THE IMPACT OF QUALITY OF TEACHING ON STUDENT OUTCOMES: IMPLICATIONS FOR POLICY AND PRACTICE ON TEACHER PROFESSIONAL DEVELOPMENT

Leonidas Kyriakides & Bert Creemers

INTRODUCTION

One of the key findings from decades of Educational Effectiveness Research (EER) is the importance of the classroom level as a predictor of pupil outcomes. Research has consistently shown not only that the classroom level can explain more of the variance in pupil outcomes than the school level, but that a large proportion of this classroom level variance can be explained by what teachers do in the classroom (Muijs & Reynolds, 2001). As a result of these findings classroom practise has become firmly integrated into theoretical and empirical models of educational effectiveness (e.g., Creemers & Kyriakides, 2008). This paper therefore aims to summarise key findings and developments in the area of Teacher Effectiveness Research (TER) and to discuss the main methodological and conceptual limitations of TER. We also refer to recent developments in the area which reveal the importance of identifying grouping of teacher factors associated with student achievement and present results of projects conducted in different countries which reveal that teaching skills can be grouped into specific developmental stages. Finally, a dynamic approach to teacher professional development is proposed and implications for policy and practice on teacher professional development are drawn.

MAJOR FINDINGS OF TEACHER EFFECTIVENESS RESEARCH

During the last 35 years, researchers have turned to teacher behaviours as predictors of student achievement in order to build up a knowledge base on effective teaching. This research has led to the identification of a range of behaviours which are positively related to student achievement (e.g., Brophy & Good, 1986; Creemers, 1994; Doyle, 1986; Galton, 1987; Muijs & Reynolds, 2000). The most consistently replicated findings of teacher effectiveness studies conducted in different countries link student achievement to the *quantity and pacing of instruction*. Amount learnt is related to opportunity to learn and achievement is maximised when teachers prioritise academic instruction and

allocate available time to curriculum-related activities. Consistent success is another significant factor associated with student achievement. To learn efficiently, students must be engaged in activities that are appropriate in difficulty level and suited to their current achievement levels and needs (Stallings, 1985).

Effective teachers are also expected to organise and manage the classroom environment as an efficient learning environment and thereby to maximise engagement rates (Creemers & Reezigt, 1996; Kyriakides, 2008). Doyle (1986) claims that key indicators of effective classroom management include: good preparation of the classroom and installation of rules and procedures at the beginning of the year, smoothness and momentum in lesson pacing, consistent accountability procedures, and clarity about when and how students can get help and about what options are available when they finish. As far as the actual teaching process is concerned, research into classroom discourse reveals that, although there is a great deal of teacher talk in the classes of effective teachers, most of it is academic rather than managerial or procedural, and much of it involves asking questions and giving feedback rather than extended lecturing (Cazden, 1986; Kyriakides & Creemers, 2008).

The findings summarised above deal with factors associated with the quantity of academic activity. The variables presented below concern the *form and quality of teacher's organised lessons* and can be divided into those that involve giving information (structuring), asking questions (soliciting) and providing feedback (reacting). In regard to the structuring factor, Rosenshine and Stevens (1986) point out that achievement is maximised when teachers not only actively present material but also structure it by: a) beginning with overviews and/or review of objectives; b) outlining the content to be covered and signalling transitions between lesson parts; c) calling attention to main ideas; and d) reviewing main ideas at the end. Summary reviews are also important since they integrate and reinforce the learning of major points. These structuring elements not only facilitate memorising of the information but allow for its apprehension as an integrated whole with recognition of the relationships between parts (Creemers & Kyriakides, 2008). Moreover, achievement is higher when information is presented with a degree of redundancy, particularly in the form of repeating and reviewing general views and key concepts. Clarity of presentation is also a consistent correlate of student achievement (Scheerens & Bosker, 1997; Seidel & Shavelson, 2007). Effective teachers are

able to communicate clearly and directly with their students without digression, speaking above students' levels of comprehension or using speech patterns that impair the clarity of what is being taught (Smith & Land, 1981; Walberg, 1986). Muijs and Reynolds (2000) indicate that the focus on teachers actively presenting materials should not be seen as an indication that traditional lecturing and drill approach is an effective teaching approach. Effective teachers ask a lot of questions and attempt to involve students in class discussion. There should also be a mix of product questions (i.e., expecting a single response from students) and process questions (i.e., expecting students to provide explanations), but effective teachers ask more process questions (Askew & William, 1995; Kyriakides & Creemers, 2008, 2009). Effective teachers also use seatwork or small group tasks since they provide needed practice and application opportunities. The effectiveness of seatwork assignments is enhanced when the teacher explains the work that students are expected to do and once the students are released to work independently the teacher circulates to monitor progress and provide help and feedback.

Classroom climate is a significant teacher factor. The climate is usually seen as associated with the behaviour of the stakeholders, whereas culture is seen as measuring the values and norms of the organization (Heck & Marcoulides, 1996; Hoy, Tater, & Bliss, 1990). Classroom climate research is described as the stepchild of psychological and classroom research (Creemers & Reezigt, 1996). The classroom effects research tradition initially focused on climate factors defined as managerial techniques (e.g., Doyle, 1986). Management is necessary to create conditions for learning and instruction, but management itself is not sufficient for student results (Creemers, 1994). On the other hand, the psychological tradition of classroom environment research paid a lot of attention to instruments for the measurement of students' perceptions of climate. Many studies report on their psychometric characteristics (Fraser, 1991). We advocate the need to integrate elements of different research traditions and search for the contribution of the teacher in creating the classroom as a learning environment. Effectiveness studies conducted during the last two decades (e.g., Kosir, 2005; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003; Slavin, 1983; Slavin & Cooper, 1999) reveal the importance of investigating the teacher's contribution in creating a learning environment in his/her classroom by taking into account the following elements of the classroom environment: teacher-

student interaction, student-student interaction, students' treatment by the teacher, competition between students, and classroom disorder (Creemers & Kyriakides, 2008; Kyriakides & Christoforou, 2011). The first two elements are important components of measuring classroom climate, as classroom environment research has shown (Cazden, 1986; den Brok, Brekelmans, & Wubbels, 2004; Fraser, 1991). However, teacher effectiveness research is concerned with the type of interactions that exist in a classroom rather than to how students perceive teacher interpersonal behaviour (Kyriakides, 2008). The other three elements refer to the attempt of teachers to create a businesslike and supportive environment for learning especially since research on teacher effectiveness reveals that the classroom environment should not be only businesslike but also needs to be supportive for the students (Walberg, 1986). Thus, effective teachers expect all students to be able to succeed and their positive expectations are transmitted to their students.

CONCEPTUAL AND METHODOLOGICAL LIMITATIONS OF TEACHER EFFECTIVENESS RESEARCH

This section deals with two conceptual problems of teacher effectiveness in the research literature. The first constraint has to do with the fact that most effectiveness studies are exclusively focused on language or mathematics rather than on the whole school curriculum aims (cognitive, meta-cognitive, and affective). This implies that EER should take into account the new goals of education and related to this their implications for teaching and learning. Moreover, new theories of teaching and learning should be used in order to specify variables associated with the quality of teaching.

Recent models of teaching and learning characterize learning as a self-regulated and constructive process (Bransford, Brown, & Cocking, 2000; Collins, Brown, & Newman, 1989). During the last decade, this characterization of teaching stimulated a substantial number of studies which were conducted in different countries around the world and are looking for the impact of new learning approaches to teaching on student outcomes (e.g., Brush, 1997; Nolen, 2003; Fuchs, Fuchs, Yazdian, & Powell, 2002; Ramsden, 1997). Although the effects of teaching on student learning which were identified by these studies were diverse and complex, two recent meta-analyses revealed that they were fairly systematic (Seidel & Shavelson, 2007; Kyriakides & Christoforou, 2011).

4

Instead of treating active and direct teaching approaches as in contrast with the new leaning approaches to teaching, an integrated approach to teaching should be adopted. In this context, the dynamic model of educational effectiveness (Creemers & Kyriakides, 2008), which was presented in the first paper at this conference, is not based only on traditional views on learning and instruction which emphasise the role of teacher as instructor responsible for providing knowledge and skills. The dynamic model takes into account new ideas on learning and instruction associated with constructivism which give emphasis to independent learning and the construction of knowledge by the learner (Simons, van der Linden, & Duffy, 2000). In the latter case, the role of the teacher gradually moves from instructing to coaching and modelling learning. Thus, the dynamic model advocates for the use of an integrated approach in defining quality of teaching and refers to those teacher factors that were found to be consistently related with student achievement irrespective of whether they are in line with one or the other approach to teaching (Creemers & Kyriakides, 2006). This assumption of the dynamic model is also in line with those indicating the limitations of using exclusively either the direct teaching approach (Steffe & Gale, 1995) or approaches associated with constructivism (Kirschner, Sweller, & Clark, 2006) to describe effective teaching. Therefore, suggestions for further research promoting an integrated approach to effective teaching can be drawn, especially since the great majority of effectiveness studies conducted during the last four decades are concerned with factors associated with only a single approach to teaching (Seidel & Shavelson, 2007).

A second constraint of the existing approaches to effective teaching is the fact that the process does not contribute significantly to teachers' professional development or to improving teachers' effectiveness. This is partly due to the fact that most studies are based on correlational research findings which look at variations in existing practices and even most of the experiments involved practices previously observed. Even if most of the results of teacher effectiveness research are transferable to the classroom and several professional development programmes such as the Active Mathematics Teaching (Good, Grouws, & Ebmeier, 1983) and the Teacher Effectiveness Enhancement Project (Muijs & Reynolds, 2000) have been developed, prescriptions for applications derived from these studies usually remain within the ranges of teacher behaviour which were observed (Kyriakides & Christoforou, 2011). This implies that we need to establish stronger links

between Teacher Effectiveness Research (TER) and research on teacher professional development and investigate the extent to which teachers and schools can make use of the knowledge base of TER in order to improve their practice.

A DYNAMIC APPROACH TO TEACHER PROFESSIONAL DEVELOPMENT

Research on teacher training and research on teacher effectiveness have been conducted apart from and without much reference to one another. Few researchers of teacher training methods rationalize their selection of teaching skills in terms of TER and very few evaluate the impact of teacher professional development on student learning. At the same time, investigators of teacher effectiveness spend little time speculating about the methods that may be used to improve teaching practice. In this context, the dynamic model of educational effectiveness has been developed in order to establish stronger links between educational effectiveness research and improvement of practice (Creemers & Kyriakides, 2006).

In regard to the use of the model for improvement purposes at teacher level (e.g., initial training and professional development), it is assumed that teaching factors refer to knowledge and skills associated with different types of teacher behaviour in the classroom. These types of behaviour may not necessarily stem from a specific approach to teaching such as direct instruction, active teaching, differentiated teaching or more constructivist approaches to teaching. The dynamic model is not promoting specific approaches but concentrates on instructional behaviour of teachers by describing them through the use of specific teaching factors. Moreover, the model is based on the assumption that teaching factors are not separate entities but some of them are interrelated (Campbell et al., 2003; Creemers, 2007; Johnson, 1997). This implies that teachers may demonstrate types of behaviour that are based on different combinations of the various teaching approaches which can be related to student outcome measures. The model also attempts to describe the complex nature of effectiveness by pointing out not only the importance of specific factors but also by searching for grouping of factors (i.e., types of teacher behaviour). This implies that the model is based on the assumption that improvement of teacher effectiveness can be focused not on the acquisition of isolated skills/competencies (Gilberts & Lignugaris-Kraft, 1997) but on helping teachers exercise

6

and/or develop types of teacher behaviour that are more effective than others. Three recent studies provided support to this assumption and are briefly described below.

1) A Study Searching for Stages - Levels of Effective Teaching

All the grade 5 students (n=2503) from each class (n=108) of 50 primary schools in Cyprus participated in this study. Student achievement in mathematics, Greek language, and religious education were measured, when the students were both at the beginning and end of grade 5. In order to collect data on teacher factors of the dynamic model, 972 observations of the 108 teachers of the student sample were conducted. Two low-inference and one high-inference observational instruments were used. These instruments were designed to collect data concerned with all the eight factors of the teacher level, in relation with the five measurement dimensions which are included in the dynamic model of EER.

By utilizing the Rasch and Saltus models, the teaching skills included in the dynamic model of educational effectiveness were grouped into 5 stages. These were situated in a developmental order and linked with student outcomes. Taking student outcomes as criteria, teachers who demonstrated competencies in the higher stages were found to be more effective than those situated at the lower stages, and thus students of teachers situated at higher stages showed better outcomes. This association is found for achievement in different subjects and for both cognitive and affective outcomes.

The five levels of the model are presented in Table 1. The first three levels are largely related to the direct and active teaching approach, by moving from the basic requirements concerning quantitative characteristics of teaching routines to the more advanced requirements concerning the appropriate use of these skills as these are measured by the qualitative characteristics of these factors. These skills also gradually move from the use of teacher-centered approaches to the active involvement of students in teaching and learning. The last two levels are more demanding since teachers are expected to differentiate their instruction (level 4) and demonstrate their ability to use the new teaching approach (level 5). Considering these five stages and the properties of the Rasch scale which were developed, one can conclude that some stages are more difficult to accomplish than

others. This supports the conclusion that the five stages are not just a grouping of effectiveness factors, but represent equivalent developmental stages of teaching proficiency.

_____

Insert Table 1 Here

_____

The findings of this study are also in line with theories related to the stage models of professional development (e.g., Dreyfus & Dreyfus, 1986; Berliner, 1988, 1992, 1994; Feiman-Nemser & Remillard, 1996; Sternberg et al., 2000). The five stages proposed by Kyriakides, Creemers and Antoniou (2009) advance on previous stage models by specifically determining the content of each stage (in terms of teaching skills), whereas previous stage models often lacked clarity on what might constitute each developmental stage.

2) The Cross-cultural Validity of the Dynamic Model at Teacher Level: A Canadian Study

The main aim of the second study was to further test the validity of the dynamic model at the teacher level, by investigating the extent to which the teaching skills of teachers in Canada could be grouped into the same stages as those reported by the study mentioned above (Kyriakides, Creemers, & Antoniou, 2009). In the first phase of the study the eight teacher factors and their dimensions were measured, by administering a questionnaire to students. Students were asked to indicate the extent to which their teacher behaved in a certain way in their classroom; a Likert scale was used to collect these data. This questionnaire has been used to collect data from Cypriot students of grades 5 and 6, and a Generalisability study (Creemers & Kyriakides, 2008) on the use of students' ratings revealed that data from almost all the questionnaire items could be used for measuring teaching quality. For the development of the French version of the questionnaire, the process of double translation was used and thus both the face and content validity of the instrument were examined. Consequently, 78 items were kept in the final version of the questionnaire.

The sample was taken from seven primary schools in the suburb area of Montreal (Canada), who agreed to participate in the study. All grade 3, 4, 5, and 6 students (n=959) from each class (n=42) of the school sample were asked to complete the questionnaire. The response rate was 73%.

The Generalisability study (G-study) revealed that the data from almost all items (63 out of 65) could be used for measuring the teaching quality of each teacher separately for each subject. It is important to note here that the student questionnaire was administered to far younger students than those participating in the Cyprus study. However, age effects on the results of the G-study were not identified. This implies that, at least in Canada, younger students could also generate reliable data on their teachers' classroom behaviours, in relation to the eight factors of the model and their five dimensions. Since the data were found to be generalisable at the teacher level, the research team calculated a score for each teacher in each of the 63 questionnaire items deemed generalisable. Specifically, for each teacher a score for each item was created by calculating the mean score from the responses of the students of their class. Following this, the Rasch model was applied to the whole sample of teachers and all 62 measures concerning their teaching skills, using the computer program Quest (Adams & Khoo, 1996). Five items did not fit the model. By analyzing the data on the other 58 items, a scale with appropriate psychometric properties was established (see Janosz, Archambault, & Kyriakides, 2011). The results of the various approaches used to test the fit of the Rasch model to the data revealed a good fit when teachers' performance in other teaching skills were analysed. Specifically, all teaching skills were found to have item infit with the range of 0.83 up to 1.20, and item outfit with the range of 0.71 up to 1.42. In addition, all the values of infit for both individuals and teaching skills were greater than -2.00 and smaller than 2.00. The procedure proposed by Yen (1993) was used to test for local independence. It found that local independence was generally not violated. Finally, the fit of the Rasch model to the existing data was also tested against alternative item response theory models, but the improvement of fit by the Two-Parameter Logistic (2PL) over the Rasch model was not statistically significant.

Subsequently, the procedure for detecting pattern clustering in measurement designs, developed by Marcoulides and Drezner (1999), was used to establish whether teaching skills were grouped into levels of difficulty corresponding to easier or more difficult types of teacher behaviours. This method of clustering teaching skills, on the basis of their difficulties from the Rasch model, showed that they are optimally clustered into four types of teacher behaviour (stages of teaching)

which were similar to those identified by the study conducted in Cyprus (Janosz, Archambault, & Kyriakides, 2011).

This study provides some support to the assumption of the dynamic model that teacher level factors are interrelated, and thus should not be treated as isolated. Moreover, the use of specific ways to describe both quantitative and qualitative characteristics of these factors assists in classifying these skills into types of teacher behaviours, which range from relatively easy to the more advanced. The four types of behaviour which emerged from this study are similar to the five levels identified by the study conducted in Cyprus. However, skills associated with the differentiation of teaching were not found to belong to a single level. The results of this study also provide support to the dynamic model's attempt to describe effective teaching using an integrated approach. Specifically, skills associated with both direct teaching and the new teaching approaches were found to belong to the same levels. Moreover, the types of teacher behaviour identified support the idea of combining teaching skills within each type of behaviour, rather than treating each skill or factor in an isolated way.

3) Searching for Stages of Teacher Skills in Assessment

The third study was an attempt to search for grouping of factors by investigating into more details a factor of the model concerned with teacher role in assessing his/her students. Teacher assessment is considered an integral part of teaching (Broadfoot & Black, 2004; Delandshere, 2002; Gipps, 1994; Harlen & James, 1997; Linn, 1993). It is defined as the systematic process of gathering information about student learning (Shepard, 2000). It involves making our expectations explicit and public; setting appropriate criteria and high standards for learning quality; systematically gathering, analyzing, and interpreting evidence to determine how well performance matches those expectations and standards; and using the resulting information to document, explain, and improve performance (Angelo, 1995). Teacher assessment's impact on learning has been widely documented in the literature (Boud, 1995; Butler & Winnie, 1995; Crooks, 1988; Nicol & Macfarlane-Dick, 2006; Sadler, 1989). Consequently, the dynamic model refers to teacher assessment as an important teacher factor. It is also stressed that assessment should serve a formative purpose, and research evidence

shows that teachers and schools which make use of assessment data for formative reasons are more effective (Brookhart, 2004; Delandshere, 2002; Krasne et al., 2006; Kyriakides, 2005). Although the formative purpose of assessment has been widely promoted by the educational community (Gipps, 1994; Stiggins, 1999; Shepard, 2000; Stobart, 2004; Popham, 2006), assessment research literature has failed to impact upon teachers' everyday assessment practices, which still appear to be outcome-oriented (Earl & Katz, 2000; Lock & Munby, 2000). Furthermore, there is little research investigating teachers' assessment skills either for formative or summative purposes (Mok, 2010; Wiliam, Lee, Harrison, & Black, 2004). This project not only attempts to find out whether teacher assessment skills can be grouped into different developmental levels, but also whether teachers mastering higher level skills are more effective than others.

In order to measure teachers' assessment skills, the project addresses the four distinct phases of the assessment cycle (see Figure 1) which show that teachers should ensure that: (a) appropriate assessment instruments are used to collect valid and reliable data; (b) appropriate procedures in administering these instruments are followed; (c) the data emerging from assessment are recorded efficiently and without losing important information, and (d) the results are reported to parents and students, to assist them in decisions on how to provide support to students to increase their learning and learning outcomes.

_____

Insert Figure 1 about here

_____

The dynamic model was also utilised in order to measure teacher assessment skills. Specifically, the five dimensions used to measure the functioning of each classroom factor were taken into account. For example, the stage dimension addresses the period in which assessment is used, and teachers are expected to use assessment at different periods to inform students about their progress and their learning needs, and not solely at the end of the school year or semester. Similarly, the differentiation dimension investigates the extent to which teachers use different approaches for reporting results to parents or students. Finally, the techniques used by teachers are taken into account when measuring their skills. The importance of using various techniques is supported in the literature, and also relates

to the dynamic model's measurement of this teacher factor through the frequency dimension. Table 2 shows the theoretical framework that was used in measuring teacher assessment skills. The research design of the project is described in the next section.

_____

Insert Table 2 about here

_____

By taking into account the theoretical framework and its dimensions, a teacher questionnaire was developed and administered to a representative sample of 10% of primary Cypriot teachers. Of the 240 teachers approached 178 responded, a response rate of 74.2%. The questionnaire was concerned with their skills in assessment of mathematics in primary school. In order to examine the internal validity of the questionnaire data, semi-structured interviews with eight teachers were also conducted. These qualitative data were analysed using the constant comparative method. Comparing the results from each interviewee with their responses to the questionnaire provided support for the internal validity of the study (see Christoforides & Kyriakides, 2011).

The Rasch model was used to analyse data from teachers' responses to the questionnaire; the scales which were created showed satisfactory psychometric properties (see Kyriakides & Christoforides, 2011). It was also found that assessment skills could be grouped into four distinct types of assessment behaviour, which range from skills associated with everyday assessment routines to more advanced skills concerned with differentiation in assessment. The stages are described below:

Stage 1: Basic skills in assessment: Teachers using assessment skills in this stage typically use assessment in order to improve students' learning. They are able to effectively use assessment routines such as:

- ✓ Enrichment or alteration of ready-made written tests
- ✓ Using different types of written questions to assess students' performance
- ✓ Assessing group work based on more than just the overall result
- ✓ Being consistent in checking homework
- ✓ Keeping records for written assessment
- ✓ Reporting assessment results in a summative way.

Stage 2: Using assessment for improvement purposes: The assessment skills included in this stage reveal that teachers using these behaviours are able to use assessment for formative purposes. The skills included in this stage are:

- ✓ Developing a specification table in order to construct written tests
- ✓ Using test items which not only ask for the final product of a task, but also the process used to reach this outcome
- ✓ Using oral assessment and observation
- ✓ Offering clarification comments during the administration of written tests
- ✓ Marking homework for formative reasons
- ✓ Keeping records using descriptive comments
- ✓ Reporting assessment results to parents.

Stage 3: Using assessment techniques to measure the three main learning domains (cognitive, affective and psychomotor) and reporting data covering these domains: Teachers using the assessment skills included in this stage are not only able to measure knowledge in mathematics, but also measure skills and abilities of students working together. Skills included in this stage are:

- ✓ Evaluating skills by developing relevant observation tools
- ✓ Assessing group work
- ✓ Reporting results derived from all assessment techniques to both parents and students
- ✓ Keeping records for the performance of students in each exercise/goal included in the specification table of the assessment instrument.

Stage 4 (Differentiation in Assessment): Teachers demonstrating the assessment skills in this stage are able to differentiate assessment procedures and tools based on their students' needs. This differentiation refers to:

- ✓ Construction and administration of written assessments
- ✓ Construction and administration of oral assessments
- ✓ Reporting to parents and students.

IMPLICATIONS FOR TEACHER PROFESSIONAL DEVELOPMENT

Implications of the findings of these three recent studies for the development of teacher professional development programmes, or specific strategies for improving teacher effectiveness, may emerge by looking at the grouping of teacher factors in the dynamic model. A question raised is the extent to which teachers can move from one stage of teaching competence to the next, by improving their teaching skills and ultimately their student achievement gains. In order to provide answers to this question, two group randomization studies were conducted and we will present shortly the first study. The dynamic approach proposed here lies between the two dominant approaches to teacher professional development (i.e., the CBA and the HA) and aims to overcome their main weaknesses (see Antoniou & Kyriakides, 2011; Creemers & Kyriakides, 2012). Particularly, the dynamic dimension of this approach is attributed to the fact that its content derives from the grouping of teaching skills included in the dynamic model and it is differentiated to meet the needs and priorities of teachers at each developmental stage. Although the content of this approach refers to teaching skills that were found to be positively related with student achievement, the participants are also engaged into systematic and guided critical reflection on their teaching practices. Thus, the effectiveness of this approach was compared with the Holistic (or reflective) approach in improving teaching skills and student outcomes in mathematics. This comparison was made because the Holistic Approach (HA) is largely considered to be the current dominant orthodoxy in teachers' training and professional development (Golby & Viant, 2007).

A total of 130 primary teachers volunteered to participate in the professional development programme. Data were also collected for all students (n=2356) of the teacher-sample. Data were collected both at the beginning and end of the intervention. The four phases of the experimental study are elaborated upon below:

Phase 1: At the beginning of the school year 2008-2009, the teaching skills of the participants were evaluated by external observers. Data on student achievement were collected using external written forms of assessment in mathematics. Observation data were then analysed using the same procedure as described by Kyriakides et al. (2009) in order to classify teachers into developmental stages according to their teaching skills. Using the Rasch and the Saltus models, it was found that

teachers could be classified into the same five developmental stages which emerged from the first study mentioned above (see Table 1).

Phase 2: The teachers at each developmental stage were randomly allocated into two groups of equal size. The first group employed the dynamic approach while the second group used the HA. For example, the 32 teachers at Stage 1 were randomly allocated into the two experimental groups, each one consisting of 16 teachers.

Phase 3: In the third phase of the study, the teachers of each experimental group began to work towards improving their teaching skills. This phase sought to initiate changes in educational practices, working with the teachers throughout the whole curriculum. It was also concerned with whether, and to what extent, teachers can develop their teaching skills and integrate them into a more self-consciously articulated model of classroom pedagogy. The interventions for experimental groups A and B are described below respectively.

i) Sessions for teachers employing the Dynamic Approach

Teachers employing the dynamic approach were assigned to four groups according to the development stage in which they were found to be situated. Supporting literature and research findings related with the teaching skills which correspond to their developmental stage were provided and the area on which each group had to concentrate their efforts for improvement was made clear. Finally, each teacher developed his/her own action plan by exchanging ideas with the research team and the members of his/her group. One session per month was scheduled until the end of the school year. This decision provided the teachers with sufficient time to implement the activities included in their action plans in their teaching and also to reflect on the effectiveness of these activities in order to revise and develop further their action plans. The monthly sessions were organized in groups (based on teachers´ stages) and teachers were strongly encouraged to cooperate and share ideas and teaching materials, to exchange and discuss their experiences and generally to share the results of their exploration. Finally, researchers regularly visited teachers at their schools to discuss emergent issues related with the implementation of their action plans into their everyday teaching and to provide support and feedback to the teachers.

ii) Sessions for teachers employing the HA

The primary aim of these sessions was to enable individuals to critically evaluate their own beliefs and practice and help them transform their experiences from a past event to an ongoing learning process. Thus, teachers had the chance to discuss in groups, identify a problem which they considered important in their teaching and formulate a plan of action to tackle this problem. After the development of the teachers' initial action plans, we scheduled one session per month until the end of the school year. This decision provided the teachers with sufficient time to implement the activities included in their action plans and to reflect on the effectiveness of these activities.

The results of the analysis evidenced the impact of the two approaches to teacher professional development on the improvement of teaching skills and student academic outcomes. A summary of the main findings is presented here whereas further information about the results of the study are reported by Antoniou & Kyriakides (2011).

First, the results of the analysis of both the initial and final data related to teaching skills suggest that the five stages of teaching skills were formulated in a consistent manner. This provides support for the generalisability of the five developmental stages of teaching skills proposed by previous research findings (Kyriakides et al., 2009). In addition, it was found that teachers demonstrating higher level competencies were more effective than those situated at the lower stages, in terms of student outcomes. Secondly, the results indicated that the dynamic approach is more effective than the HA in improving teaching skills, for all teachers. By comparing the two experimental groups it was found that, overall, teachers employing the HA neither made statistically significant progress nor moved from one stage to another. On the other hand, statistically significant progress in teaching skills was found for the teachers employing the approach based on the grouping of teaching skills in the dynamic model. Thirdly, it was found that employing the dynamic approach had a reasonable and statistically significant effect on student achievement, compared with employing the HA.

The above findings support that teachers can improve and ultimately progress to the next developmental stage of teaching skills, by undertaking appropriate interventions and participating in effective professional development programmes. As this study demonstrated, teachers employing the

dynamic approach improved their teaching skills, whereas those employing the HA did not. In addition, the use of the dynamic approach had a significant impact upon student achievement gains in mathematics. A similar argument was made by King and Kitchener (1994). They argued that stage growth was most apparent for teachers who continued their informal education and participated in effective professional development programmes. This is an important reminder that teacher improvement and stage growth do not unilaterally unfold, but also require a stimulating and supportive environment.

The issue concerning the content of teacher professional development programmes has been addressed in this study, by drawing from a validated theoretical model of EER. In particular the dynamic model of educational effectiveness emphasises not only the importance of specific factors, but also the grouping of factors, when addressing the complex nature of effectiveness. This implies that improvement of teacher effectiveness cannot be focused solely on the acquisition of isolated skills or competencies (Gilberts & Lignugaris-Kraft, 1997), nor on reflection across the whole teaching process to help teachers get "greater fulfillment as a practitioner of the art" (of teaching) (Clarke & Hollingsworth, 2002, p. 948).

At the same time, the results of this study indicate that reflection is more effective when teachers' priorities for improvement are taken into account, and when they are encouraged to develop action plans which address their professional needs; these were identified through a relevant empirical investigation. Although both interventions encouraged and utilised teachers' critical reflections of their teaching practices, teachers employing the dynamic approach were asked to reflect on those aspects which related to their priorities for improvement based on their developmental stage. These stages were defined by taking into account the knowledge-base of EER, especially teacher factors found to be associated with student achievement. On the other hand, teachers employing the HA adopted a less focused reflection strategy, which allowed teachers to reflect on any aspect of their teaching practice irrespective of the stage on which they were situated. For example, some teachers at stage 1 employing the HA developed action plans aiming to differentiate their instruction; yet their attempts to incorporate this into their teaching were not successful. This may be attributed to the fact that they did not possess basic skills corresponding to their stage, such as classroom management and

structuring, which could be considered pre-requisites for the differentiation of teaching. Therefore the HA does not take into account research evidence supporting the grouping of teacher factors and their dimensions, grouped into stages, structured in a developmental order and associated with student outcomes. It must be emphasised that the importance of thinking and critical analysis are important, and thus those aspects of the HA were utilised in the development of the dynamic approach. However, complimenting reflection with the knowledge-base of EER, which addresses the needs of specific groups of teachers, could help us to establish more effective approaches to teacher professional development which will have an impact on improving learning outcomes.

References

Adams, R.J., & Khoo, S. (1996). *Quest: The interactive test analysis system, Version 2.1*. Melbourne: ACER.

Angelo, T.A. (1995). Reassessing and defining assessment. AAHE Bulletin (Nov.), 7–9.

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.

Askew, M., & William, D. (1995). *Recent Research in Mathematics Education 5-16.* London: Office for Standards in Education, 53.

Bransford, J., Brown, A., & Cocking, R. (2000). *How People Learn: Brain, Mind, and Experience & School.* Washington, DC: National Academy Press.

Berliner, D. (1988). *The Development of Expertise in Pedagogy*. Charles W. Hunt Memorial Lecture for the American Association of Colleges in Teacher Education, New Orleans, LA.

Berliner, D. (1992). Expertise in teaching. In F. Oser, J. -L. Patry, & A. Dick (Eds.), *Effective and responsible teaching,* (pp. 227–249). San Francisco: Jossey-Bass.

Berliner, D. (1994). Expertise: The wonder of exemplary performances. In J. Mangieri & C. Block (Eds.), *Creating powerful thinking in teachers and students: Diverse perspectives,* (pp. 161–186). Fort Worth, TX: Harcourt Brace College.

Boud, D. (1995). Assessment and learning: contradictory or complementary? In P. Knight (Ed.) *Assessment for Learning in Higher Education*. London: Kogan Page.

Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of Assessment *in Education*. *Assessment in Education, 11*(1), 7–27.

Brookhart, S.M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, *106*(3), 429–458.

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

Brush, T.A. (1997). The effects on student achievement and attitudes when using integrated learning systems with cooperative pairs. *Etr & D-Educational Technology Research and Development, 45*(1), 51-64.

Butler, D.L., & Winne, P.H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*, 245–281.

Campbell, R.J., Kyriakides, L., Muijs, R.D., & Robinson, W. (2003). Differential Teacher Effectiveness: Towards A Model For Research And Teacher Appraisal. *Oxford Review of Education, 29*(3), 347–362.

Cazden, C.B. (1986). Classroom Discourse. In M. C. Wittrock (Ed.) *Handbook of Research on Teaching* (pp. 432-463). New York: MacMillan.

Christoforides, M., & Kyriakides, L. (2011, January). *Using the dynamic model to identify stages of teachers' skills in assessment.* Paper presented at the 24th International Congress for School Effectiveness and Improvement (ICSEI) 2011. Limassol, Cyprus.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*(8), 947–967.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing and mathematics. In L.B. Resnick (Ed.), *Knowing learning and instruction* (pp. 453–495). Hillsdale, MI: Lawrence Erlbaum.

Creemers, B. (1994). *The Effective Classroom.* London: Cassell.

Creemers, B.P.M. (2007). Educational effectiveness and improvement: the development of the field in mainland Europe. In Townsend, T. (ed) *International Handbook of School Effectiveness and Improvement* (pp. 223-242). New York, Springer.

Creemers, B.P.M., & Kyriakides, L. (2006). Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement, 17*(3), 347–366.

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. (2012). *Improving Quality in Education: Dynamic Approaches to School Improvement*. London: Routledge.

Creemers, B.P.M., & Reezigt, G.J. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement*, *7*(3), 197–228.

Crooks, T.J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, *58*, 438–481.

Delandshere, G. (2002). Assessment as Inquiry. *Teachers College Record*, *104*(7), 1461– 1484.

Den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School Effectiveness and School Improvement*, *15*(3-4), 407–442.

Doyle, W. (1986). Classroom organization and management. In M.C. Wittrock (Ed.), *Handbook of Research on Teaching, Third Edition* (pp. 392–431). New York: Macmillan.

Dreyfus, H.L., & Dreyfus, S.E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer.* New York: Free Press.

Earl, L., & Katz, S. (2000). Changing classroom assessment: Teachers' struggles. In N. Bascia & A. Hargreaves (Eds.), *The sharp edge of educational change* (pp. 97–111). London: Routledge.

Feiman-Nemser, S., & Remillard, J. (1996). Perspectives on learning to teach. In F.B. Murray (Ed.), *The teacher educator's handbook*, (pp. 63–91). San Francisco: Jossey-Bass.

Fraser, B.J. (1991). Two Decades of Classroom Environment Research. In B.J. Fraser & H.J. Walberg (Eds.) *Educational Environments: Evaluation, Antecedents and Consequences,* (pp. 3-29). Oxford: Pergamon.

Fuchs, L.S., Fuchs, D., Yazdian, L., & Powell, S.R. (2002). Enhancing first-grade children's mathematical development with peer-assisted learning strategies. *School Psychology Review, 31*(4), 569-583.

Galton, M. (1987). An ORACLE chronicle: A decade of classroom research. *Teaching and Teacher Education, 3*(4), 299-313.

Gilberts, G.H., & Lignugaris-Kraft, B. (1997). Classroom management and instruction competencies for preparing elementary and special education teachers. *Teaching and Teacher Education, 13*(6), 597–610.

Gipps, C. (1994). *Beyond Testing*. RoutledgeFalmer, London.

Golby, M., & Viant, R., (2007). Means and ends in professional development. *Teacher Development*, *11*(2), 237–243.

Good, T.L., Grouws, D.A., & Ebmeier, H (1983). *Active Mathematics Teaching*. New York: Longman.

Harlen, W., & James, M. (1997). Assessment and Learning: Differences and relationships between formative and summative assessment. *Assessment in Education*, *4*(3), 365-379.

Heck, R.A., & Marcoulides, G.A. (1996). School culture and performance: Testing the invariance of an organizational model. *School Effectiveness and School Improvement, 7*(1), 76–106.

Hoy, W.K., Tater, J.C., & Bliss, J.R. (1990). Organizational climate, school health, and effectiveness: A comparative analysis. *Educational Administration Quarterly*, *26*(3), 260-279.

Janosz, M., Archambault, I., & Kyriakides, L. (2011, January). *The cross-cultural validity of the dynamic model of educational effectiveness: A Canadian study.* Paper presented at the 24th International Congress for School Effectiveness and Improvement (ICSEI) 2011. Limassol, Cyprus.

Johnson, B. (1997). An organizational analysis of multiple perspectives of effective teaching: Implications of teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 69–87.

King, P.M., & Kitchener, K.S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults.* San Francisco: Jossey-Bass.

Kirschner, P.A., Sweller, J., & Clark, R.E. (2006). Why minimal guidance does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Kosir, K. (2005). The influence of teacher's classroom management style on pupils' self regulative behaviour. *Studia Psychologica*, *47*(2), 119-143.

Krasne, S., Wimmers, P.F., Relan, A., & Drake, T.A. (2006). Differential effects of two types of formative assessment in predicting performance of first-year medical students. *Advances in Health Sciences Education*, *11*(2), 155–171.

Kyriakides, L. (2005). Extending the Comprehensive Model of Educational Effectiveness by an Empirical Investigation. *School Effectiveness and School Improvement*, *16*(2), 103–152.

Kyriakides, L. (2008). Testing the validity of the comprehensive model of educational effectiveness: a step towards the development of a dynamic model of effectiveness. *School Effectiveness and School Improvement, 19*(4), 429–446.

Kyriakides, L., & Christoforides, M. (2011, October). *Searching for stages of teacher skills in assessment: Implications for research on teacher professional development.* Paper presented at the 37th International Association for Educational Assessment Annual Conference (IAEA) 2011. Manila, Philippines.

Kyriakides, L., & Christoforou, Ch. (2011, April). *A Synthesis of Studies Searching for Teacher Factors: Implications for Educational Effectiveness Theory.* Paper presented at the American Educational Research Association (AERA) 2011 Conference. New Orleans.

Kyriakides, L., & Creemers, B.P.M. (2006). Using the dynamic model of educational effectiveness to introduce a policy promoting the provision of equal opportunities to students of different social groups. In McInerney, D.M., Van Etten, S. & Dowson, M. (Eds.) *Research on Sociocultural Influences on Motivation and learning, Vol. 6: Effective schooling.* Information Age Publishing, Greenwich CT.

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

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

Linn, R.L. (1993). Educational assessment: Expanded expectations and challenges. *Educational Evaluation and Policy Analysis*, *15,* 1–16.

Lock, C.L., & Munby, H. (2000). Changing assessment practices in the classroom: A study of one teacher's change. *The Alberta Journal of Educational Research*, *46,* 267–279.

Marcoulides, G.A., & Drezner, Z. (1999). A procedure for detecting pattern clustering in measurement designs. In M. Wilson, & G. Engelhard, Jr. (Eds.), *Objective measurement: Theory into practice* (Vol. 5, pp. 261-277)*.* Ablex Publishing Corporation.

Mok, M.M.C. (2010). *Self-directed Learning Oriented Assessment: Assessment that Informs Learning & Empowers the Learner*. Hong Kong: Pace Publications Ltd.

Muijs, D., & Reynolds, D. (2000). School Effectiveness and Teacher Effectiveness in Mathematics: Some preliminary Findings from the Evaluation of the Mathematics Enhancement Programme (Primary). *School Effectiveness and School Improvement, 11*(3), 273–303.

Muijs, D., & Reynolds, D. (2001). *Effective Teaching: evidence and practice*. London: Sage.

Nicol, D.J., & Macfarlane-Dick, D. (2006). Formative Assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218.

Nolen, S.B. (2003). Learning environment, motivation, and achievement in high school science. *Journal of Research in Science Teaching, 40*(4), 347-368.

Popham, W.J. (2006). Phony formative assessments: Buyer beware! *Educational Leadership, 64*(3), 86–87.

Ramsden, J.M. (1997). How does a context –based approach influence understanding of key chemical ideas at 16+? *International Journal of Science Education, 19*(6), 697-710.

Rohrbeck, C.A., Ginsburg-Block, M.D., Fantuzzo, J.W., & Miller, T.R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95*(2), 240-257.

Rosenshine, B., & Stevens, R. (1986). Teaching Functions. In M.C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed., pp. 376–391). New York: Macmillan.

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144.

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

Seidel, T., & Shavelson, R.J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*(4), 454-499.

Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational researcher*, *29*(7), 4–14.

Simons, R.J., van der Linden, J., & Duffy, T. (2000). New Learning: Three ways to Learn in a New Balance. In R.J. Simons, J. van der Linden, & T. Duffy (Eds) *New Learning* (pp. 1-20). Dordrecht, The Netherlands: Kluwer.

Slavin, R.E. (1983). When does cooperative learning increase student-achievement? *Psychological Bulletin, 94*(3), 429–445.

Slavin R.E., & Cooper, R. (1999). Improving intergroup relations: Lessons learned from cooperative learning programs. *Journal Of Social Issues*, *55*(4), 647–663.

Smith, L., & Land, M. (1981). Low-inference verbal behaviors related to teacher clarity. *Journal of Classroom Interaction, 17*, 37-42.

Stallings, J. (1985). Effective elementary classroom practices. In M.J. Kyle (Ed.), *Reaching for excellence: An effective sourcebook* (pp. 14-42). Washington, DC: US Governing Printing Office.

Steffe, L., & Gale, J. (Eds.). (1995). *Constructivism in education*. Hillsdale, N.J: Lawrence Erlbaum Associates, Inc.

Sternberg, R.J., Forsythe, G.B., Hedlund, J., Hovath, A.J., Wagner, R.K., Williams, W.M., Snook, S.A., & Grigorenko, E.L. (2000). *Practical intelligence in everyday life.* New York: Cambridge University Press.

Stiggins, R.J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, *18*(1), 23–27.

Stobart, G. (2004, June). *The formative use of summative assessment: possibilities and limits.* Paper presented at the 30th Annual IAEA Conference, Philadelphia.

Walberg, H.J. (1986). What works in a nation still at risk. *Educational Leadership*, *44*(1), 7–10.

Wiliam, D., Lee, C., Harrison, C., & Black, P.J. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice, 11*(1), 49–65.

Yen, W. (1993). Scaling and performance assessments: strategies for managing local item dependence. *Journal of Educational Measurement, 30,* 187–213.

Table 1. The five stages of teaching skills included in the Dynamic Model

Stages	Teaching Skills
1. Basic elements of direct teaching	• Frequency management time • Stage management of time • Frequency structuring • Frequency application • Frequency assessment • Frequency questioning • Frequency teacher-student relation
2. Putting aspects of quality in direct teaching and touching on active teaching	• Stage structuring • Quality application • Stage questioning • Frequency student relations • Focus application • Stage application • Quality of questions
3. Acquiring quality in active / direct teaching	• Stage student relations • Stage teacher-student relation • Stage assessment • Frequency teaching modelling • Frequency orientation • Focus student relations • Quality: feedback • Focus questioning • Focus teacher-student relation • Quality structuring • Quality assessment
4. Differentiation of teaching	• Differentiation structuring • Differentiation time management • Differentiation questioning • Differentiation application • Focus assessment • Differentiation assessment • Stage teaching modelling • Stage orientation
5. Achieving quality and differentiation in teaching using different approaches	• Quality teacher-student relation • Quality student relations • Differentiation teacher-student relation • Differentiation student relations • Focus orientation • Quality orientation • Differentiation orientation • Quality of teaching modelling • Focus teaching modelling

Table 2. The theoretical framework for measuring teacher assessment skills

Assessment phases	Assessment techniques	Measuring dimensions of the dynamic model
1) Planning/construction of tools	1) Written assessment	1) Frequency
2) Assessment administration	2) Oral assessment	2) Focus
3) Recording of assessment information	3) Observation	3) Stage
4) Reporting	4) Performance assessment	4) Quality
		5) Differentiation

Planning/construction of tools
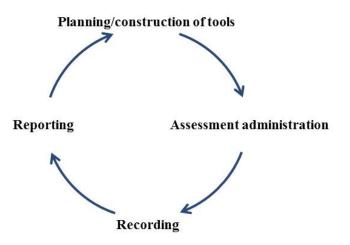
Reporting

Assessment administration

Recording

Figure 1: The assessment cycle illustrating the four phases of assessment