Routledge
Taylor & Francis Group

Journal of Education for Students Placed at Risk (JESPAR)

ISSN: 1082-4669 (Print) 1532-7671 (Online) Journal homepage: http://www.tandfonline.com/loi/hjsp20

Can Schools Achieve Both Quality and Equity? Investigating the Two Dimensions of Educational Effectiveness

L. Kyriakides & B. P. M. Creemers

To cite this article: L. Kyriakides & B. P. M. Creemers (2011) Can Schools Achieve Both Quality and Equity? Investigating the Two Dimensions of Educational Effectiveness, Journal of Education for Students Placed at Risk (JESPAR), 16:4, 237-254, DOI: 10.1080/10824669.2011.610269

To link to this article: http://dx.doi.org/10.1080/10824669.2011.610269

Journal of Education for Students Placed at Risk, 16: 237–254, 2011

ISSN: 1082-4669 print/1532-7671 online
DOI: 10.1080/10824669.2011.610269

Routledge
Taylor & Francis Group

Can Schools Achieve Both Quality and Equity? Investigating the Two Dimensions of Educational Effectiveness

L. Kyriakides

Department of Education, University of Cyprus, Cyprus

B. P. M. Creemers

Faculty of Psychology, Education, and Sociology, University of Groningen, The Netherlands

This article investigates the extent to which schools can achieve both equity and quality. Data emerged from two effectiveness studies in teaching mathematics and Greek language, which were conducted to test the validity of the dynamic model of educational effectiveness. Separate multilevel analyses for each subject were conducted and it was found that the effectiveness status of schools does not change significantly when the two dimensions (equity and quality) are used to measure their effectiveness status in each subject. Changes in their effectiveness status in terms of each dimension of measuring effectiveness were also examined. In each subject, schools that were found to improve their effectiveness status in terms of the equity dimension were also found to improve their effectiveness status in terms of the quality dimension. Moreover, no school that was found to improve its effectiveness status in terms of one dimension had declining effectiveness in terms of the other dimension of effectiveness. Implications of these findings for the development of educational effectiveness research are drawn.

At the outset of instruction in any topic, students of any age and in any culture will differ from one another in various intellectual and psychomotor skills, generalized and specialized prior knowledge, interests and motives, socio-economic background, and personal styles of thoughts and work during learning (Tomlinson, 1999). This argument has a strong history in Educational Effectiveness Research (EER). The first effectiveness studies during the 1970s (Edmonds, 1979; Rutter, Maughan, Mortimore, Ouston, & Smith, 1979) were concerned with examining evidence and making an argument about the potential power of schooling to make a difference to students' life chances. During the last 3 decades, publication of these studies was followed by numerous studies in different countries into school effectiveness and school improvement efforts, aimed at putting the results of research into practice (Teddlie & Reynolds, 2000; Townsend, 2007). A major aim of effectiveness studies was to support teachers and schools in their attempt to provide equal opportunities to their students with different learning

Correspondence should be addressed to Leonidas Kyriakides, Department of Education, University of Cyprus, PO Box 20537, Nicosia, CYPRUS 1678. E-mail: kyriakid@ucy.ac.cy

needs arising from their background and personal characteristics. Coming from the history of research in 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 amount in 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) and have not been able to contribute significantly to the improvement of education for different groups of students (Creemers & Kyriakides, 2006). Thus, critics of EER argue that there are really no grounds for thinking that EER can overcome the effects of social inequality. Although greater effectiveness may somewhat improve the absolute performance of disadvantaged groups, critics argue it will not improve their relative performance against more advantaged groups (Thrupp, 2001). However, Reynolds and Teddlie (2001, p. 112) claimed that: ''With their pessimism, passivity and inability to do anything more than talk about change, it is the critics that are the true conservatives now as in the 1960s.'' Nevertheless, researchers of educational effectiveness may have to examine further issues dealing with equity in education. For example, research into differential teacher and school effectiveness (e.g., Kyriakides, 2004; Strand, 2010) 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, p. 141) 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.'' Findings concerning differential school effects emerging from studies conducted in the United States support the conclusion that schools matter most for underprivileged and/or initially low-achieving students (Scheerens & Bosker, 1997). Thus, research into differential effectiveness may raise issues regarding the development and implementation of policy on educational equality. If schools differ significantly in terms of their effectiveness for particular pupil groups, 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). The identification of these factors may also be useful for policymakers attempting to design and implement policies on equal opportunities. In this context, the dynamic model of educational effectiveness (Creemers & Kyriakides, 2008) has been developed, attempting to demonstrate the complexity of improving educational effectiveness by taking into account the major findings of research on differential teacher and school effectiveness (e.g., Campbell et al., 2004; Kyriakides, 2004; Strand, 2010). The dynamic model is also based on the assumption that both the quality and the equity dimension of educational effectiveness should be considered in establishing criteria for measuring effectiveness.

DIMENSIONS OF EFFECTIVENESS: QUALITY AND EQUITY

It is expected generally in society that education should achieve high results in different domains of learning and subject areas. This means that the criteria for effectiveness will assess

the levels obtained by individual students, classes, and schools with respect to those objectives—excellence.

However, it is also possible to look at the effectiveness of a school from a different angle, equity, especially through investigating how far schools and teachers managed to reduce unjustifiable differences in outcomes of schooling (Sammons, 2010). This results in educational objectives and criteria for educational effectiveness that are not related to a specific objective and specific students, but related to different groups of students in relationship to each other. The idea is that education can contribute to social justice and democracy by closing the gap between students with regard to their background, especially their abilities and the sociocultural status of their family. As a consequence, the early school effectiveness research and school improvement projects were determined, more or less, by the idea of creating effective schools for the urban poor (Edmonds, 1979). The 1980s saw heavy criticism of this kind of school improvement and research with its conspicuous sampling biases (Firestone & Herriot, 1982; Good & Brophy, 1986; Purkey & Smith, 1983; Ralph & Fennessey, 1983; Rowan, Bossart, & Dwyer, 1983). As a result, EER is nowadays more realistic and modest in its beliefs about the contributions educational effectiveness can have in promoting equity. However, research into educational effectiveness reveals that teachers and schools matter most for underprivileged and/or initially low-achieving students (Kyriakides, 2004; Scheerens & Bosker, 1997). This reveals the importance of using both dimensions of measuring effectiveness—quality and equity—in building theoretical models of educational effectiveness.

Because schools are primarily places where learning takes place, the objectives of education are primarily student learning outcomes. These can be found in the cognitive domain and also in the affective, social, and aesthetic domains. Student learning in noncognitive areas is determined overwhelmingly by other actors in the society, whereas the cognitive domain is determined less by other social agents. Thus, schools have a specific role in the cognitive domain and, consequently, objectives in this area are crucial for the educational system in general. Moreover, achievement of cognitive outcomes determines, to some extent, achievement in other domains, like motivation and well-being (Steinmayr & Spinath, 2009; van Der Werf, Opdenakker, & Kuyper, 2008). However, this does not mean that education should be restricted to cognitive objectives, because only a partial relationship between achievement of cognitive and noncognitive domains exists (Kyriakides, 2005; Opdenakker & Van Damme, 2000). Therefore, the cognitive basis for noncognitive areas can be offered. Moreover, schools may act as social agents. They can provide a social and aesthetic environment in which social behavior and aesthetic attitudes can be developed. Schools and teachers should be supported in such a way that objectives are reached and educational quality becomes a fact. In this context, research can offer insight into which factors and variables contribute to student results. Thus, in this article, the term *learning outcomes* is used in a broader sense.

Both quality and equity are treated as criteria for measuring effectiveness. In the case of quality, student achievement gains in both the cognitive domain and other domains are examined (Creemers, Kyriakides, & Sammons, 2010). In regard to the equity dimension, it is taken into account that the concept of equity is subject to several interpretations (Demeuse, Crahay, & Monseur, 2001; Organisation for Economic Co-operation and Development, 2004). Philosophers have been struggling for a long time to clarify what might be meant in social policy by the term *equity*. However, there is general agreement that the aim of public policy cannot and should not be equality in the sense that everyone is the same or achieves the same outcomes—a statement that

appears to be both impossible and undesirable (see Organisation for Economic Co-operation and Development, 2004; Levin, 2003). Rather, a commitment to equity suggests that differences in outcomes should not be attributable to differences in areas such as wealth, income, power, or possessions (Levin, 1995). The question is always a practical one, then, of what state or degree of inequality is acceptable. The answer to this question will always be a contested one, fought out in political arenas of all kinds. The grounds of that struggle seem to have shifted in the last 30 years toward reducing the gap in outcomes between the top and bottom by helping those at the bottom move up (Demeuse, 2004; Vernez, Krop, & Rydell, 1999). Thus, Creemers and Kyriakides (2008) argued that the equity dimension can be measured by looking at the extent to which schools and teachers manage to reduce unjustifiable differences in outcomes of schooling. Figure 1 illustrates the way that measures of effectiveness in relation to dimensions and domains of learning can be conceptualized. This implies that effectiveness studies should search for any interaction between the two dimensions, as well as among the domains of learning.

However, almost all effectiveness studies measure school effectiveness in relation to the quality dimension (Sammons, 2010). As a consequence, a debate about the use of appropriate value-added models to measure the effectiveness of schools is in place in most countries (Goldstein & Spiegelhalter, 1996). Moreover, accountability systems have been developed in several countries that treat student progress as the main criterion for evaluating teachers and schools (Ray, 2006; Sanders & Horn, 1994; Teddlie, Stringfield, & Burdett, 2003). Furthermore, the great majority of effectiveness studies attempt to identify teacher- and school-level factors that are associated with student achievement. For example, the various studies that were conducted to test the validity of the dynamic model were concerned with the impact of teacher and school factors upon student achievement in different cognitive and affective aims of schooling (e.g., Creemers & Kyriakides, 2010a, 2010b; Kyriakides & Creemers, 2008, 2009). None of these studies was looking at the extent to which the model can explain variation in teacher and school contribution toward the reduction of gaps among students of different ability groups—equity.

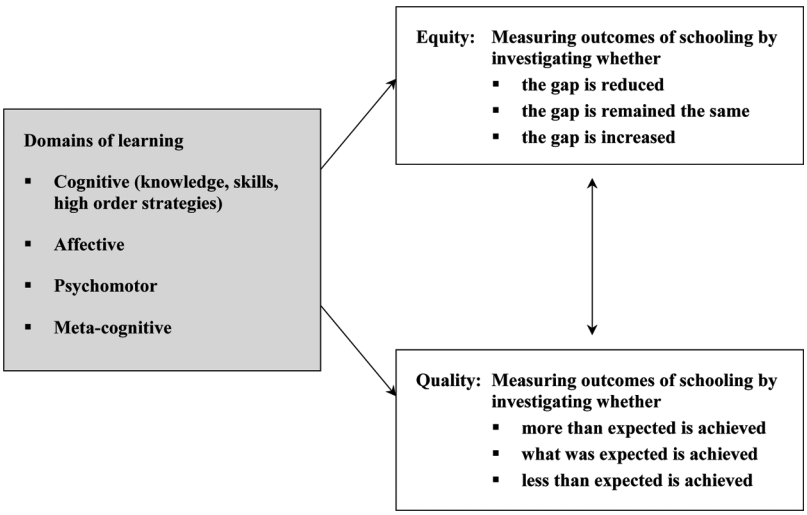

FIGURE 1 Dimensions of measuring effectiveness concerned with outcomes of schooling.

RESEARCH AIMS

This article proposes an approach to measuring school effectiveness in relation to the equity dimension. Moreover, we attempt to identify the extent to which the effectiveness status of each school changes or does not change significantly when the two dimensions of effectiveness (i.e., equity and quality) are used to measure their effectiveness status. To achieve this aim, we draw from the data emerging from the original and the follow-up study conducted to test the validity of the dynamic model. Because the follow-up study took place in the same 50 schools where the original study took place, we also attempt to identify changes in the effectiveness status of our school sample in terms of each dimension of measuring effectiveness, and to investigate the extent to which similar changes in their effectiveness status in terms of quality and equity can be identified. In this way, we not only investigate the extent to which schools can achieve both equity and quality, but we also try to identify whether similar changes in terms of the equity and the quality dimension of effectiveness can be observed over time.

METHODS

Because the major aim of conducting the follow-up study to test the validity of the dynamic model was to search for changes in the functioning of school factors and relate them to changes in the effectiveness status of the school (see Creemers & Kyriakides, 2010a), the design of the follow-up study was identical to several aspects of the original study conducted in order to test the validity of the dynamic model (Creemers & Kyriakides, 2010b; Kyriakides & Creemers, 2008). First, this study took place in the same schools. Since the student body and most teaching staff of these schools has changed, by collecting data from the same 50 schools where the original study was conducted, it was possible to measure the changes or stability in the effectiveness status of these schools in relation to each dimension. Second, the same instruments to measure teacher and school factors were used. By using the same instruments, it was possible to ensure that differences between the results of the two studies are not due to different measurement methods. Finally, the same age group of students is used in order to avoid problems that may arise due to the fact that some factors may have differential effects on achievement students of different ages. The next part of the methods section discusses the participants and methods used to conduct the original and the follow-up study.

Participants

The school sample of the follow-up study was identical to the one used for the original study and consisted of 50 primary schools. In both studies, all sixth-grade students from each class of the school sample were chosen. The chi-square test did not reveal any statistically significant difference between the sample of each study and the population in terms of students' sex (original study: $\chi^2 = 0.84$, $df = 1$, $p = 0.36$; follow-up study: $\chi^2 = 0.69$, $df = 1$, $p = 0.79$). 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 (original: $t = 1.21$, $df = 107$, $p = 0.22$; follow-up study: $t = 1.62$, $df = 111$, $p = 0.11$). Although these two studies refer to other variables such as the socioeconomic status (SES) of students and their achievement levels in different outcomes

of schooling, there is no data about these characteristics of the Greek Cypriot sixth-grade students. Therefore, it was not possible to examine whether the sample of each study was nationally representative in terms of any characteristics other than sex and class size. However, in each study, a nationally representative sample of Cypriot sixth-grade students, in terms of these two characteristics, was drawn.

The overall teacher and student sample size for each study is shown in Appendix A, along with a variety of descriptive statistics on the student background variables of the study. For each study, students with missing prior attainment or background data (less than 7% of the original sample of each study) were excluded from the analyses. This table also shows that there is no statistically significant difference in any of the background characteristics between the samples of these two studies. It is, however, important to note that in Cyprus, teacher appointments in all public primary schools are the responsibility of the educational service committee and each teacher is appointed to a school for a maximum period of 5 years (Kyriakides, 1999). Therefore, the great majority of teachers and headteachers who were teaching at a schools in 2004–2005 were not serving at the same school when the follow-up study took place. Despite the instability this policy creates, the trade union of Cypriot teachers, very strong politically, is in favor of this policy. As a consequence, during the last 20 years, no Minister of Education has ever attempted to change the teacher appointment and transfer system (Kyriakides, Demetriou, & Charalambous, 2006).

Variables

Output measures. 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 sixth grade and at the end of sixth grade. The construction of the tests was subject to controls for reliability and validity. For each subject, the Extended Logistic Model of Rasch (Andrich, 1988) was used to analyze the data at the beginning and at the end of school year, separately, and two scales per subject were created. Analysis of the data on student achievement revealed that each scale had relatively satisfactory psychometric properties. Specifically, for each scale the indexes of cases (i.e., students) and item separation were higher than 0.80, 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. It can, therefore, be claimed that each analysis revealed that there was a good fit to the model (Keeves & Alagumalai, 1999). Thus, for each student participating in each of these two studies, it was possible to generate two different scores for achievement in each subject at the beginning of year 6 and at the end of year 6, by calculating the relevant Rasch person estimate.

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., primary school, secondary school, or college/university), the social status of father's job, the social status of mother's job, and the economic situation of the family. Following the classification of occupations used by the Ministry of Finance, it was possible

to classify parents' occupations into three groups with relatively similar sizes: working class, middle class, and upper-middle class. Representative parental occupations for the working class are farmer, truck driver, and machine operator in a factory; for the middle class are police officer, teacher, and bank officer; and for the upper-middle class are doctor, lawyer, and business executive. Relevant information for each child was taken from the school records. Then standardized values of these five variables were calculated, resulting in the SES indicator.

Data on teacher- and school-level factors included in the dynamic model were also collected. Eight factors describe teachers' instructional role: orientation, structuring, questioning, teaching-modeling, applications, management of time, teacher role in making the classroom a learning environment, and classroom assessment. These eight factors do not only refer to one approach to teaching such as the direct teaching model or the constructivist approach. An integrated approach in defining quality of teaching is adopted. In addition, the dynamic model refers to school factors that are related to the same key concepts of quantity of teaching, provision of learning opportunities, and quality of teaching that were used to define teacher-level factors. In particular, the following four overarching school factors are included in the model and were taken into account for the design of these studies:

1. School policy for teaching and actions taken to improve teaching practice,
2. Evaluation of school policy for teaching and of actions taken to improve teaching,
3. Policy for creating a school learning environment (SLE) and actions taken to improve the SLE, and
4. Evaluation of the SLE.

Because data on these factors are not drawn to estimate the effectiveness status of schools, the instruments and the processes used to measure these factors are not presented in this article. However, information on the measurement of teacher and school factors are reported in papers presenting results of these studies concerned with the impact of these factors on student achievement (see Creemers & Kyriakides, 2010a, 2010b; Kyriakides & Creemers, 2008).

Method of Analysis

For the purposes of analyzing the data that emerged from each study, separate multilevel analyses for each subject were conducted. To measure the effectiveness status of schools in terms of the quality dimension, the following approach was used. For each subject, the first step in the analysis was to determine the variance of final achievement at individual, class, and school level without explanatory variables (empty model). For the purposes of this article, prior achievement and background factors were only controlled to estimate the schools' value-added contributions. These are typically referred to as the effectiveness scores of schools, but they also reflect other unmeasured factors (outside the control of the school) which were not controlled for in the analysis (Thomas, 2001). Therefore, based on the results of Model 1, which emerged by adding student prior attainment and background factors into the empty model (see Appendix B), the difference between the expected and the actual score for each school was plotted. The standard error of estimate for each school was also taken into account and is represented by the length of a vertical line. This line can be conceptualized as the range within which we are 95% confident that the true estimate of the school's residual lies (Goldstein, 2003). Thus, where this

vertical line does not cross the horizontal zero line and is also situated below the zero line, the school it represents is considered as one of the least effective schools of our sample. On the other hand, where this line does not cross the horizontal zero line and is situated above the zero line, the school it represents is characterized as one of the most effective schools. All the other schools are characterized as typical.

To estimate the effectiveness status of schools in terms of the equity dimension, the following approach was used. Initially, the variance of final achievement of students at the level of classroom was calculated. The variance at classroom level of student achievement emerging from the baseline test was also calculated. Then, the difference between the two variances for each classroom was calculated. In this way, it was possible to determine the extent to which the initial variance on student achievement was reduced by the end of the school year. This measure was seen as an indicator of the equity dimension of effectiveness in each subject and was treated as our dependent variable. At the next stage, a two-level model was used to determine the variance of our dependent variable at class, and school level without explanatory variables (empty model). Then context variables at classroom, and school level were added to the empty model. Based on the results of this model (see Appendix C), it was possible to plot the difference between the expected and the actual score for each school. As in the case of the quality dimension, the standard error of estimate for each school was taken into account and was represented by the length of a vertical line. Thus, where this vertical line does not cross the horizontal zero line and is also situated below the zero line, the school it represents is considered as one of the least effective schools of our sample in terms of the equity dimension. On the other hand, where this line does not cross the horizontal zero line and is situated above the zero line, the school it represents is characterized as one of the most effective schools in terms of the equity dimension. All the other schools are characterized as typical.

RESULTS

Following the approach previously described, we classified the school sample of each study according to their effectiveness status in relation to the quality and equity dimension in each subject separately. Tables 1 and 2 refer to the results of the analyses that emerged from the data of the original study in teaching mathematics and Greek language respectively. Specifically, these tables illustrate the distribution of the school sample according to their effectiveness status in

TABLE 1
The Distribution of the School Sample According to Their Quality and Equity Status in Regard to Teaching Mathematics Emerged from the Original Study

	Quality			
Equity	*Least Effective*	*Typical*	*Most Effective*	*Total Number of Schools*
Negative effect	8	1	0	9
No effect	3	22	5	30
Positive effect	0	3	8	11
Total number of schools	11	26	13	50

TABLE 2
The Distribution of the School Sample According to Their Quality and Equity Status in Regard to Teaching Greek Language Emerged from the Original Study

	Quality			
Equity	*Least Effective*	*Typical*	*Most Effective*	*Total Number of Schools*
Negative effect	9	1	0	10
No effect	2	24	4	30
Positive effect	0	1	9	10
Total number of schools	11	26	13	50

TABLE 3
The Distribution of the School Sample According to Their Quality and Equity Status in Teaching Mathematics Emerged from the Replication Study

	Quality			
Equity	*Least Effective*	*Typical*	*Most Effective*	*Total Number of Schools*
Negative effect	9	1	0	10
No effect	3	24	4	31
Positive effect	0	0	9	9
Total number of schools	12	25	13	50

relation to the quality and equity dimension in each subject separately. The relevant distributions that emerged from analyzing the data of the follow-up study are presented in Tables 3 and 4.

These four tables show that no school was considered as among the most effective in terms of one dimension of measuring effectiveness and at the same time among the least effective in terms of the other. Moreover, the majority of the schools that were considered as among the most effective in terms of the quality dimension were also found to contribute to the reduction of initial achievement gaps among students. For example, the figures of the fourth column of Table 1 reveal that eight out of the 13 most effective schools in mathematics had a positive impact on reducing the initial achievement gap among their students. Furthermore, the majority

TABLE 4
The Distribution of the School Sample According to Their Quality and Equity Status in Regard to Teaching Greek Language Emerged from the Replication Study

	Quality			
Equity	*Least Effective*	*Typical*	*Most Effective*	*Total Number of Schools*
Negative effect	9	1	0	10
No effect	2	23	3	28
Positive effect	0	2	10	12
Total number of schools	11	26	13	50

TABLE 5
The Distribution of the School Sample According to Their Effectiveness Status in Terms of the Quality Dimension During the School Year 2004–2005 and During the School Year 2008–2009 per Subject

Groups of Schools	*Mathematics*	*Language*
A) Stability		
Remain typical	14	15
Remain least effective	6	6
Remain most effective	7	6
B) Improvement		
From least effective to typical	5	5
From least effective to most effective	0	0
From typical to most effective	6	7
C) Declining		
From most effective to typical	6	6
From typical to least effective	6	4
From most effective to least effective	0	1

of the least effective schools in each subject had a negative impact on reducing the achievement gap identified at the beginning of the school year. These findings imply that the measured effectiveness of schools does not change dramatically when the two dimensions (quality and equity) are used to measure their effectiveness in each subject.

At the next step, for each subject, it was possible to compare the effectiveness status of each school during the 2004–2005 school year in each dimension of measuring effectiveness with its relevant effectiveness status during the 2008–2009 school year. Table 5 illustrates the distribution of changes in the quality dimension of the effectiveness status of our school sample in each subject separately; Table 6 illustrates the relevant distribution of changes in the effectiveness status of schools in terms of the equity dimension.

TABLE 6
The Distribution of the School Sample According to Their Effectiveness Status in Terms of the Equity Dimension During the 2004–2005 School Year and During the 2008–2009 School Year per Subject

Groups of Schools	*Mathematics*	*Language*
A) Stability		
Remain typical	18	18
Remain least effective	3	4
Remain most effective	4	6
B) Improvement		
From least effective to typical	6	6
From least effective to most effective	0	0
From typical to most effective	5	6
C) Declining		
From most effective to typical	7	4
From typical to least effective	7	6
From most effective to least effective	0	0

TABLE 7
The Distribution of the School Sample According to the Type of Changes Observed in Their Effectiveness Status in Each Dimension in Teaching Mathematics

	Quality			
Equity	*Decline*	*Stability*	*Improvement*	*Total Number of Schools*
Decline	9	5	0	14
Stability	3	21	1	25
Improvement	0	1	10	11
Total number of schools	12	27	11	50

The following observations arise from these two tables. First, for each subject, no change in the effectiveness status of at least 50% of our school sample in either the quality or the equity dimension can be observed. For example, 27 out of 50 schools managed to remain equally effective in terms of the quality dimension in mathematics (see Table 5). We also observed that 25 out of 50 schools managed to remain equally effective in terms of the equity dimension in mathematics (see Table 6). Second, in each subject, more than 10 schools managed to improve their effectiveness statuses in regard to either the equity or the quality dimension. For example, 12 schools managed to improve their effectiveness status in terms of the quality dimension in the subject of language and 11 schools managed to improve their effectiveness status in terms of the equity dimension in mathematics. In each subject, a decline in the effectiveness status of almost a quarter of the school sample was also observed in terms of each dimension. Third, extreme changes in the effectiveness status of the schools are observed in only one school, which dropped down from the most to least effective in terms of the quality dimension in language.

Finally, for each subject, we attempted to compare the schools in terms of the type of change in their effectiveness status on each of the two dimensions of measuring effectiveness (see Tables 7 and 8). It is important to note that no school both improved its effectiveness status in terms of one dimension and saw its effectiveness status decline in terms of the other dimension of measuring effectiveness. Moreover, all schools that managed to improve their effectiveness status in terms of the equity dimension in Greek language managed to improve their effectiveness status in terms of the quality dimension (see Table 8). In regard to teaching mathematics, 10 out of 11 schools that managed to improve their effectiveness status in terms of the

TABLE 8
The Distribution of the School Sample According to the Type of Changes Observed in Their Effectiveness Status in Each Dimension in Teaching Greek Language

	Quality			
Equity	*Decline*	*Stability*	*Improvement*	*Total Number of Schools*
Decline	8	2	0	10
Stability	3	25	0	28
Improvement	0	0	12	12
Total number of schools	11	27	12	50

equity dimension managed also to improve their effectiveness status in terms of the quality dimension. In addition, for both subjects, the great majority of schools that became less effective in terms of the quality dimension also saw their effectiveness status decline in terms of the equity dimension.

DISCUSSION

Implications of the findings for the development of EER are drawn. Reanalysis of the two studies testing the validity of the dynamic model reveals that the effectiveness status of schools does not change dramatically when the two dimensions (quality and equity) are used to define their effectiveness status in each of the two core subjects of the primary curriculum in Cyprus. Moreover, no school was found to be both among the most effective schools in terms of the equity dimension and among the least effective schools in terms of the quality dimension. However, only a few schools managed to be among the most effective in terms of both the quality and the equity dimension in each subject. These findings seem to support the more realistic and modest belief of the current phase of EER about the contribution that schools can have in promoting quality and equity. Nevertheless, the findings of this study also reveal that schools that are among the most effective in terms of equity do not risk being among the least effective in terms of the quality dimension. Therefore, we should seek to expand the agenda of EER to identify factors that can explain why some schools that are among the most effective in terms of quality can also reduce initial achievement gaps whereas other schools are only effective in terms of the quality dimension. Longitudinal and case studies should be conducted to identify effective strategies used by some schools that helped them to improve not only in terms of the quality dimension but also in terms of the equity dimension. For example, a case study conducted by Fisher, Frey, and Lapp (2011) shows how a school managed to improve its effectiveness in terms of equity by focusing on attendance and student engagement. By developing a schoolwide plan that ensured that attendance was noticed, corrected, and celebrated, students became involved in learning and their achievement improved. These initiatives, which seem to be in line with the factor of the dynamic model concerned with the school policy of teaching, further closed the achievement gap. By identifying factors associated with both the quality and the equity dimension of school effectiveness, EER may contribute toward the development and dissemination of research-based school improvement designs that promote effectiveness in terms of both the quality and equity dimensions (e.g., Datnow & Stringfield, 2000; Jordan, McPartland, Legters, & Balfanz, 2000; Slavin & Madden, 2000).

Looking at the changes in our schools, similar observations can be drawn when either the quality or the equity dimension is used to measure the effectiveness status of schools. Changes in the effectiveness status of a significant number of schools in terms of either quality or equity can be observed. Although almost 55% of schools remained equally effective and dramatic changes were not observed, improvement in the effectiveness status of more than 20% of our schools was also observed. This finding provides support to studies investigating school effectiveness over a long period of time, which reveal that there are limits to improvement over a long period of time and that a relatively small proportion of schools appear to have significant improvement patterns (Thomas, Peng, & Gray, 2007). Although these studies look only at the quality dimension, the same observations can be drawn by looking at changes in equity. But

although improvement appears to be difficult, the reanalyses of these two studies also reveal that improvement in both quality and equity is not impossible.

Tables 7 and 8 reveal that almost all schools that managed to improve in quality managed also to improve on equity. Thus, promoting one dimension of school effectiveness does not negatively influence the other dimension, but is likely to influence it positively. Thus, our findings indicate that schools should provide extra support to those who need it, to improve in equity and quality.

This article also draws attention to the importance of looking at changes in the effectiveness status of schools in terms of both dimensions. Rather than treating achievement of a single group of students as the dependent variable implying stability in school effectiveness, our attempt to search for changes in the effectiveness status of schools in each dimension reveals that the main aim of effectiveness studies should move from understanding variation in the effectiveness status of a number of schools to understanding why changes in the effectiveness status of schools in both quality and equity are observed. Such studies could help researchers expand the theoretical framework of educational effectiveness by identifying factors contributing to changes in the effectiveness status of schools in terms of both quality and equity. Because all school teachers and head teachers had been recently changed, changing school personnel does not explain changes in quality or equity, especially because most schools remained equally effective in terms of both the quality and the equity dimensions (see Tables 7 and 8). However, Creemers and Kyriakides (2010a) showed that changes in the effectiveness status of schools in terms of quality can be explained by looking, first of all, at changes in quality of teaching practice and then at changes in the functioning of most school factors of the dynamic model. Further studies may also help identify factors that can explain changes in the effectiveness status in terms of both quality and equity.

Finally, some suggestions for further research can be provided. First, further research is indicated to test the degree to which the findings reported here may be generalized. These studies may investigate the extent to which, in different phases of schooling (e.g., preprimary, primary, or secondary), schools can achieve both quality and equity, and whether similar changes in the effectiveness status of schools can be observed when both quality and equity are used as dimensions of evaluating schools. Moreover, the impact of the outcomes used to define the two measurement dimensions can be investigated. For example, further research could investigate the relation between equity and quality when achievement in affective outcomes of learning is considered. Finally, a group randomization study could be conducted to identify the impact of using a theory-driven and evidence-based approach to promote equity in education. For example, a framework based on research on equity in education and on the dynamic model of educational effectiveness could be offered to one group of schools to help them identify what could be achieved, and how, to promote equity. A second group of schools could be encouraged to establish school self-evaluation mechanisms and develop their own improvement strategies to promote equity. Based on the results of this experimental study, researchers should not only measure the impact of the proposed approach to promoting equity but also reveal its added value by comparing its impact with the impact of an approach to school improvement which is advocated widely. In this way, research investigating the quality and equity dimensions of educational effectiveness may not only contribute to the development of the theoretical framework of EER but may also provide suggestions on how school policy and practice could be improved in order to promote both quality and equity in education.

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. doi:10.1207/s15324818ame0104_7.

Campbell, R. J., Kyriakides, L., Muijs, R. D., & Robinson, W. (2004). *Assessing teacher effectiveness: A differentiated model*. London, England: Routledge Falmer.

Creemers, B. P. M., & Kyriakides, L. (2006). A critical analysis of the current approaches to modelling educational effectiveness: the importance of establishing a dynamic model. *School Effectiveness and School Improvement*, *17*, 347–366. doi:10.1080/09243450600697242.

Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness*. London, England: Routledge.

Creemers, B. P. M., & Kyriakides, L. (2010a). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement*, *21*, 409–427. doi:10.1080/09243453.2010.512795.

Creemers, B. P. M., & Kyriakides, L. (2010b). School factors explaining achievement on cognitive and affective outcomes: Establishing a dynamic model of educational effectiveness. *Scandinavian Journal of Educational Research*, *54*(3), 263–294. doi:10.1080/00313831003764529.

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

Datnow, A., & Stringfield, S. (2000). Working together for reliable school reform. *Journal of Education for Students Placed At Risk*, *5*, 183–204. doi:10.1080/10824669.2000.9671386.

Demeuse, M. (2004). A set of equity indicators of the European education system—A synthesis. In L. Moreno Herrera & G. Francia (Eds.), *Educational policies: Implications for equity, equality and equivalence* (pp. 43–47). Reports from the Department of Education, 6). Örebro, Sweden: Örebro University.

Demeuse, M., Crahay, M., & Monseur, C. (2001). Efficiency and equity. In W. Hutmacher, D. Cochrane, & N. Botani (Eds.), *In pursuit of equity in education: Using international indicators to compare equity policies* (pp. 65–91). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership*, *37*(10), 15–24. Retrieved from http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_197910_edmonds.pdf.

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.

Firestone, W. A., & Herriott, R. E. (1982). Prescriptions for effective elementary schools don't fit secondary schools. *Educational Leadership*, *40*, 51–53. Retrieved from http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_198212_firestone.pdf.

Fisher, D., Frey, N, & Lapp, D. (2011). Focusing on the participation and engagement gap: A case study on closing the achievement gap. *Journal of Education for Students Placed at Risk (JESPAR)*, *16*, 56–64. doi:10.1080/10824669.2011.545976.

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

Goldstein, H., & Spiegelhalter, D. J. (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society, Series A*, *159*, 385–443. Retrieved from http://www.jstor.org/stable/2983325.

Good, T. L., & Brophy, J. E. (1986). School effects. In M. C. Wittrock (Ed.). *Handbook of research on teaching* (3rd ed., pp. 570–602). New York: MacMillan.

Jordan, W., McPartland, J., Legters, N., & Balfanz, R. (2000). Creating a comprehensive school reform model: The Talent Development High School with career academies. *Journal of Education for Students Placed At Risk*, *5*, 159–181. doi:10.1080/10824669.2000.9671385.

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, England: Pergamon.

Kyriakides, L. (1999). The management of curriculum improvement in Cyprus: A critique of a 'centre-periphery' model in a centralised system. In T. Townsend, P. Clarke, & M. Ainscow (Eds.), *Third millennium schools. A world of difference in school effectiveness and school improvement* (pp. 107–124). Lisse, The Netherlands: Swets and Zeitlinger.

Kyriakides, L. (2004). Differential school effectiveness in relation to sex and social class: Some implications for policy evaluation. *Educational Research and Evaluation*, *10*, 141–161. doi:10.1076/edre.10.2.141.27907.

Kyriakides, L. (2005). Extending the comprehensive model of educational effectiveness by an empirical investigation. *School Effectiveness and School Improvement*, *16*, 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, The Netherlands: Springer.

Kyriakides, L., & Creemers, B. P. M. (2008). Using a multidimensional approach to measure the impact of classroom level factors upon student achievement: A study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, *19*, 183–306. doi:10.1080/09243450802047873.

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*, 61–86. doi:10.1080/19415530903043680.

Kyriakides, L., Demetriou, D., & Charalambous, C. (2006). Generating criteria for evaluating teachers through teacher effectiveness research. *Educational Research*, *48*, 1–20. doi:10.1080/00131880500498297.

Levin, B. (1995). Education and poverty. *Canadian Journal of Education*, *20*, 211–224.

Levin, B. (2003, August). *Approaches to equity in policy for lifelong learning*. Paper commissioned by the Education and Training Policy Division, OECD, for the Equity in Education Thematic Review.

Opdenakker, M. C., & Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes. *School Effectiveness and School Improvement*, *11*, 65–196. doi:10.1076/0924-3453(200006)11:2;1-Q;FT165.

Organisation for Economic Co-operation and Development. (2004). *Equity in education: Students with disabilities, learning difficulties and disadvantages*. Paris: Centre for Educational Research and Innovation Education & Skills.

Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *Elementary School Journal*, *83*, 427–452. Retrieved from http://www.jstor.org/stable/1001168.

Ralph, J. H., & Fennessey, J. (1983). Science or reform: Some questions about the effective schools model. *Phi Delta Kappan*, *64*, 689–694.

Ray, A. (2006). *School value added measures in England: A paper for the OECD project on the development of value-added models in education systems*. London, England: Department of Education Skills.

Reynolds, D., & Teddlie, C. (2001). Reflections on the critics and beyond them. *School Effectiveness and School Improvement*, *12*, 99–114. doi: 10.1076/sesi.12.1.99.3460.

Rowan, B., Bossart, S. T., & Dwyer, D. C. (1983). Research on effective schools. A cautionary note. *Educational Researcher*, *12*(4), 24–32. doi:10.3102/0013189 × 012004024.

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

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, England: Elsevier.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, *8*, 299–311. doi:10.1007/BF00973726.

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

Slavin, R. E., & Madden, N. A. (2000). Roots & wings: Effects of whole-school reform on student achievement. *Journal of Education for Students Placed At Risk (JESPAR)*, *5*, 109–136. doi:10.1080/10824669.2000.9671383.

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*, 80–90. doi:10.1016/j.lindif.2008.05.004.

Strand, S. (2010). Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement*, *21*, 289–314. doi:10.1080/09243451003732651.

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

Teddlie, C., Stringfield, S., & Burdett, J. (2003). International comparisons of the relations among educational effectiveness and improvement variables: An overview. *Journal of Personnel Evaluation in Education*, *17*, 5–19. doi:10.1023/A:1025020928735.

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.

Thomas, S., Peng, W. J., & Gray, J. (2007). Modelling patterns of improvement over time: Value added trends in English secondary school performance across ten cohorts. *Oxford Review of Education*, *33*(3), 261–295. doi: 10.1080/03054980701366116.

Thrupp, M. (2001). Recent school effectiveness counter-critiques: Problems and possibilities. *British Educational Research Journal*, *27*(4), 443–457. doi:10.1080/01411920120071452.

Tomlinson, C. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: Association for Supervision and Curriculum Development.

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

Van Der Werf, G., Opdenakker, M. C., & Kuyper, H. (2008). Testing a dynamic model of student and school effectiveness with a multivariate multilevel latent growth curve approach. *School Effectiveness and School Improvement*, *19*(4), 447–462. doi: 10.1080/09243450802535216.

Vernez, G., Krop, R., & Rydell, P. (1999). *Closing the education gap: Benefits and costs*. Santa Monica: RAND.

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

APPENDIX A

Table for Descriptive Statistics for the Dataset of Each Study and Statistical Figures of Tests Used to Compare the Background Characteristics of the Two Samples

Characteristics of Sample	*Original Study*	*Follow-Up Study*	*Statistical Figures Emerged from Comparing the Two Samples*
Original sample			
Number of pupils	2503	2716	NA
Number of teachers of grade 6	108	112	NA
Number of teachers	364	387	NA
Sample used in the analysis			
Number of pupils	2369	2682	NA
Percentage of girls	1239 (52.3%)	1432 (53.4%)	Chi-square test: ($\chi^2 = 0.60$, $df = 1$, $p = 0.44$)
Educational background of father			
Graduate of a primary school	829 (35%)	912 (34%)	Kolmogorov-Smirnov two sample test (K-S Z = 0.364, $p = 0.999$)
Graduate of secondary school	900 (38%)	1073 (40%)	
Graduate of a college/university	640 (27%)	697 (26%)	
Educational background of mother			
Graduate of a primary school	805 (34%)	912 (34%)	Kolmogorov-Smirnov two sample test (K-S Z = 0.341, $p = 0.999$)
Graduate of secondary school	995 (42%)	1100 (41%)	
Graduate of a college/university	569 (24%)	670 (25%)	
Father occupation			
Occupations held by working class	782 (33%)	939 (35%)	Kolmogorov-Smirnov two sample test (K-S Z = 0.710, $p = 0.695$)
Occupations held by middle class	876 (37%)	965 (36%)	
Occupations held by upper-middle class	711 (30%)	778 (29%)	
Mother occupation			
Occupations held by working class	875 (37%)	938 (35%)	Kolmogorov-Smirnov two sample test (K-S Z = 0.682, $p = 0.740$)
Occupations held by middle class	877 (37%)	1073 (40%)	
Occupations held by upper-middle class	617 (26%)	671 (25%)	
Financial situation of the family	$M = 2.02$ $SD = 1.12$	$M = 1.98$ $SD = 1.09$	*t*-test for independent samples ($t = 1.28$, $df = 5049$, $p = 0.20$)

APPENDIX B
Parameter Estimates and (Standard Errors) for the Analyses of Student Achievement in Greek Language and Mathematics of Each Study

	Follow-Up Study				*Original Study*			
	Greek Language		*Mathematics*		*Greek Language*		*Mathematics*	
Factors	*Model 0*	*Model 1*	*Model 0*	*Model 1*	*Model 0*	*Model 1*	*Model 0*	*Model 1*
Fixed part/intercept	.42 (.06)	.31 (.05)	−.41 (.09)	−.33 (.08)	−.31 (.08)	−.22 (.08)	.35 (.05)	.28 (.05)
Student level								
Prior knowledge		.43 (.11)		.38 (.05)		.39 (.05)		.45 (.10)
Sex (boys = 0, girls = 1)		.12 (03)		N.S.S.*		.19 (.08)		−.14 (.06)
Socioeconomic status (SES)		.19 (.10)		.19 (.07)		.30 (.06)		.30 (.12)
Classroom level								
Context								
Average knowledge		.10 (.04)		.10 (.04)		.12 (.05)		.28 (.10)
Average SES		NSS		NSS		.08 (.03)		.12 (.05)
Percentage of girls		NSS		NSS		NSS		−.05 (.02)
School level								
Context								
Average SES		NSS		NSS		NSS		NSS
Average knowledge		.07 (.03)		.06 (.03)		.09 (.04)		.11 (.05)
Percentage of girls		NSS		NSS		NSS		NSS
Variance								
School	9.5%	8.8%	.1%	8.7%	9.0%	8.2%	11.2%	9.8%
Class	18.0%	14.9%	6.8%	14.1%	14.7%	10.3%	14.8%	10.0%
Student	72.5%	29.0%	4.1%	29.5%	76.3%	31.3%	74.0%	30.2%
Explained		47.3%		47.7%		50.2%		50.0%
Significance test								
χ^2	924.3	605.2	915.6	499.6	815.6	507.2	1144.9	795.5
Reduction		319.1		416.0		308.4		349.4
Degrees of freedom		5		4		6		7
p-value		.001		.001		.001		.001

Note. Each model was estimated without the variables that did not have a statistically significant effect at .05 level.
NSS = No statistically significant effect at .05 level.

APPENDIX C
Parameter Estimates and (Standard Errors) for the Analyses on Reduction of Variance in Each Subject Emerged From the Original and the Follow-Up Study

	Greek Language				*Mathematics*			
	Original Study		*Follow-Up Study*		*Original Study*		*Follow-Up Study*	
Factors	*Model 0*	*Model 1*	*Model 0*	*Model 1*	*Model 0*	*Model 1*	*Model 0*	*Model 1*
Intercept	.39 (.05)	.33 (.05)	.36 (.04)	.30 (.04)	.26 (.05)	.22 (.05)	.20 (.04)	.17 (.04)
Classroom level								
Context								
Variance in socioeconomic status (SES)		−.32 (.09)		−.25 (.05)		−.28 (.09)		−.15 (.05)
Variance in gender = (0.50 − % of girls)^2		−.16 (.04)		−.14 (.04)		−.19 (.04)		−.08 (.04)
School level								
Context								
Variance in SES		−.14 (.03)		−.13 (.03)		−.11 (.03)		−.10 (.03)
Variance in gender = (0.50 − % of girls)^2		−.09 (.03)		−.08 (.03)		N.S.S.		−.05 (.02)
Variance								
School	25.9%	21.1%	27.3%	24.2%	27.8%	24.1%	28.9%	25.2%
Class	74.1%	54.2%	72.7%	50.6%	72.2%	52.1%	71.1%	52.8%
Explained		24.7%		25.2%		23.8%		22.0%
Significance test								
Deviance	1224.7	1015.2	763.9	661.7	824.3	715.2	503.9	401.7
Reduction		209.5		102.2		109.1		102.2
Degrees of freedom		4		4		3		4
p-value		.001		.001		.001		.001