techniques

✉ Leonidas Kyriakides
kyriakid@ucy.ac.cy

Extended author information available on the last page of the article

1 Introduction

Education is a powerful mechanism for improving several aspects of a person's life, including socio-economic standing and welfare (Kyriakides et al. 2018). School failure has a negative long-lasting impact on a child's life since leaving school without sufficient qualifications could result in difficulties to secure equal participation in the financial, civic and social aspects of modern society (Micklewright and Schnepf 2007). However, children are not all equal when it comes to education failure. Evidence shows that children coming from socially disadvantaged backgrounds are more likely to have lower school results and to drop out of school more frequently than children coming from better-off families (Kyriakides et al. 2018; Sirin 2005). International evaluation studies like PISA revealed that in Europe approximately 20% of students are not equipped with basic skills in mathematics and that a 15-year-old student from a relatively disadvantaged home is 2.37 times more likely to be a poor performer (obtaining a score below level 2 that measures basic skills in mathematics) than a student from an affluent family (see OECD 2012). PISA studies also report that 40% of the variation in student performance is found between schools within a country. This suggests that issues of both quality and equity are at play in these schools and that interventions aiming to improve the quality of underperforming schools are needed. Moreover, research shows that interventions supporting primary school students who are at risk have stronger effects than those addressing students at secondary school level (Creemers and Kyriakides 2012; Scheerens and Bosker 1997; Townsend 2007). Furthermore, various syntheses of effectiveness programmes aiming to improve the attainment of primary students with low basic skills reveal that whole school interventions are more effective (e.g. Borman et al. 2003; Hattie 2009). It should however be acknowledged that international large-scale studies also reveal that educational outcomes vary widely between countries. For example, the education systems in Korea and Shanghai, China not only have above-average performance in reading but also rather small differences between the scores of the highest- and the lowest-achieving students. Similarly, in 10 out of the 17 countries that have above-average performance in reading, variations in performance are smaller than the average variation observed across OECD countries (OECD 2013). Another important finding from the international evaluation studies is that average achievement levels among 15-year-old students between the top- and bottom-performing countries differ significantly by more than two standard deviations (corresponding of 6–8 years of learning) (Woessmann 2016). Considerably, the study reported in this paper has been conducted in four European countries (i.e. Cyprus, England, Greece and Ireland) to find out whether a whole school improvement approach can contribute in promoting not only quality but also equity in different educational settings.

2 Research aims

Given that children in socially disadvantaged areas are more likely to have lower achievement levels, the study reported here investigated the extent to which a whole school intervention based on the *Dynamic Approach to School Improvement* (DASI) (Creemers and Kyriakides 2012) could promote both quality and equity in socially

disadvantaged primary schools in four European countries. This was done by investigating the extent to which schools and teachers were able to reduce the gaps in schooling outcomes among students with different socio-economic background characteristics. The basic rationale is that education can contribute to social justice and democracy by closing the gap in learning outcomes between students from different socioeconomic backgrounds, particularly when their abilities and the socio-cultural status of their family are taken into consideration (Kelly 2012; Sammons et al. 2018). Specifically, this project aimed to develop further the DASI and evaluate its impact on (1) improving the functioning of school factors, (2) promoting student achievement gains (*quality dimension of school effectiveness*) and (3) reducing the impact of the socioeconomic status (SES) on student achievement in mathematics (*equity dimension of school effectiveness*). Thus, the rationale and the main steps of DASI are presented in the next section of this paper. The importance of using DASI to promote not only quality but also equity in education is also discussed.

At this point, it should be noted that the four European countries participating in this study were chosen due to the fact that the promotion of equity is considered a priority in each one of these countries (see OECD 2010, 2012). It should also be acknowledged that the dynamic model of educational effectiveness (see Creemers and Kyriakides 2008), which is the theoretical framework upon which the intervention of the present study is based, has been empirically tested in these four countries through several international and national studies (e.g. Kyriakides et al. 2015; Panayiotou et al. 2014, 2016; Christoforidou and Xirafidou 2014). In addition, a number of experimental studies have been conducted in three out of the four participating countries (i.e. Cyprus, England and Greece) in order to identify the impact of DASI on promoting student learning outcomes (e.g. Kyriakides et al. 2014). However, schools participating in the experimental studies investigating the impact of DASI on promoting quality in education were not situated in socially disadvantaged areas. Given that early effectiveness studies of the field of Educational Effectiveness Research (EER) were concerned with identifying ways to help schools in disadvantaged areas to achieve better learning outcomes (Edmonds 1979; Rutter et al. 1979), it is important to find out whether DASI can help schools in these areas to become more effective. In addition, there is a variation between these four countries regarding the way that educational policy is applied to schools in order to support students coming from low socioeconomic backgrounds. For example, Ireland has a specific policy that focuses on providing special support and learning opportunities to these students whereas Cyprus and Greece have no clear policy on promoting equity in education. Lastly, these countries have been affected by the recent economic crisis and/or treat the promotion of equity in education as a policy priority. Therefore, this study aimed to raise awareness among policy-makers and practitioners of these countries and help them emphasise on factors that can promote not only quality but also equity in education.

3 The dynamic approach to school improvement: Rationale and major steps

Creemers and Kyriakides (2015) emphasise that 'the main assumption of DASI is that school improvement projects can have an impact on student learning outcomes only

when these projects are based on valid theories' (p. 105). Researchers have stressed the importance of providing a studious and theoretical framework when designing an improvement programme (e.g. Bierman et al. 2008; Buczynski and Hansen 2010; Domitrovich et al. 2009; Yoon et al. 2007). Consequently, DASI is developed on the basis of a theoretical framework, which refers to factors of educational effectiveness that need to be carefully examined when introducing a change at the school level (Creemers and Kyriakides 2010a; Kyriakides 2017). More specifically, DASI makes use of the *dynamic model of educational effectiveness* (Creemers and Kyriakides 2008), which has been systematically tested through empirical studies and meta-analyses which revealed that the effectiveness factors included in the dynamic model were associated with student achievement gains (e.g. Antoniou and Kyriakides 2011, 2013; Azigwe et al. 2016; Azkiyah et al. 2014; Christoforidou et al. 2014; Christoforidou and Xirafidou 2014; Creemers and Kyriakides 2010b; Kyriakides et al. 2016; Kyriakides et al. 2013a, b; Kyriakides and Creemers 2008, 2009; Kyriakides et al. 2010; Panayiotou et al. 2014).

Second, 'DASI assumes that each school should develop its own strategies and action plans for improvement by taking into account the knowledge base provided by the dynamic model, which attempts to describe the complexity of educational effectiveness' (Creemers and Kyriakides 2015, p. 106). It is important to note that while the dynamic model focuses on its effectiveness factors, it simultaneously gives opportunities to school stakeholders to address their improvement priorities in an adjustable way (Creemers and Kyriakides 2010b; Heck and Moriyama 2010; Hofman et al. 2010; Sammons 2009). This is also empirically supported as the aforementioned studies testing the validity of the dynamic model revealed variation in the functioning of factors in different school settings. Therefore, this finding leads to the conclusion that one should take into account the abilities of the school stakeholders (i.e. teachers, students, parents and school management team) and their professional needs in the process for improving the functioning of each effectiveness factor (see Kyriakides et al. 2018).

Third, Creemers and Kyriakides (2015) argue that 'effective schooling is seen as a dynamic, ongoing process' (p. 106). This means that schools are expected to adapt to the changing contexts, needs and diverse priorities in order to be considered as effective. Likewise, less effective schools may be encouraged by the community and local school boards to take actions to improve their students' learning outcomes. This notion relies on the contingency theory (Donaldson 2001; Mintzberg 1979) and can be viewed as one of the main assumptions upon which the dynamic model is based (Scheerens 2013). Therefore, DASI is concerned with the process of improving the functioning of the school level factors in each school. Moreover, DASI is based on the assumption that even schools which are among the most effective should take actions to improve further their policy for teaching and their school learning environment in order to remain effective (Creemers and Kyriakides 2010b).

Fourth, DASI stresses that each school should develop its own strategies and improvement plans, but support to schools should be offered by an Advisory and Research Team (A&R Team), who has the appropriate technical expertise and theoretical knowledge base for addressing the school level factors in order to help schools improve their policy for teaching and their policy for creating a school learning environment (Creemers and Kyriakides 2015; Kyriakides 2017). This means that each

school is encouraged to ask for the support of the A&R Team and of any other available resource within and/or outside the school when designing and implementing their action plans. In this way, a systematic research-based approach to design, implement and evaluate improvement efforts (Bryk et al. 2010; Rowan et al. 2009) is promoted by DASI (Kyriakides 2017).

Fifth, DASI is based on the assumptions that both the A&R Team and the school stakeholders should be involved in the design and the evaluation of the improvement project (Creemers and Kyriakides 2015). In this way, we avoid the problems that may arise in conducting experimental studies where practitioners see themselves as those who are expected to follow in a rather strict way an intervention designed by others, and thereby ownership is not established. In the case of DASI, school stakeholders are those who take decisions on which improvement actions and tasks should be carried out. By using this approach, not only ownership of the improvement project is likely to be established but also the stakeholders' experiences and the special situation of the school are taken into account (Creemers and Kyriakides 2012). At the same time, the A&R Team has an important role to play since this team is expected to share its expertise and knowledge with practitioners and help them develop strategies and action plans that are in line with the relevant knowledge base of EER.

Finally, DASI is concerned with the impact of the improvement project on student learning outcomes and thereby a summative evaluation is undertaken searching for the impact of the intervention on student learning outcomes. This implies that DASI promotes an innovative approach to improvement where stakeholders and the A&R Team are expected to collaborate in order to develop improvement strategies and action plans, implement those and evaluate their impact on student learning outcomes. At the same time, each party has a specific role and expertise to bring to the improvement intervention, so they are not expected to play the same role in the school improvement process.

Figure 1 illustrates the major steps of DASI and reveals that school stakeholders and the A&R Team should be actively involved in each step of DASI. Their ability to work together and exchange skills, expertise and experience is critical to the success of the school improvement project. The first two steps are concerned with the establishment of clarity and consensus about the general purpose and the aims of the school improvement project. Initially, the importance of promoting student learning is stressed (step A). This implies that school stakeholders and the A&R Team should define the aims of their intervention by taking into account the fact that promotion of student learning must be the ultimate aim of any school improvement effort (Creemers and Kyriakides 2010a). In the case of the study reported here, the importance of promoting both quality and equity was stressed at this stage of the intervention to both the experimental and the control group, so as to secure that the two groups had equivalent and sufficient information regarding the purpose of the study. In the second step, the effectiveness factors of the dynamic model of educational effectiveness are presented to school stakeholders. This may help school stakeholders understand *how* and *why* addressing these school factors may promote student learning outcomes. In the third step, stakeholders of the schools in the experimental group (with the support of the A&R Team) should develop their own evaluation mechanisms in order to collect and analyse data about the functioning of school factors and identify their priority area(s) for improvement. The fourth step shows that the A&R Team should work closely with the school stakeholders to help them develop their own improvement strategies and

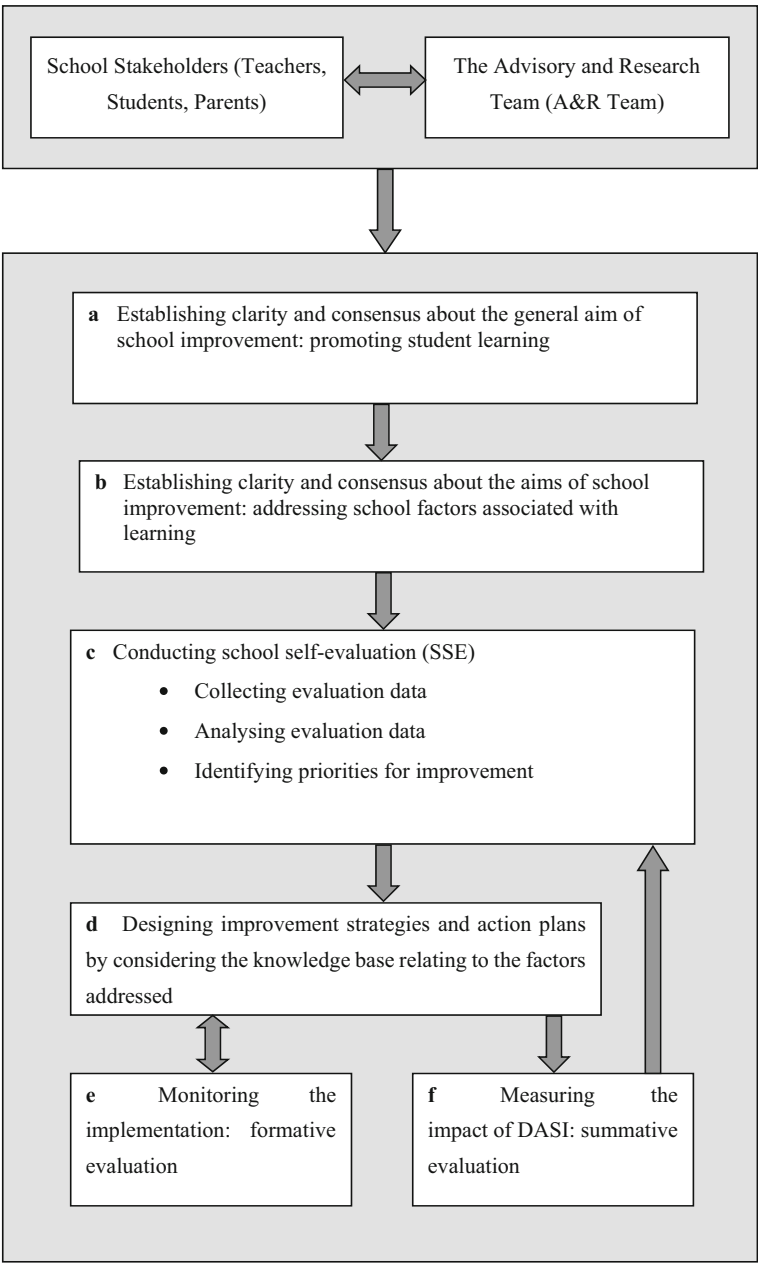

Fig. 1 The major steps in the Dynamic Approach to School Improvement (DASI)

action plans. School stakeholders are expected to take into account the available knowledge base of EER during this process. A handbook indicating actions that could be taken to improve the functioning of each school factor in relation to their improvement priorities is provided by the A&R Team. School stakeholders are expected to adjust these guidelines to their school context (with the support of the A&R Team).

Then, school stakeholders and the A&R Team should develop mechanisms for monitoring the implementation of the intervention (step E). At this point, the role of formative evaluation is stressed and the importance of using evaluation data to further develop the school improvement strategies and action plans is emphasised. Finally, the A&R Team and the school stakeholders should measure the impact of DASI on promoting student learning outcomes and reducing the learning differences between students coming from different socioeconomic backgrounds (step F). This step may also reveal new areas for improvement. In particular, if the evaluation results reveal that a school has managed substantially to improve the functioning of the factor(s) addressed during the first year of the implementation of the project, school stakeholders and the A&R Team may decide to collect evaluation data and identify a new priority improvement area. By collecting data on the functioning of school factors (i.e. moving back to step C), the new priority area may be identified and a new improvement project could be developed and implemented during the second year of the improvement project. This implies that improving school effectiveness is an ongoing and never-ending process, irrespective of how effective a school might be. Improvement efforts are continuous, cyclical in nature and embedded in a wider process of overall school development. In the next section, we advocate the use of DASI to promote not only quality but also equity.

4 Using the dynamic approach to school improvement to promote quality and equity

This section advocates for the use of DASI to promote quality and equity in education. In regard to its impact on quality, four experimental studies revealed that DASI had a stronger impact on improving student learning outcomes than the participatory approach to teacher and school improvement which gives emphasis to professional experience (see Creemers and Kyriakides 2015; Kyriakides 2017). It is important to note here that three out of the four experimental studies were concerned with the use of DASI for promoting cognitive learning outcomes. However, the fourth study took place in five European countries (i.e. Belgium, Cyprus, England, Greece and the Netherlands) and was concerned with the use of DASI for reducing bullying (Kyriakides et al. 2014). In each country, schools which made use of DASI managed to reduce bullying at a significantly higher level than the schools of the control group.

Recent research findings suggest that there are almost no data on the impact of DASI on promoting equity in education (Kyriakides et al. 2018). Moreover, the participating schools in the four studies mentioned above were not situated in socially disadvantaged areas and the main aim of these studies was to search for the impact of DASI on promoting student learning outcomes rather than reducing the SES gap in student achievement. This can be attributed to the fact that during the last three decades, EER was mainly concerned with identifying factors associated with student achievement gains rather than searching for the contribution of schools and teachers to promote equity. There is research to suggest, for example, little conscious awareness by teachers of the equity dimension to their role, in spite of observed differences in their practices and expectations with different groups of children on the basis of SES, ethnicity and gender (Devine and MacGillicuddy 2016; Lingard 2007). It should, however, be acknowledged that EER has

revealed that student learning outcomes are associated with various student factors, including background factors, such as SES, gender and ethnicity, as well as other factors that are not likely to change, such as personality traits. This implies that even if students are provided with the same learning opportunities, variation in student learning achievement gains can be detected and this variation can partly be explained by student background factors beyond their cognitive abilities and the effort that they put in to achieve these outcomes (Kyriakides and Luyten 2009; Lim 2013). Even if students are given the same opportunities within the schools, not all students will manage to progress at the same rate given the strong influence of home environment (Lareau 2016) and the intersection with patterns of structural inequality in the society at large (Sammons et al. 2018). Effective schools are therefore expected to provide further support for those disadvantaged groups of students (based on their background characteristics, such as SES, gender and ethnicity) in order to ensure that differences in learning outcomes are substantially reduced (Kelly and Downey 2010; Kyriakides and Creemers 2011; OECD 2012). This implies that positive discrimination of different groups is not only legitimated in order to obtain equity in education but is also seen as a characteristic of effective education. We also argue that effective school-based interventions are not only those that contribute to the promotion of learning outcomes for all (*quality*) but also those that reduce differences in student learning outcomes between groups of students with different background characteristics (*equity*). It should however be acknowledged that differences in learning outcomes between different groups of students cannot be completely eliminated since these gaps can be attributed to other hidden mechanisms in society over which schools have no control. Thus, the study reported here moves a step forward to investigate the impact of DASI on promoting quality and equity in socially disadvantaged schools. Given that early effectiveness studies were concerned with identifying ways to help schools in disadvantaged areas to achieve improved learning outcomes, it is important to find out whether DASI can help schools in disadvantaged areas in four different countries to become more effective by making use of a theoretical framework that has been empirically tested in these countries. In addition, the impact of DASI on promoting equity should be investigated.

5 Methods

5.1 Participants

At the beginning of school year 2015–2016, a sample of 72 primary schools in four European countries (i.e. Cyprus, England, Greece and Ireland) in socially disadvantaged areas was selected. In each country, participating schools were randomly allocated into two groups: the experimental and the control group. A pre-measure of achievement in mathematics of all grade 4, 5 and 6 students ($n = 5560$) in the participating schools, and of the functioning of school factors, was also conducted. Table 1 provides some descriptive data on the background factors of students participating in this study and their prior achievement by country. In each country, the t test did not reveal any statistically significant difference at 0.05 level between the two groups in relation to two of the background characteristics of their students (i.e. SES, age) and their prior achievement in mathematics. With regard to the other two student

background factors (i.e. gender and ethnicity), the χ^2 test did not also reveal any statistically significant difference at 0.05 level.

Both groups of schools were asked to develop improvement strategies and action plans to promote quality and equity. The next part of this section provides a brief description of the intervention that took place in these schools in order to promote quality and equity. At the end of the school year, student achievement in mathematics and the functioning of school factors were measured.

Table 1 Descriptive data about the background factors of the students in the experimental and the control group per country

Student background factors	Experimental		Control		*t* test		
	Mean	SD	Mean	SD	*t*	*df*	*p*
Cyprus							
Prior achievement	0.64	1.11	0.68	1.13	−0.76	1788	0.449
Post achievement	1.09	1.15	0.87	1.08	4.16	1788	0.001
SES	0.85	0.73	0.80	0.68	1.49	1788	0.135
Age in days	3805	393	3830	391	−1.58	1788	0.115
England							
Prior achievement	0.78	1.41	0.74	1.38	0.45	988	0.647
Post achievement	1.22	1.40	0.98	1.39	2.69	988	0.007
SES	0.61	0.56	0.64	0.40	−0.97	988	0.331
Age in days	3472	327	3496	342	−1.13	988	0.259
Greece							
Prior achievement	0.72	1.43	0.68	1.35	0.51	1284	0.597
Post achievement	0.98	1.25	0.81	1.16	2.52	1284	0.012
SES	0.65	0.57	0.70	0.58	−1.56	1284	0.123
Age in days	3730	329	3742	328	−0.65	1284	0.513
Ireland							
Prior achievement	0.90	1.32	0.85	1.34	0.72	1492	0.469
Post achievement	1.24	1.29	0.99	1.13	4.14	1492	0.001
SES	0.55	0.53	0.51	0.66	1.29	1492	0.194
Age in days	3926	335	3938	344	−0.69	1492	0.498
Overall							
Prior achievement	0.76	1.34	0.73	1.30	0.85	5558	0.40
Post achievement	1.13	1.30	0.91	1.25	6.42	5558	0.001
SES	0.68	0.65	0.67	0.66	0.57	5558	0.559
Age in days	3765	367	3776	363	−1.13	5558	0.262

Number of participants in the experimental and control groups per country:

Cyprus: Experimental ($n = 930$) / Control ($n = 860$)

England: Experimental ($n = 489$) / Control ($n = 501$)

Greece: Experimental ($n = 677$) / Control ($n = 609$)

Ireland: Experimental ($n = 803$) / Control ($n = 691$)

5.2 The treatment offered to the experimental group

At the beginning of the school year 2015–2016, a seminar for the head teachers of the schools participating in the experiment group ($n = 36$) was organised to reach a consensus on the general purpose of the intervention and to discuss about the main phases of the project, the role of the A&R Team and the role of the school stakeholders (see steps A and B of DASI). A handbook was also produced which presented the theoretical framework of the intervention and provided suggestions for schools on how to build school evaluation mechanisms that aim to improve educational practices at school and classroom level. The handbook also included the rationale of the project and clarified the role of the A&R Team. It was made clear that researchers should assist them in carefully setting up their own strategies and action plans for promoting both quality and equity. The handbook can be accessed from the web page of the project (www.ucy.ac.cy/promqe).

In the next step of the intervention, the A&R Team provided support to the schools to help them establish school evaluation mechanisms (see step C of DASI). In addition, the A&R Team analysed the data that emerged from the teacher questionnaire and supplied each school with feedback indicating its priorities for improvement. Subsequently, each country's A&R Team visited the schools of the experimental group and participated in staff meetings to present the results of the teacher questionnaire. In this way, school stakeholders had the chance to discuss the findings of school evaluation and decide whether their action plans would address one or a combination of priorities concerning the factors included in the dynamic model which were found to perform less well in their schools. It was strongly recommended that students and parents should be actively involved in selecting the school improvement area(s). For this reason, schools were encouraged to establish a committee with representatives of parents, students and teachers to discuss the results and gradually reach a consensus about the priorities of the school and how to address them. The final decision was announced to the whole school community and feedback was provided which helped schools to produce a clear definition of their improvement area.

At the next stage, school stakeholders (in collaboration with the A&R Team) developed their improvement strategies and action plans (step D of DASI). At this point, school stakeholders were reminded to make use of the suggestions and additional reading resources provided in the handbook in order to specify the activities involved in their improvement strategies and action plans. Examples of action plans designed by the schools can be found in the website of the project (www.ucy.ac.cy/promqe). Beyond designing action plans, school stakeholders were further asked to make decisions regarding the monitoring of the implementation of their strategies and action plans (see step E of DASI). In the next stage, the intervention was implemented. The implementation of DASI lasted for approximately 8 months and the A&R Team provided support to the school stakeholders by helping them overcome difficulties and problems that emerged during the implementation of their action plans. Specifically, the country teams visited the schools in the experimental group once every 6 weeks to provide feedback and support in the implementation and/or in modifying the action plans. A network within and across countries of schools addressing the same factors was also developed in order to share experiences during the implementation of their school improvement strategies. Moreover, the A&R Team helped school stakeholders to use the evaluation data in order to modify their strategies and

action plans, according to the circumstances and specific needs of different groups of the school population. The appropriate and timely modification of action plans was found to reduce the chance of a school finding out at a very at late stage that no progress had been made during the school year due to poor implementation of its action plans.

5.3 Support offered to the schools of the control group

In order to evaluate the impact of DASI, the A&R Team gave feedback to the control group of primary schools ($n = 36$) regarding the results that emerged from a pre-measure concerned with student achievement in mathematics and the functioning of effectiveness factors in their school (but without mentioning what their improvement priorities were) as revealed by teacher responses to the questionnaire measuring the school factors of the dynamic model. These schools were also offered support to develop their own strategies and actions to promote quality and equity, but without using the DASI. Support was, therefore, provided by the A&R Team and the same amount of time and effort was allocated to each treatment group. In addition, accounts of the effort that schools put into the project were collected, but we did not identify any statistically significant difference between the two groups. By following this approach, we were able to provide equal support for each group and to control for the Hawthorne effect in two ways: all groups put the same amount of effort into their specific intervention and schools in each group were not aware of the other interventions thus avoiding compensatory rivalry or resentful demoralisation on the part of either of the two groups (Shadish et al. 2002).

5.4 Measures

Student achievement in mathematics at the beginning and at the end of the intervention For each year group of students, criterion-reference tests in mathematics were constructed in order to measure their knowledge and skills in mathematics in relation to the objectives of the national curriculum in the four European countries. Specifically, a specification table covering the basic skills in mathematics expected to be taught to students in grades 3 to 6 by all countries was developed. The specification table was developed in collaboration with expert teachers and ministry officials from each country so as to represent adequately the mathematics skills and knowledge included in the mathematics curriculum of each country. Based on this table, a battery of written tests was developed. The written tests administered during the main study were subject to control for reliability and validity. The face and content validity of each test was evaluated by a group of expert teachers and teaching mathematics academics in each country. It is important to note that none of the respondents achieved a full score, and none scored zero. Moreover, less than 3% of the students achieved over 80% of the maximum score, and less than 9% of the students achieved over 70% of the maximum score. Based on the range of the results, the ceiling and floor effects in the attainment data were not observed.

Equating of tests The test administered to grade 6 students when they were at the end of the school year was purposefully more difficult than the one administered to grade 4 students when they were at the beginning of the school year so as to correspond to their

age skills, maturity stage and level of mathematics knowledge. In order to make the comparison of test scores meaningful, the scores had to be made comparable. Equating was done using Item Response Theory (IRT) modelling (Hambleton and Swaminathan 1985). The method of equating followed the same procedure as that used in PISA studies. However, in PISA, equating is horizontal (equating the different versions of tests), whereas in this study the equating was vertical. Specifically, the scores were transformed into the same scale on the basis of characteristics of IRT models with students' latent level of ability (y) and difficulty level of an item (b) being identical when certain preconditions were fulfilled (Bond and Fox 2001). The latent ability level for each student could be determined in every version as long as there were so-called anchoring items connecting the versions. For the purposes of this study, we used sufficient common items (i.e. approximately 15% of anchoring items across all tests) with representative content to be measured (Kolen and Brennan 1995). Estimation was made using the Extended Logistic Model of Rasch (Andrich 1988), which revealed that each scale had satisfactory psychometric properties. Specifically, for each scale, the indices of cases (i.e. students) and item separation were higher than 0.82, indicating that the separability of each scale was satisfactory (Wright 1985). Moreover, the infit mean squares and the outfit mean squares of each scale were near one and the values of the infit *t* scores and the outfit t scores were approximately zero. Furthermore, each analysis revealed that all items had item infit with the range 0.84 to 1.19. Therefore, for each assessment period, achievement in mathematics was estimated by calculating the Rasch person estimates.

Student background factors Information was collected on four student background factors: gender (0 = boys, 1 = girls), ethnicity, language spoken at home (0 = other language, 1 = language of instruction at school) and SES. Five SES variables were available: father's and mother's education level, the social status of father's job, the social status of mother's job and the main elements of the home learning environment. Specifically, it was possible to classify parents' occupation into three groups: occupations followed by the working class (63%), occupations followed by the middle class (28%) and occupations followed by the upper-middle class (9%). All five variables were considered in establishing an SES score since variation across the five variables per student can be observed. For example, it is possible that a father has a professional job whereas a mother does not. It may also be possible that a father has a university degree but currently has a blue-collar job. It was for this reason that data emerged from all five variables were taken into account in generating an SES score. We even took into account data emerged from another part of the student questionnaire which was concerned with the main elements of the home learning environment (i.e. learning materials available at home and learning opportunities offered at home). The Extended Logistic Model of Rasch (Andrich 1988) was used to analyse the ordinal data that emerged from the questionnaire. Thus, a scale which referred to the student SES score was created and analysed for reliability, fit with the model, meaning and validity. Analysis of the data revealed that the scale had relatively satisfactory psychometric properties. Specifically, the indices of cases (i.e. students) and item separation were higher than 0.87, indicating that the separability of the scale was satisfactory (Wright 1985). Furthermore, the infit mean squares and the outfit mean squares of each scale were near one and the values of the infit *t* scores and the outfit *t* scores were

approximately zero. The analysis revealed that there was a good fit with the model (Keeves and Alagumalai 1999). Thus, an overall score for the SES of each student was calculated using the relevant Rasch person estimate in the overall SES scale.

Using a teacher questionnaire to measure school factors The explanatory variables which refer to the school level factors of the dynamic model were measured by asking the teachers in the school sample ($n = 762$) to complete a questionnaire. This questionnaire can be accessed from the web page of the project (www.ucy.ac.cy/promqe). Since it is expected that teachers within a school are or should be aware of the policy of their school and the evaluation mechanisms of their school similarly, but differently from teachers in other schools, a generalisability study was initially conducted. For each participating country, it was found that in all the questionnaire items, the object of measurement was the school. Reliability was then computed for each of the dimensions of the school factors by calculating multilevel λ (Snijders and Bosker 1999) and the Cronbach alpha for data aggregated at the school level. The value of the Cronbach alpha represented consistency across items, whereas multilevel λ represented consistency across groups of teachers. For all factors, the reliability coefficients were high (around 0.80). It was also found that the percentages of variance at the school level were between 29 and 38%.

To test the validity of the questionnaire, separate Structural Equation Modelling (SEM) analyses were conducted for each of the three overarching factors: (1) *school policy on teaching*, (2) *policy on the school learning environment* and (3) *school evaluation*. The first overarching factor was *school policy on teaching* and consisted of factors measuring the quantity of teaching (e.g. Whole school records are kept concerning student absenteeism), provision of learning opportunities (e.g. Our school participates in programmes/projects (e.g. Erasmus, action research projects, collaboration with other schools, pilot initiatives) that focus on providing learning opportunities beyond those offered by the formal curriculum) and quality of teaching (e.g. Our school takes into consideration the professional experience, skills and aptitudes of each individual teacher in designing and implementing our school policy/policies for teaching). The second overarching factor was *policy on the school learning environment* and comprised factors measuring student behaviour outside the classroom (e.g. When supervising students on playground, teachers are encouraged to interact with children who may require support (e.g. children who are upset, isolated or display challenging behaviour)), teacher collaboration (e.g. The teachers in our school cooperate with each other by exchanging ideas and materials when teaching specific units or series of lessons), partnership policy (e.g. At staff meetings, we make decisions on how parents/guardians can be involved in learning activities) and provision of sufficient learning resources (e.g. There is material on noticeboards in the school relevant to the effective use of a range of educational resources for teaching). Finally, the third overarching factor was *policy on evaluation* and was composed of factors measuring evaluation of the school policy on teaching (e.g. Teachers' capacity to implement school policy on teaching (e.g. quantity of education, quality of education, provision of learning opportunities for students) is evaluated within the school) and evaluation of the learning environment (e.g. Our school regularly reviews and revises policies concerned with the broader learning environment of school). Additionally, for each of the three overarching factors, another model was tested in order to compare its fitting indices with the data from the three proposed theoretical models (i.e. model 1). In

these alternative models (model 2), all items that were used for the SEM analyses of each of the three overarching factors were considered to belong to a single factor. The fit indices of each model per school factor are shown in Table 2, where it can be seen that model 1 was found to have the best fitting and that in each case (i.e. overarching factor) the fit indices of this model were satisfactory.

For each overarching factor, Multiple Group Confirmatory Factor Analysis (MGCFA) was then conducted to test whether the teacher questionnaire elicit similar response patterns across the four countries. Specifically, measurement invariance can be examined on three sequential levels: configural, metric and scalar (Kline 2015). Configural invariance investigates the extent to which the pattern of fixed and free factor loadings among and between factors and items is the same and a value of RMSEA is expected to be smaller than 0.05 (Wu et al. 2007). For each overarching factor, configural invariance was supported since for each country the values of RMSEA were found to be smaller than 0.05 and the values of CFI were higher than 0.94. The second step of invariance involves the examination of metric invariance by comparing the baseline model (which allows the factor loadings to be freely estimated across multiple groups) and the invariance model (which expects the factor loadings to be equal across the four country groups). Differences between the two nested models can be examined with the χ^2 difference test (Muthén and Muthén 2012) and the ΔCFI (Cheung and Rensvold 2002). It was first of all found that for each overarching factor, the baseline model fit the data adequately (i.e. CFI bigger than 0.96 and RMSEA smaller than 0.06). Then, all the factor loadings across the four country groups were constrained to be equal, but for two overarching factors (i.e. policy for school learning environment and school evaluation), the data did not fit adequately to the relevant models (i.e. CFI smaller than 0.90 and RMSEA bigger than 0.15). Moreover, for each overarching factor, the corrected χ^2 difference test indicated that the factor loading invariant model was significantly worse than the baseline model. In addition, the ΔCFI was much bigger than 0.01, indicating that the metric invariance of the teacher questionnaire was not supported for any of the three scales measuring the overarching school factors.

The lack of metric and scalar invariance makes factor score comparisons invalid since differing response mechanisms seem to underlie country-group answers to the

Table 2 Fit indices of the models that emerged from the SEM analyses of the teacher questionnaire used to measure each overarching school factor

Models	χ^2	*df*	χ^2/df	*p* value	CFI	RMSEA	Range RMSEA
School policy on teaching							
Model 1	140	16	8.75	0.001	0.992	0.051	0.045–0.058
Model 2 (one-factor model)	493	20	24.7	0.001	0.941	0.093	0.085–0.099
Policy on the school learning environment							
Model 1	679	96	7.1	0.001	0.967	0.052	0.045–0.063
Model 2 (one factor model)	3888	135	28.8	0.001	0.738	0.099	0.096–0.107
School evaluation							
Model 1	544	57	9.54	0.001	0.969	0.056	0.048–0.060
Model 2 (one-factor model)	1545	65	23.8	0.001	0.895	0.093	0.089–0.096

items of each overarching factor (Brown et al. 2017). However, the purpose of this study was not to compare the overarching factor scores across the four countries but to measure the effect of DASI in each country. Even if measurement invariance was established, comparison across the four countries would have not been conducted especially since the sample in each country was not nationally representative. Nevertheless, we decided to conduct four separate within-country analyses to measure the effect of DASI on student achievement gains in each country. Since configural invariance was achieved, it was also decided to generate factor scores by taking into account teacher responses to the equivalent questionnaire items by considering the SEM model emerged in each country. At the next step, for each school, separate analysis of the teacher responses to the questionnaire items was conducted, and those factors which had the lowest mean rank values were identified, indicating each school priorities for improvement. The results were reported to each school and stakeholders in the experimental group were encouraged to develop their strategies and action plans in order to improve the functioning of those factors for which lower mean rank values were estimated. A similar approach was used in analysing teachers' responses to the questionnaire at the end of the intervention. The reports sent to the schools at the end of the intervention made suggestions regarding the improvement areas that each school could consider in developing its own strategies and action plans during the next school year (i.e. 2016–2017).

6 Findings

The results section is split into three parts. The first part refers to the impact of the intervention upon the functioning of school factors since these factors were directly addressed through the action plans developed by the experimental group schools. The second part investigates the extent to which the intervention had an impact on student achievement gains (*quality dimension*) in each country, and the results of the four within-country multilevel regression analyses are presented. The impact of this intervention on promoting *equity* in each country is examined in the last part of this section.

6.1 The impact of the intervention on improving school factors

Table 3 presents the means and standard deviations of the three school-level overarching factor scores before the implementation of the intervention and at the end of the intervention in the experimental and control schools in each of the four countries participating in this study. Although a repeated measures MANOVA of treatment (following DASI/not following the proposed approach) according to time (before (i.e. pre)/end (i.e. post)) could have been carried out with the three factor scores (i.e. policy on teaching, school learning environment and school evaluation) as dependent variables, we decided to compare the school factor scores of these two groups by using non-parametric statistical tests due to our small sample size at the school level (i.e. less than 15 schools in each group). Siegel and Castellan (1988) argue that when the sample size is small, non-parametric tests are preferable to parametric tests, even when interval data have been collected. Thus, the Kolmogorov–Smirnov two-sample test was initially employed to identify any statistically

Table 3 Means and standard deviations of the functioning of each overarching school factor in the experimental and control schools and values of the Kolmogorov–Smirnov two-sample test

Overarching school factor	Experimental school		Control school		K-S Z	p value
	Mean	SD	Mean	SD		
Cyprus						
(A) Before the intervention						
School policy on teaching	2.96	0.87	2.98	0.69	−0.766	0.601
School learning environment	3.04	0.77	3.03	0.55	0.656	0.782
School evaluation	2.77	0.73	2.79	0.65	−0.774	0.587
(B) At the end of intervention						
School policy on teaching	3.48	0.63	3.04	0.68	1.474	0.028
School learning environment	3.80	0.83	3.09	0.76	1.992	0.009
School evaluation	3.15	0.78	2.86	0.73	1.413	0.029
England						
(A) Before the intervention						
School policy on teaching	3.11	0.86	3.15	0.82	−0.832	0.493
School learning environment	3.05	0.80	3.07	0.90	−0.799	0.547
School evaluation	2.95	0.96	2.93	0.89	0.661	0.765
(B) At the end of intervention						
School policy on teaching	3.46	0.90	3.17	0.84	1.389	0.038
School learning environment	3.42	0.86	3.08	0.88	1.989	0.007
School evaluation	3.29	0.91	2.91	0.89	1.467	0.031
Greece						
(A) Before the intervention						
School policy on teaching	3.05	0.83	2.98	0.89	0.914	0.874
School learning environment	3.10	0.72	3.13	0.73	−0.616	0.799
School evaluation	2.71	0.83	2.74	0.28	−0.963	0.312
(B) At the end of intervention						
School policy on teaching	3.39	0.80	3.00	0.81	1.713	0.019
School learning environment	3.46	0.74	3.11	0.73	1.450	0.038
School evaluation	3.11	0.71	2.75	0.83	1.389	0.041
Ireland						
(A) Before the intervention						
School policy on teaching	2.98	0.76	3.05	0.69	−0.821	0.502
School learning environment	2.89	0.86	2.84	0.79	0.799	0.547
School evaluation	2.81	0.82	2.87	0.83	−0.963	0.312
(B) At the end of intervention						
School policy on teaching	3.29	0.74	2.97	0.72	1.934	0.015
School learning environment	3.18	0.73	2.91	0.73	1.656	0.024
School evaluation	3.12	0.24	2.91	0.25	1.611	0.035

significant difference between the two groups in terms of the functioning of the three overarching school factors before the intervention. No statistically significant difference was

identified at the 0.05 level. This implies that the two groups were performing equally well in relation to the functioning of the three overarching school factors. At the end of the intervention, the Kolmogorov–Smirnov two-sample test revealed statistically significant differences at the 0.05 level between the two groups of schools in each country in relation to each school factor (see Table 3). Moreover, the Wilcoxon test was used to identify whether there was any statistically significant progress in the performance of each group of schools in relation to the three overarching school factors. In each country, it was found out that only the schools in the experimental group managed to improve the functioning of their school factors at a statistically significant level.

6.2 The impact of the intervention on improving student achievement gains in mathematics

In each country, we conducted multilevel regression analysis to identify the impact of DASI on student achievement at the end of the intervention. Specifically, an empty model consisting of student, class and school levels was initially used. In subsequent steps, explanatory variables at different levels were added, starting at the student level. Explanatory variables, apart from grouping variables, were centred as Z-scores with a mean of 0 and a standard deviation of 1. Grouping variables were entered as dummies with one of the groups as the baseline (e.g. boys = 0). The models presented in Tables 4, 5, 6 and 7 were estimated without the variables that had no statistically significant effect at level 0.05.

In model 1, the context variables at each level (i.e. prior achievement, gender, SES, age and ethnicity) were added to the empty model. Tables 4, 5, 6 and 7 reveal that in each country, model 1 explained at least 35% of the total variance, and most of the explained variance was at the student level. Moreover, the likelihood statistic revealed a statistically significant change between the empty model and model 1 ($p < 0.001$). Each analysis also revealed that the effects of all contextual factors, other than ethnicity and age, were significant at 0.05 level. Moreover, prior knowledge was the only contextual variable which had a significant effect on student achievement when aggregated at the school level. In model 2, the impact of DASI was tested by adding to model 1 a dummy variable. By considering the control group as a reference group, it was found that the group which made use of DASI managed to achieve better results than the control group in each participating country (see Tables 4, 5, 6 and 7).

The findings of the four multilevel regression analyses seem to provide support for the use of DASI in socially disadvantaged schools to promote quality in education in each participating country. Table 8 summarises the main results concerning the impact of DASI that emerged from each within-country analysis. The fixed effects obtained from multilevel analysis can readily be converted into standardised effects or 'Cohen's *d*' by dividing them by the standard deviations in the 'treatment group' which made use of the DASI to promote student learning outcomes in mathematics. Thus, Table 8 presents the effect sizes of using DASI according to country, and we can observe some differences in the reported effect sizes. The impact of DASI in some countries (i.e. Greece and Ireland) was found to be bigger, whereas in England its impact was smaller. These differences could be attributed to differences between countries in terms of the support that schools in socially disadvantaged areas received from the system level to promote quality. Differences in the extent to which the school factors of the dynamic model are addressed by national policies and the relevant school evaluation

Table 4 Parameter estimates and standard errors for the analysis of mathematics achievement (students within classes, within schools) of students in Cyprus

Factors	Model 0	Model 1	Model 2
Fixed part			
Intercept	0.85 (0.05)*	0.81 (0.05)*	0.56 (0.05)*
Student level			
Prior achievement		0.68 (0.02)*	0.67 (0.02)*
Gender (0 = boy, 1 = girl)		− 0.07 (0.02)*	− 0.07 (0.02)*
SES		0.18 (0.04)*	0.17 (0.04)*
Age		0.06 (0.04)	
Ethnicity (0 = other, 1 = immigrant background)		− 0.05 (0.04)	
Class level			
Average prior achievement		0.11 (0.04)*	0.10 (0.04)*
Percentage of girls		− 0.03 (0.04)	
Average SES		0.07 (0.04)	
Average age		0.04 (0.04)	
Percentage of students with immigrant background		− 0.05 (0.03)	
School level			
Context			
Average prior achievement		0.14 (0.06)*	0.13 (0.06)*
Percentage of girls		− 0.02 (0.04)	
Average SES		0.06 (0.04)	
Average age		0.04 (0.04)	
Percentage of students with immigrant background		− 0.03 (0.03)	
DASI (0 = control, 1 = experimental)			0.24 (0.02)*
Variance components			
School	11.2%	9.8%	4.1%
Class	17.1%	14.2%	12.1%
Student	71.7%	36.3%	35.1%
Explained		39.7%	48.7%
Significant test			
χ^2	6604.4	4862.3	4341.1
Reduction		1742.1	521.2
Degrees of freedom		5**	1
p value		0.000	0.000

*Statistically significant effect at 0.05 level

**The models presented in this table were estimated without the variables that did not have a statistically significant effect at 0.05 level

mechanisms should also be considered. For example, in Cyprus, the Ministry of Education refers explicitly to the school factors of the dynamic model and expects head teachers to develop action plans in order to improve the functioning of school factors, and this might have had an effect on schools in the control group.

Table 5 Parameter estimates and standard errors for the analysis of mathematics achievement (students within classes, within schools) of students in England

Factors	Model 0	Model 1	Model 2
Fixed part			
Intercept	0.99 (0.05)*	0.81 (0.05)*	0.66 (0.05)*
Student level			
Prior achievement		0.62 (0.03)*	0.62 (0.02)*
Gender (0 = boy, 1 = girl)		−0.08 (0.02)*	−0.07 (0.02)*
SES		0.45 (0.04)*	0.44 (0.04)*
Age		0.05 (0.04)	
Ethnicity (0 = other, 1 = immigrant background)		−0.06 (0.04)	
Class level			
Average prior achievement		0.13 (0.04)*	0.12 (0.04)*
Percentage of girls		−0.04 (0.04)	
Average SES		0.08 (0.04)*	0.08 (0.04)*
Average age		0.06 (0.04)	
Percentage of students with immigrant background		−0.06 (0.04)	
School level			
Context			
Average prior achievement		0.17 (0.06)*	0.16 (0.06)*
Percentage of girls		−0.02 (0.04)	
Average SES		0.13 (0.04)*	0.13 (0.04)*
Average age		0.04 (0.04)	
Percentage of students with immigrant background		−0.03 (0.03)	
DASI (0 = control, 1 = experimental)			0.16 (0.03)*
Variance components			
School	14.2%	12.8%	9.5%
Class	19.1%	15.2%	12.1%
Student	66.7%	33.3%	32.1%
Explained		38.7%	46.3%
Significant test			
χ^2	3051.7	2021.3	1841.1
Reduction		1030.4	180.2
Degrees of freedom		7**	1
p value		0.000	0.000

*Statistically significant effect at 0.05 level

**The models presented in this table were estimated without the variables that did not have a statistically significant effect at 0.05 level

6.3 The impact of the intervention on equity

To investigate the impact of DASI on promoting equity, two separate within-country multilevel analyses were conducted for each group of schools to identify the direct impact of SES on prior and final student achievement in mathematics. In this way, it

Table 6 Parameter estimates and standard errors for the analysis of mathematics achievement (students within classes, within schools) of students in Greece

Factors	Model 0	Model 1	Model 2
Fixed part			
Intercept	0.81 (0.07)*	0.67 (0.06)*	0.46 (0.06)*
Student level			
Prior achievement		0.52 (0.03)*	0.52 (0.03)*
Gender (0 = boy, 1 = girl)		− 0.05 (0.02)*	− 0.05 (0.02)*
SES		0.35 (0.06)*	0.34 (0.06)*
Age		0.07 (0.05)	
Ethnicity (0 = other, 1 = immigrant background)		− 0.04 (0.04)	
Class level			
Average prior achievement		0.21 (0.04)*	0.19 (0.04)*
Percentage of girls		− 0.03 (0.04)	
Average SES		0.05 (0.04)	
Age		0.04 (0.04)	
Percentage of students with immigrant background		− 0.04 (0.04)	
School level			
Context			
Average prior achievement		0.17 (0.06)*	0.16 (0.06)*
Percentage of girls		− 0.01 (0.04)	
Average SES		0.05 (0.04)	
Average age		0.02 (0.04)	
Percentage of students with immigrant background		− 0.01 (0.03)	
DASI (0 = control, 1 = experimental)			0.28 (0.02)*
Variance components			
School	13.6%	11.8%	7.1%
Class	16.1%	12.2%	8.1%
Student	70.3%	36.0%	34.1%
Explained		40.0%	50.7%
Significant test			
χ^2	2790.4	2100.3	1850.2
Reduction		690.1	250.1
Degrees of freedom		5**	1
p value		0.000	0.000

*Statistically significant effect at 0.05 level

**The models presented in this table were estimated without the variables that did not have a statistically significant effect at 0.05 level

was possible to discover whether the direct effect of SES on achievement became smaller in the experimental and/or in the control group. The results of the two separate multilevel analyses of student achievement at the beginning of the intervention revealed that two background factors (i.e. SES and gender) were associated with the achievement of each group of students at the beginning of the intervention in each country. In

Table 7 Parameter estimates and standard errors for the analysis of mathematics achievement (students within classes, within schools) of students in Ireland

Factors	Model 0	Model 1	Model 2
Fixed part			
Intercept	0.82 (0.06)*	0.69 (0.05)*	0.45 (0.05)*
Student level			
Prior achievement		0.51 (0.03)*	0.51 (0.03)*
Gender (0 = boy, 1 = girl)		− 0.05 (0.02)*	− 0.05 (0.02)*
SES		0.31 (0.04)*	0.31 (0.04)*
Age		0.03 (0.05)	
Ethnicity (0 = other, 1 = immigrant background)		− 0.03 (0.04)	
Class level			
Average prior achievement		0.06 (0.04)	
Percentage of girls		− 0.02 (0.04)	
Average SES		0.03 (0.04)	
Average age		0.02 (0.04)	
Percentage of students with immigrant background		− 0.02 (0.03)	
School level			
Context			
Average prior achievement		0.13 (0.06)*	0.12 (0.06)*
Percentage of girls		− 0.01 (0.02)	
Average SES		0.03 (0.04)	
Average age		0.02 (0.03)	
Percentage of students with immigrant background		− 0.01 (0.04)	
DASI (0 = control, 1 = experimental)			0.32 (0.02)*
Variance components			
School	12.1%	10.8%	7.1%
Class	16.1%	13.2%	9.3%
Student	71.8%	37.1%	34.5%
Explained		38.9%	49.1%
Significant test			
χ^2	5458.2	4157.1	3937.0
Reduction		1301.1	220.1
Degrees of freedom		4**	1
p value		0.000	0.000

*Statistically significant effect at 0.05 level

**The models presented in this table were estimated without the variables that did not have a statistically significant effect at 0.05 level

order to estimate the relative importance of the SES on student achievement at the beginning of the intervention for each group of students, the fixed effect obtained from each multilevel analysis was converted to standardised effects or 'Cohen's *d*' by following the approach proposed by Elliot and Sammons (2004). By using this approach, it was found that the effect size was equally high for each group of students

Table 8 Effect of using the DASI approach on student achievement gains in mathematics

Country	Effect	Pooled SD	Cohen's *d*
Cyprus	0.24	0.73	0.33
England	0.16	0.71	0.23
Greece	0.28	0.67	0.42
Ireland	0.32	0.84	0.38

in each participating country (see Table 9). For each group, separate within-country analyses investigating the impact of student background factors on achievement at the end of the intervention were conducted. These analyses revealed that achievement at the end of the intervention was associated with all student background factors apart from age and ethnicity (i.e. prior achievement, SES and gender). With regard to the direct effect of SES on student achievement at the end of the intervention, it was found that, for students in the control group, the effect of SES was bigger than for those in the experimental group (see Table 9). By comparing the effect size of SES at the beginning and at the end of the intervention, one can see that the direct effect of SES was reduced in the schools of the experimental group in each participating country (see Table 9). Further analysis was also conducted to measure the total effect of SES in each group of schools by taking into account the fact that SES has an indirect effect on final achievement through its impact on prior achievement. Table 9 revealed differences in the total effect of SES on achievement between the experimental and the control group. Specifically, in each country, schools which made use of DASI not only managed to reduce the direct effect of SES on final achievement in mathematics but also had smaller total effects of SES on achievement by the end of the intervention. However, by comparing the effect of SES on student achievement at the beginning of the intervention with the total effect of SES on achievement at the end of the intervention, one can see that neither a reduction nor an increase was observed in the schools of the experimental group.

7 Discussion

This European study highlighted the importance of utilising DASI to improve not only the quality but also equity in primary schools. Implications of findings for research, policy and practice are drawn. First, previous studies revealed that DASI had an effect on promoting student learning outcomes. It was also found that interventions based on DASI had both a direct and indirect effects (through improving school factors) on student learning outcomes. The issue of equity was however not addressed. This experimental study seems to reveal that by using DASI not only the quality dimension of school effectiveness but also equity can be promoted. Specifically, at the beginning of the intervention, the effect of SES on student achievement was equally strong in the experimental and control groups (with an effect size comparable to what has been reported in relevant meta-analyses, e.g., Sirin 2005; White 1982). At the end of the intervention, the direct effect of SES on student achievement became smaller only in the schools of the experimental group. In regard to the total effect of SES on student achievement in mathematics, a difference in the two groups was also identified. It is, however, important to note that by comparing the effect of SES on initial achievement of

Table 9 Effects (in Cohen's *d* values) of SES on achievement at the beginning and at the end of the intervention of students at the schools of the control and experimental group per county

	Experimental group	Control group
Cyprus		
Effect of SES on initial achievement	0.31	0.29
Direct effect of SES on final achievement	0.24	0.30
Total effect of SES on final achievement	0.32	0.39
England		
Effect of SES on initial achievement	0.28	0.27
Direct effect of SES on final achievement	0.22	0.28
Total effect of SES on final achievement	0.29	0.39
Greece		
Effect of SES on initial achievement	0.35	0.36
Direct effect of SES on final achievement	0.29	0.35
Total effect of SES on final achievement	0.34	0.44
Ireland		
Effect of SES on initial achievement	0.27	0.28
Direct effect of SES on final achievement	0.20	0.26
Total effect of SES on final achievement	0.28	0.34

students in the experimental with the total impact of SES on their final achievement, one can see that it remained equally high. Nevertheless, one should not underestimate the impact of DASI on promoting equity since the total effect of SES in the control schools increased substantially in each country, and this finding is in line with the results of longitudinal studies conducted in different countries, which reveal that the total effect of SES gradually increases over time (Hansen et al. 2011; Sammons 2008). One might then claim that the intervention had a positive effect on equity since, by comparing the impact of SES on achievement at two points in time, one can see that there was no difference before the intervention, whereas at the end of the intervention SES was found to matter less in the experimental schools than in the schools in the control group.

The results of longitudinal studies as the ones mentioned above and also of the present study that took place in schools for only one school year, demonstrating an increase of the impact of SES over time, reveal that policy-makers and school stakeholders should consider as a priority the promotion of equity at school level. At the same time, these results reveal the complexity of evaluating interventions aiming to promote equity. One could claim that DASI was not in a position to promote equity since the impact of SES on student achievement was not reduced during the intervention period. However, the fact that the impact of SES was substantially increased in the control schools reveals that the intervention can be considered effective since the impact of SES on achievement at least remained the same in the schools of the experimental group. At the same time, implications for measuring the effectiveness status of schools in terms of the equity dimension can be drawn. We argue for collecting longitudinal data in order to establish formative school evaluation

mechanisms and measure changes in the impact of background factors on student achievement over time. This suggestion seems to be in line with the claim that value-added approaches are needed in measuring the quality dimension and comparing the contribution that each teacher/school/system makes to student achievement gains rather than comparing schools on the basis of final student learning outcomes (Creemers et al. 2010). In the case of equity, we should also seek changes over time of the impact that SES (or other background factors such as gender and ethnicity) have on achievement and compare the changes in individual schools with those in the whole population. In case an increase in the impact of SES is observed at country level, any school with a much smaller increase (or no increase) could still be considered effective in terms of equity. Thus, we argue for the need to establish continuous formative evaluation mechanisms about student achievement gains and the impact of SES on achievement over time in order to help teachers, schools and educational systems identify changes in their effectiveness status in terms of both dimensions of educational effectiveness (quality and equity).

Third, an essential difference of the study reported here from all studies investigating the impact of DASI has to do with the fact that DASI was offered in schools in socially disadvantaged areas. In general school improvement interventions in this type of schools are less likely to promote student learning outcomes (Sammons 2010; Townsend 2007). Thus, this study seems to reveal the potential of DASI to promote quality and equity especially since by using DASI for one school year statistically significant effects on quality and equity in mathematics were identified in socially disadvantaged schools in each participating country. Further research is, however, needed to investigate the impact of using DASI for a longer period in different educational settings. A study looking at the impact of a 3-year intervention based on DASI on quality of teaching (Kyriakides et al. 2017) reported a small effect of DASI (i.e. $d = 0.17$) during the first year of the intervention, but its effect was increased when the intervention was offered for 3 years ($d = 0.39$). This implies that DASI interventions may not reach an optimal point where it can have no further effect when offered for a long period. On the contrary, by offering the intervention for a period of 3 years, a medium effect of the intervention can be identified. Further research is, therefore, needed to search for the added value of offering DASI for a longer period in terms of promoting both quality and equity. Studies investigating the sustainability of DASI are also needed especially since one of the most important parts of an intervention programme is not only the investigation of its immediate impact on the functioning of school factors, on promoting student learning outcomes and on reducing the impact of SES on achievement, but also an exploration of the sustainability of its effects. Data emerged from the study reported in this paper were collected over the course of only one school year and the intervention lasted for approximately 8 months. Therefore, changes in school policy and/or the impact of these changes on the final student outcomes were only identified with respect to this period. This study reveals the potential benefit of investigating the impact of using DASI for a longer period to promote quality and equity, and, consequently, there is a need to conduct longitudinal studies to identify changes in the effectiveness status of schools in terms of both quality and equity, even after the end of the intervention.

It should finally be stressed that DASI supports the view that authentic changes designed to improve equity may come from interventions taking place at the school

rather than at the system level. One could therefore argue that the use of DASI stimulates a special approach to improvement whereby each party has a specific role in, and expertise that they contribute to, the intervention and thus ownership is accomplished. The relationship established between the school and the A&R Team reveals the main difference between DASI and other school improvement approaches that follow a top-down approach giving emphasis only to available knowledge that has emerged from educational effectiveness studies and not to the existing problems, situations, professional needs and abilities of the schools' stakeholders (teachers, students, parents). Thus, this dynamic approach can be used by stakeholders, especially when it is necessary to deal with improving the effectiveness status of schools situated in disadvantaged areas, since these schools have to face problems that require special attention and handling according to their context. Policy-makers should be able to support schools in implementing such an approach by providing them with all the necessary learning resources and especially an A&R Team that can help them identify improvement priorities, and then design, implement and evaluate school improvement strategies and action plans that take into account the knowledge base of school effectiveness research.

Acknowledgements The study presented in this paper is part of a three-year project (2014-2017) entitled 'Promoting Quality and Equity: A Dynamic Approach to School Improvement (PROMQE) (Agreement Number: 2014-1-CY01-KA200-000281)', under the Erasmus+ Key Action 2, Cooperation for innovation and the exchange of good practices (Field: School Education). This project has been funded with support from the European Commission. This publication 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.

References

Andrich, D. (1988). A general form of Rasch's extended logistic model for partial credit scoring. *Applied Measurement in Education, 1*(4), 363–378.

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.

Antoniou, P., & Kyriakides, L. (2013). A dynamic integrated approach to teacher professional development: Impact and sustainability of the effects on improving teacher behavior and student outcomes. *Teaching and Teacher Education, 29*(1), 1–12.

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.

Azkiyah, S. N., Doolaard, S., Creemers, B. P. M., & Van Der Werf, M. P. C. (2014). The effects of two intervention programs on teaching quality and student achievement. *Journal of Classroom Interaction, 49*(1), 4–11.

Bierman, K., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in head start REDI program. *Development and Psychopathology, 20*(3), 821–843.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah: Lawrence Erlbaum Associates.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and student achievement: A meta-analysis. *Review of Educational Research, 73*(2), 125–230.

Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: Identifying and understanding non-invariance. *International Journal of Research & Method in Education, 40*(1), 66–90.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago: University of Chicago Press.

Buczynski, S., & Hansen, C. B. (2010). Impact of professional development on teacher practice: Uncovering connections. *Teaching and Teacher Education, 26*(3), 599–607.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255.

Christoforidou, M., & Xirafidou, E. (2014). Using the dynamic model to identify stages of teacher skills in assessment. *Journal of Classroom Interaction, 49*(1), 12–25.

Christoforidou, M., Kyriakides, L., Antoniou, P., & Creemers, B. P. M. (2014). Searching for stages of teacher skills in assessment. *Studies in Educational Evaluation, 40*, 1–11.

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

Creemers, B. P. M., & Kyriakides, L. (2010a). Using the dynamic model to develop an evidence-based and theory-driven approach to school improvement. *Irish Educational Studies, 29*, 5–23.

Creemers, B. P. M., & Kyriakides, L. (2010b). 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.

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

Creemers, B. P. M., Kyriakides, L., & Sammons, P. (2010). *Methodological advances in educational effectiveness research*. London: Routledge.

Devine, D., & MacGillicuddy, D. (2016). Positioning pedagogy—A matter of children's rights? *Oxford Review of Education, 42*(4), 424–443.

Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J. A., & Jones, D. J. (2009). Fostering high quality teaching with an enriched curriculum and professional development: Head start REDI. *American Educational Research Journal, 46*(2), 567–597.

Donaldson, L. (2001). *The contingency theory of organizations: Foundations for Organisational science*. Thousand Oaks: Sage.

Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership, 37*(10), 15–24.

Elliot, K., & Sammons, P. (2004). Exploring the use of effect sizes to evaluate the impact of different influences on child outcomes: Possibilities and limitations. In I. Schagen & K. Elliott (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 6–24). National Foundation for Educational Research: Slough.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.

Hansen, K. Y., Rosén, M., & Gustafsson, J.-E. (2011). Changes in the multi-level effects of socio-economic status on reading achievement in Sweden in 1991 and 2001. *Scandinavian Journal of Educational Research, 55*(2), 197–211.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.

Heck, R. H., & Moriyama, K. (2010). Examining relationships among elementary schools' contexts, leadership, instructional practices, and added-year outcomes: A regression discontinuity approach. *School Effectiveness and School Improvement, 21*(4), 377–408.

Hofman, R. H., Hofman, W. H., & Gray, J. M. (2010). Institutional contexts and international performances in schooling: Comparing patterns and trends over time in international surveys. *European Journal of Education, 45*(1), 153–173.

Keeves, J. P., & Alagumalai, S. (1999). New approaches to measurement. In G. N. Masters & J. P. Keeves (Eds.), *Advances in measurement in educational research and assessment* (pp. 23–42). Oxford: Pergamon.

Kelly, A. (2012). Measuring 'equity' and 'equitability' in school effectiveness research. *British Educational Research Journal, 38*(6), 977–1002.

Kelly, A., & Downey, C. (2010). Value-added measures for schools in England: Looking inside the 'black box' of complex metrics. *Educational Assessment, Evaluation and Accountability, 22*(3), 181–198.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York: Guilford.

Kolen, M. J., & Brennan, R. L. (1995). *Test equating: Methods and practices*. New York: Springer.

Kyriakides, L. (2017). The dynamic approach to school improvement: Main features and impact on promoting quality and equity in education. *Keynote presentation at the ECER 2017 Conference, Network: 09. Assessment, Evaluation, Testing and Measurement*. Copenhagen, Denmark. 22–25 August.

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

Kyriakides, L., & 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), 61–86.

Kyriakides, L., & Creemers, B. P. M. (2011). Can schools achieve both quality and equity? Investigating the two dimensions of educational effectiveness. *Journal of Education for Students Placed at Risk, 16*(4), 237–254.

Kyriakides, L., & Luyten, H. (2009). The contribution of schooling to the cognitive development of secondary education students in Cyprus: An application of regression-discontinuity with multiple cut-off points. *School Effectiveness and School Improvement, 20*(2), 167–186.

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

Kyriakides, L., Archambault, I., & Janosz, M. (2013a). Searching for stages of effective teaching: A study testing the validity of the dynamic model in Canada. *Journal of Classroom Interaction, 48*(2), 11–24.

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

Kyriakides, L., Creemers, B. P. M., Muijs, D., Rekers-Mombarg, L., Papastylianou, D., Van Petegem, P., & Pearson, D. (2014). Using the dynamic model of educational effectiveness to design strategies and actions to face bullying. *School Effectiveness and School Improvement, 25*(1), 83–104.

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.

Kyriakides, L., Anthimou, M., & Charalambous, C.Y. (2016). Searching for the Impact of Teacher Behavior on Promoting Students' Cognitive and Metacognitive Skills. *Paper presented at the American Educational Research Association (AERA) 2016 Conference "Public Scholarship to Educate Diverse Democracies"*. Washington, DC, April 8–April 12.

Kyriakides, L., Christoforidou, M., Panayiotou, A., & Creemers, B. P. M. (2017). The impact of a three-year teacher professional development course on quality of teaching: Strengths and limitations of the dynamic approach. *European Journal of Teacher Education, 40*(4), 465–486.

Kyriakides, L., Creemers, B. P. M., & Charalambous, E. (2018). *Equity and quality dimensions in educational effectiveness*. Dordrecht: Springer.

Lareau, A. (2016). Invisible inequality—Social class and child rearing in black families and white families. *American Sociological Review, 67*(5), 747–776.

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.

Lingard, B. (2007). Pedagogies of indifference. *International Journal of Inclusive Education, 11*(3), 245–266.

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: Oxford Univ. Press.

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

Muthén, L.K., & Muthén, B.O. (2012). *Chi-square difference testing using the Satorra-Bentler Scaled Chi-Square*. Retrieved from http://www.statmodel.com/chidiff.shtml.

OECD. (2010). *PISA 2009 results: Overcoming social background: Equity in learning opportunities and outcomes (volume II)*. Paris: PISA, OECD Publishing Available in https://www.oecd.org/pisa/pisaproducts/48852584.pdf.

OECD. (2012). *Equity and quality in education: Supporting disadvantaged students and schools*. Paris: OECD Publishing. https://doi.org/10.1787/9789264130852-en.

OECD (2013). *PISA in focus*. Paris: OECD Publishing. Available in https://www.oecd.org/pisa/pisaproducts/pisainfocus/pisa%20in%20focus%20n27%20%28eng%29%2D%2DFINAL_version2.pdf

Panayiotou, A., Kyriakides, L., Creemers, B. P. M., McMahon, L., Vanlaar, G., Pfeifer, M., Rekalidou, G., & Bren, M. (2014). Teacher behavior and student outcomes: Results of a European study. *Educational Assessment, Evaluation and Accountability, 26*, 73–93.

Panayiotou, A., Kyriakides, L., & Creemers, B. P. M. (2016). Testing the validity of the dynamic model at school level: A European study. *School leadership and management, 36*(1), 1–20.

Rowan, B., Correnti, R., Miller, R. J., & Camburn, E. M. (2009). School improvement by design: Lessons from a study of comprehensive school reform programs. In G. Sykes & B. Schneider (Eds.), *Handbook on educational policy research* (pp. 637–651). London: Routledge.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge: Harvard University Press.

Sammons, P. (2008). Zero tolerance of failure and new labour approaches to school improvement in England. *Oxford Review of Education, 34*(6), 651–664.

Sammons, P. (2009). The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools. *School Effectiveness and School Improvement, 20*(1), 123–129.

Sammons, P. (2010). Equity and educational effectiveness. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 5, pp. 51–57). Oxford: 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.

Scheerens, J. (2013). The use of theory in school effectiveness research revisited. *School Effectiveness and School Improvement, 24*(1), 1–38.

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

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

Siegel, S., & Castellan, N. J. (1988). *Non-parametric statistics: For the behavioural sciences* (2nd ed.). New York: McGraw-Hill.

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

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

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

White, K. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3), 461–481.

Woessmann, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *Journal of Economic Perspectives, 30*(3), 3–32.

Wright, B. D. (1985). Additivity in psychological measurement. In E. E. Roskam (Ed.), *Measurement and personality assessment* (pp. 101–112). Amsterdam: Elsevier Science Publishers BV.

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation, 12*(3), 1–26.

Yoon, K. S., Garet, M., Birman, B., & Jacobson, R. (2007). *Examining the effects of mathematics and science professional development on teachers' instructional practice: Using professional development activity log*. Washington, DC: Council of Chief State School Officers.

Affiliations

Leonidas Kyriakides[1] · Evi Charalambous[1] · Bert P. M. Creemers[2] · Panayiotis Antoniou[3] · Dympna Devine[4] · Dona Papastylianou[5] · Declan Fahie[4]

Evi Charalambous
charalambous.e.evi@ucy.ac.cy

Bert P. M. Creemers
hubertus.creemers@gmail.com; b.p.m.creemers@rug.nl

Panayiotis Antoniou
pa241@cam.ac.uk

Dympna Devine
dympna.devine@ucd.ie

Dona Papastylianou
adpapast@ppp.uoa.gr; papastd@gmail.com

Declan Fahie
declan.fahie@ucd.ie

[1] Department of Education, University of Cyprus, P. O. Box 20537, 1678 Nicosia, Cyprus

[2] Faculty of Behavioural and Social Sciences, Department of Pedagogy and Educational Science, University of Groningen, Grote Rozenstraat 38, 9712 TJ Groningen, The Netherlands

[3] Faculty of Education, University of Cambridge, 184 Hills Road, Cambridge CB2 8PQ, England

[4] School of Education, University College Dublin, Roebuck Castle, Belfield, Dublin 4, Ireland

[5] Department of Philosophy, Pedagogy and Psychology, Section of Psychology, National and Kapodistrian University of Athens, H 505, 5th Floor, University Campus—Ilissia, 157 84 Athens, Greece