Creemers, Bert P. M.

Educational effectiveness and improvement in developing societies. Some experiences from the Primary Education Quality Improvement Project in Indonesia.

Tertium comparationis 5 (1999) 1, S. 32-51

urn:nbn:de:0111-opus-28908

Educational Effectiveness and Improvement in Developing Societies: Some Experiences from the Primary Education Quality Improvement Project in Indonesia

Bert P.M. Creemers

University of Groningen, the Netherlands

Abstract

The improvement of education in developing societies might benefit from theory and research on educational effectiveness. They provide a useful instrument to various interventions and improvement projects but also a knowledge base that can be used for the design of projects. The research evidence points at the importance of factors at the classroom level and the relatively small possibilities that the school and the above school level have to influence those factors at the classroom level. This is illustrated by the evaluation of the primary education quality improvement project in Indonesia, a project that aimed at the improvement of education through teacher professional development, provision of textbooks, community participation and management of schools. The results tend to support the general feeling about educational effectiveness. Conclusions stress the importance of the development of knowledge by (inter)national consultants, the content of the intervention – educational effectiveness and improvement and the adaptation of the knowledge to national and local circumstances – and procedural and technical knowledge how to design, implement and evaluate educational interventions.

1 Introduction

Unlike other studies, the present study does not derive from international comparative research nor from consultancy work in developing countries. The origins of this study are theory and research on educational effectiveness from an international perspective. For more than a decade, the Groningen Institute for Educational Research and myself are involved in theory and research on educational effectiveness. During that period research was carried out in the Netherlands taking as point of departure a more and more elaborated framework about educational effectiveness and using statistical procedures to analyse complex data sets. Although effective school research finds its origins in research in the United States and the United Kingdom, it is rooted more or less in research on teacher behaviour, quality of instruction, curriculum research, grouping procedures, and so on. In this way, it was possible for the Groningen Institute to combine different trends of research into the field of educational effectiveness.

With the start of the International Congress for School Effectiveness and Improvement (ICSEI) and the companion journal School Effectiveness and School Improvement, the theory and research on educational effectiveness turned more and more international.

That was important, because some of the factors that were significant in specific countries appeared to be of no importance in other countries. This pointed at the importance of contextual factors. Furthermore, international comparative research on student outcomes showed the need to search for factors within classrooms and schools and other levels of the educational system, that may cause differences between student outcomes within and between countries.

In our research in The Netherlands, it also became obvious that educational effectiveness research and theory provide useful instruments for evaluation research. We use designs based on educational effectiveness approach for the evaluation of different school reforms in The Netherlands, similar to studies done in the USA and the UK (Werf 1995). Based on intellectual curiosity (into characteristics of effective schools across countries) and our experiences in basic projects in the country, it was a challenge to take part in the evaluation of the *Primary Education Quality Improvement Project (PEQIP)* in Indonesia, using our knowledge base and methodology in educational effectiveness research. In addition to the research questions related to the PEQIP evaluation, i.e. the assessment of the impact of PEQIP in Indonesia, the Groningen Institute for Educational Research formulated the following questions: Do factors related to effectiveness travel across countries and is it possible to use more or less the same research methodology for evaluation of this kind of projects in developing societies?

In the following sections the design of the evaluation study and the results will be described shortly. In addition, we will address whether educational effectiveness and improvement can work in developing societies, and which lessons can be drawn from our experience in this project.

2 The History of Educational Effectiveness Research and Improvement

The history of educational effectiveness research can be described in different ways. A favourite way is to look at educational effectiveness as a reaction to the quite pessimistic views on teachers, schools and education in general, brought forward by the disappointing results of research. Another, quite different interpretation of the history of educational effectiveness research considers this research as a natural prolongation of research from the past with respect to teaching, instruction, curriculum, school organisation, and so on. Depending on one's views of history, different godfathers for educational effectiveness research are named, like Coleman et al. (1966). For most educational effectiveness research the work of Edmonds (1979) and Brookover, Beady, Flood and Schweitzer (1979) in the United States, and of Rutter, Maughan, Mortimore and Ouston (1979) in the United Kingdom are important starting points for educational effectiveness research.

Those studies revealed that schools differ from each other with respect to the outcomes they generate for comparable groups of students. This holds for academic outcomes as well as other outcomes (attitudes towards school, motivation, juvenile delinquency, etc.). The crucial question is what kind of characteristics distinguish effective schools from less effective ones.

Table 1: Effective Schools' Characteristics Identified in Two Recent Reviews

Levine and Lezotte (1990) Characteristics	Sammons, Hillman and Mortimore (1995) Factors
1. Outstanding *Leadership*	1. Professional *Leadership*
a. Superior Instructional Leadership	a. The Leading Professional
b. Support for Teachers	b. A Participative Approach
c. High Expenditure of Time and Energy for School Improvement	c. Firm and Purposeful
d. Vigorous Selection and Replacement of Teachers	

e. 'Maverick' Orientation and Buffering
f. Frequent, Personal Monitoring of School Activities and Sense-making
g. Acquisition of Resources
h. Availability and Effective Utilisation of Instructional Support Personnel

2. Effective Instructional Arrangements of Implementation
   a. Effective Teaching
   b. Successful Grouping and Related Organisational Arrangements
   c. Classroom Adaptation
   d. Active/enriched learning
   e. Emphasis on HOTs in Assessing Instructional Outcomes
   f. Co-ordination in Curriculum and Instruction
   g. Easy Availability of ... Instructional Materials
   h. Stealing Time for Reading, Language, Maths ...

2. *Purposeful Teaching*
   a. Efficient Organisation
   b. Structured Lessons
   c. Adaptive Practice
   d. Clarity of practice

3. *Focus on* Student Acquisition of Central *Learning* Skills
   a. Maximum Availability and Use of Time for Learning
   b. Emphasis on Mastery of Central Learning Skills

3. *Concentration* on Teaching and *Learning*
   a. Maximisation of Learning Time
   b. Academic Emphasis
   c. Focus on Achievement

4. *Productive* School Climate and *Culture*
   a. *Orderly Environment*
   b. Faculty Commitment to a *Shared* and Articulated *Mission* Focused on Achievement
   c. Faculty Cohesion ... and Collegiality
   d. Schoolwide Emphasis on *Recognising Positive Performance*
   e. Problem Solving Orientation
   f. Faculty Input Into Decision

4. Learning Environment
   a. An *Orderly Environment*
   b. An Attractive Working Environment

5. *High* Operationalised *Expectations* and *Requirements for Students*

5. *Shared Vision* and Goals
   a. Unity of Purpose
   b. Consistency of Practice
   c. Collegiality and Co-operation

6. Appropriate *Monitoring* of Student *Progress*

6. *Positive Reinforcement*
   a. Clear and Fair Discipline
   b. Feedback

7. Practice Oriented *Staff Development* at the *School* Site

7. *High Expectations*
   a. High Expectations All Around
   b. Communication Expectations
   c. Providing Intellectual Change

8. Salient *Parental Involvement*

8. *Pupil* Rights and *Expectations*
   a. Raising Pupil Self-Esteem
   b. Positions of Responsibility
   c. Control of Work

9. Others
   a. Student Sense of Efficacy/Futility
   b. Multicultural Instruction and Sensitivity
   c. Personal Development of Students

9. *Monitoring Progress*
   a. Monitoring Pupil Progress
   b. Evaluating School Performance

10. A Learning Organisation (*School Based Staff Development*)

11. Home-school Partnership (*Parental Involvement*)

Five characteristics are repeatedly mentioned in the literature as correlates of educational achievement. These are:
- educational leadership;
- high expectations of student achievement;
- emphasis on basic skills;
- safe and orderly climate;
- frequent evaluation of pupils' progress (see Edmonds 1979).

The status of the five factor-model is correlational. This means that, although the five

factors are usually seen as causes of student achievement, in the strict methodological sense no such statements concerning causal ordering are warranted (see Ralph & Fennessey 1983; Scheerens & Creemers 1989). This so called five factor-model theory was further elaborated in research resulting in more and more characteristics for educational effectiveness (see for a review, Levine & Lezotte 1990). Each of the major characteristics in this review also contains some subcharacteristics, that more or less define the main characteristic. Based on their overview of research Sammons, Hillman and Mortimore (1995) distinguished 11 key characteristics of effective schools, with some further indications for each of the factors. Table 1 provides an overview of the two studies. There is a remarkable overlap between the two, except that in the Sammons, Hillman and Mortimore overview, more attention is placed on vision and goals and pupils' rights and expectations. The other characteristics are more or less identical although subcharacteristics defining the characteristics differ between the two sets. The nine characteristics both reviews have in common are reformulated by Reynolds and Teddlie (forthcoming) as processes of effective schools and as components of the process and further described there (see Table 2).

Table 2: The Processes of Effective Schools

Process	Components of the Process
1. The Processes of Effective Leadership	a. Being Firm and Purposeful b. Involving Others in the Process c. Exhibiting Instructional Leadership d. Frequent, Personal Monitoring e. Selecting and Replacing Staff
2. The Processes of Effective Teaching	a. Maximising Classtime b. Successful Grouping and Organisation c. Exhibiting Best Teaching Practices d. Adapting Practice to Particulars of Classroom
3. Developing and Maintaining a Pervasive Focus on Learning	a. Focusing on Academics b. Maximising School Learning Time
4. Producing a Positive School Culture	a. Creating a Shared Vision b. Creating an Orderly Environment c. Emphasising Positive Reinforcement
5. Creating High (and Appropriate) Expectations for All	a. For Students b. For Staff
6. Emphasising Student Responsibilities and Rights	a. Responsibilities b. Rights
7. Monitoring Progress at All Levels	a. At the School Level b. At the Classroom Level c. At the Student Level
8. Developing Staff Skills at the School Site	a. Site Based b. Integrated with Ongoing Professional Development
9. Involving Parents in Productive and Appropriate Ways	a. Buffering Negative Influences b. Encouraging Productive Interactions with Parents

Next to a search for characteristics of effective education, interest developed during the last decade for models and theories on effective education. A major breakthrough in this respect was the concept of multiple levels in the system, for example, (1) the individual student level, (2) the classroom/group level, (3) the department level (especially in secondary schools), (4) the school level, (5) the district level, and maybe other levels in the educational system.

The second major component of those theories was the idea of multiple factors at each level, explaining variance within levels and between levels. Furthermore, the importance of the context of education was stressed, such as the culture of a society, the educational policy, and so on. Examples of such theories/models are Scheerens (1992), Creemers and Scheerens (1994), Creemers (1994), Scheerens and Bosker (1997), Slavin (1987, 1996), Stringfield (1994), Reynolds et al. (1994), Gray and Wilcox (1995), Hargreaves (1995). All these models and the more or less explicit theories included depend heavily on theories about learning and school organisation, for example, Carroll's theory about learning and schools, organisational theories from Minz berg and others (see for example, Scheerens 1998). These models are more related to theories developed with respect to learning in schools and school organisation. However, they include also the characteristics of effective schools described above, in order to develop a comprehensive theory about educational effectiveness. An example of such a theory is the comprehensive model for educational effectiveness, outlined in Figure 1.

Figure 1: Basic Model of Educational Effectiveness (Creemers 1994)
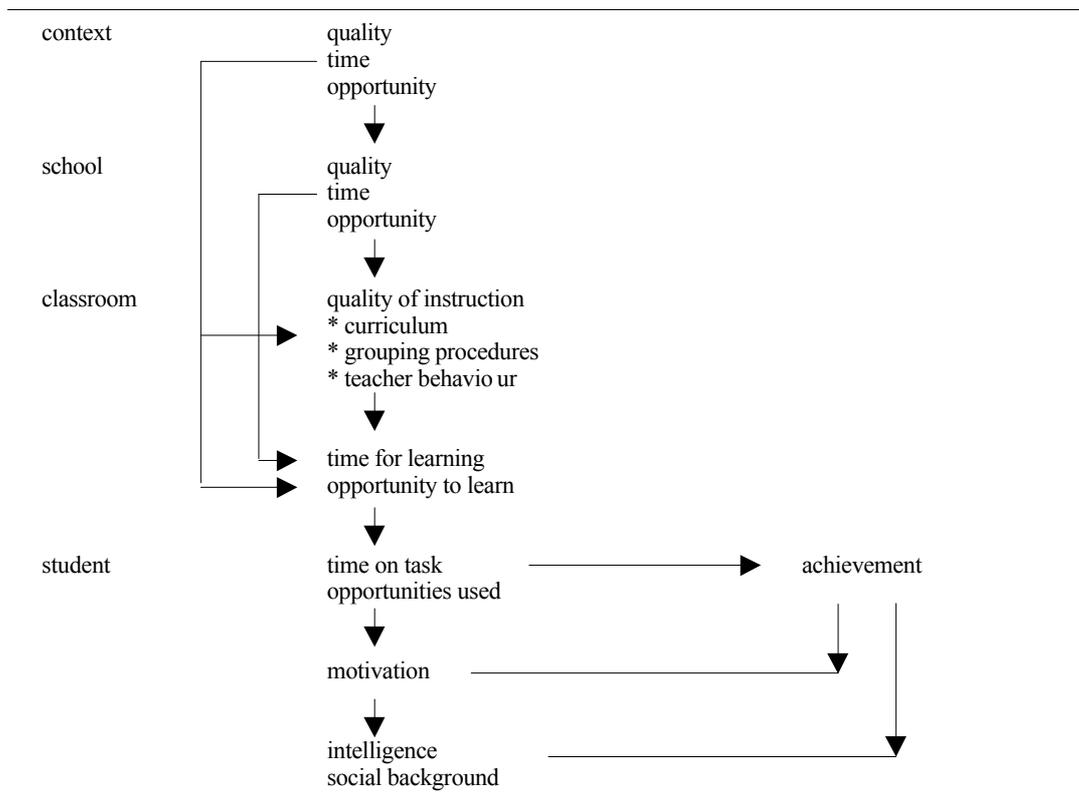
The model shows how the levels in the model influence student outcomes. *Time on task* and *opportunities used* (at student level) are influenced by *time for learning* and *opportunity to learn* provided by the teacher (at classroom level), and these are influenced by the *quality of instruction*. The more adequate the instruction, the more time students can spend on learning and the more opportunities to learn they will have. For example, more school subjects or topics within subjects will be offered. But even when teachers achieve very high levels of time for learning and opportunity to learn, ultimately the students decide how much time they will actually spend on their school work and how much opportunities to learn they will actually use. Therefore, these concepts are directly related to outcomes. In addition, outcomes are determined by student factors such as aptitudes, social background, and motivation. Quality, time and opportunity at the classroom level are influenced by factors at the school level and the school level is influenced by factors at the context level. Outcomes therefore cannot be seen as either an accomplishment of classroom factors only (as in many studies on teacher behaviour) or of school factors only (as in many studies of school policies). The influences of the context and school levels are indirect and mediated by the classroom level.

At the student level, background, motivation and aptitude strongly determine outcomes. Time on task is determined not only by motivation, but also by the time provided by the school and by processes at the school and classroom levels. Time on task has to be filled by opportunities to learn. These refer to the supply of learning material, experiences, and exercises. The opportunities offered in the instructional process and students' actual use of the offered experiences are not necessarily equivalent.

At the classroom level, teacher behaviour is central. Teachers make use of curricular materials and they carry out grouping procedures in their classrooms. However, teachers need curriculum material, which should be consistent with grouping procedures such as mastery learning or co-operative learning. Examples of such factors with respect to the curriculum are explicitness and ordering of goals and content and structure and clarity of content. Examples of teacher behaviour are management, clarity of presentation, evaluation, feedback and corrective instruction.

At the school level, factors that deal directly with activities at classroom level (such as rules and agreements about aspects of classroom instruction) are integrated in the model so are organisational factors which determine what happens in classrooms from a greater distance (such as a school policy professionalising and supervising teachers to comply with the school standards, or a school culture inducing and supporting effectiveness). Also conditions for time (such as a school policy on homework and student absenteeism) and for opportunity to learn (such as rules and agreements about how to implement the curriculum) are specified.

At the context level, quality, time and opportunity to learn are specified by factors such as a (national) policy that focuses on the effectiveness of education, national guidelines with respect to time schedules, and guidelines with respect to a core curriculum or a national curriculum.

Although focusing on the effects of factors, the elaborated model of Creemers (1994) makes tentative statements about their joint impact on student outcomes by introducing the formal criteria of consistency, cohesion, constancy and control. These formal criteria are considered as a major improvement over other models, because these criteria connect the other factors in the model and can explain the joint impact of factors which constitute learning environments. The formal criteria concern the relationships between the factors of instruction (textbook, grouping procedures and teaching behaviour), the stability of factors over time and mechanisms to 'impose' the factors. The idea behind the formal

criteria is that the influence of the factors at a certain level and between factors at different levels can be enforced to occur by the fact that these factors are pursued for a longer period of time and are co-ordinated with each other. According to the *consistency* principle the effectiveness of classrooms, schools and contexts is enhanced when the factors at these levels are consistent with each other and supporting each other. For example, a mathematics curriculum offering tests and corrective instruction might promote evaluative behaviour of the teacher much more strongly than a curriculum not offering these facilities. At the school level, consistency is an important condition for instruction. When all members of a school team take care of consistency, *cohesion* is created which means that every team member is aware of the need for consistency and acts according to schoolwide agreements in this respect. To stabilise outcomes, schools should not change rules and policies every other year. This implies the *constancy* principle, which can be demonstrated in longitudinal research by comparing school level factors from year to year. The *control* principle not only refers to the fact that student outcomes should be evaluated, but also to the quiet atmosphere in the school. Control also refers to teachers holding themselves and others responsible for effectiveness. At the context level, consistency, constancy and control are again important formal characteristics emphasising the importance of factors over time and of mechanisms to ensure effectiveness.

The empirical research with respect to educational effectiveness carries on providing information about the size, the stability of effects, characteristics of effective schools and the use that can be made of the knowledge base to improve education. The knowledge base with respect to educational effectiveness is also used for the design of evaluation studies, for example, in the UK and in The Netherlands for the evaluation of Educational Priority Program (Werf 1995).

3 Educational Effectiveness and Improvement in Developing Societies

A vast amount research is already available concerning components in education, like teacher behaviour, text books, grouping procedures, school organisation, though the origin of effective school research, later called effective education research, is strongly connected with two major studies. One in the United States by Brookover et al. (1979) and one in the United Kingdom by Rutter et al. (1979). Initially, in the United States research addressed the characteristics of effective schools and the factors that can induce effective education (see reviews of Levine & Lezotte 1990; Stringfield 1994; Reynolds et al. 1994).

Due to the fact that there was more attention for the effects in terms of student outcomes in educational policy and practice, effective school research received more and more attention in other countries. Originally, the majority of studies addressed the same issues as in USA and UK research: to find out which kind of factors where important for effective education, and to verify the findings from studies in other countries. Later on, as explained in section 2, more attention was placed on the construction of models and on theories of effective education and on finding empirical support for those models. Results were also used to improve education and to design the evaluation of ongoing school improvement projects (see for overviews: Reynolds et al. 1994; Scheerens & Bosker 1997).

A number of studies were done also in developing countries dealing with the factors that contribute to the outcomes of education in different countries. Some of those studies sought to investigate the relationships between educational factors and educational out-

comes. The majority of the studies were in one way or another related to ongoing school improvement projects. Reviews of research literature are already available (Fuller & Heyneman 1989; Riddell 1989). Until recently, there were just a few number of studies that utilise multilevel methodologies (Riddell 1989, in Zimbabwe; Lockheed & Longford 1991, in Thailand), but this number is increasing (Bashir 1994; Singh et al. 1995; Raudenbush, Kidchanapanish & Kang 1991). In the single level studies the proportion of the block of school factors explained major proportions of variance in achievement (Heyneman & Loxley 1983). On this basis Heyneman and Loxley concluded that different paradigms of school effectiveness exist in developing countries where schools have a greater impact on achievement than home background factors. The same was found by Fuller (1987), but in fact the studies on which those conclusions were based on suffer from severe methodological and statistical weaknesses, which may invalidate their results (Fuller 1987; Riddell 1989). For example, Lockheed, Fuller and Nyirongo (1989), carried out a study of family effects on achievement in Thailand and Malawi, which aimed at using more culturally relevant family background measures and also included motivational variables. They found that in Thailand family background and prior achievement effected pupils' educational experiences, perceptions of ability and effort, which in turn influence later performance. In Malawi, they found, after using twenty specific measures of family background, that the family background had a more significant influence on achievement than global Westerns proxies (however, the Malawi-studies should also be treated cautiously, because they were cross-sectional and included only 103 pupils of 21 schools). In multilevel-studies high proportions of variance were found due to differences between schools (Raudenbush et al. 1991).

For Indonesia, Kaluge (1998) found that the proportion of total variance in pupil attainment for language at the school level ranged from 14% to nearly 23%. In mathematics, school influence was even larger, ranging from 20% to 29% of the variance was explained by schools.

The majority of educational effectiveness studies took place within countries and variance between schools and classrooms can be partly explained by characteristics of factors that are related to educational outcomes. Such research increasingly makes use of an international body of knowledge and the same methodology (cohort studies, multilevel analysis, etc.) At present, more comparative international work becomes available. This may take different forms, e.g. studies that assess and compare achievement of students and relate them to some background and school factors, like the studies carried out by the International Association for the Evaluation of Educational Achievement (IEA), and the studies by the Educational Testing Service (USA) entitled the International Assessment of Educational Progress (IAEP). IEA conducted the first and second science studies, the first and second mathematics studies, the classroom environment study, the study of written composition and the first literacy studies, and recently, the third international mathematics and science study (TIMSS). IEAP has conducted the first study of science and mathematics achievement cross-culturally.

A final example of international comparative research in the field of educational effectiveness is the International School Effectiveness Research Project (ISERP) (Creemers & Werf forthcoming; Reynolds et al. 1994). These studies are all important for several reasons:

1. They add to t he development of the knowledge base on educational effectiveness; with respect to the conceptualisation, the operationalisation, and theory development.

2. They provide further empirical support for the factors that travel and do not travel across countries and the understanding of it. For example, researchers have been generally unable to validate, within their specific context, the role of assertive principal instructional leadership, perhaps one of the best supported of all *American* correlates. This points at the importance of the societal and political context of education, next to the levels in the educational system and the factors within the different levels.
3. Further internationalisation of educational effectiveness research can promote theoretical and methodological developments with respect to statistical analysis, and the design of studies.
4. Finally, international comparative research can provide a better knowledge base for the relationship between educational effectiveness research on the one hand, and school improvement and improvement of education in general, on the other hand. Of course, these studies have to be carefully carried out in different countries, taking into account the context of the particular countries (Riddell 1997).

After twenty years of research, an exhaustive knowledge base and many empirical studies within and sometimes between countries are available, which can be used for the further design of studies, but which also need further exploration especially with respect to the contextuality between countries and between contexts within countries. One way to do this is to make use of the knowledge base for the improvement of education and for the design of the evaluation of those improvement projects, using concepts and designs of educational effectiveness research. This approach was taken in the evaluation of the Primary Education Quality Improvement Project in Indonesia, which will be described in the next section.

4 The Evaluation of the Primary Education Quality Improvement Project in Indonesia

4.1 Background and Research Questions

The government initiated the Primary Education Quality Improvement Project (PEQIP) in 1992 to improve the outcomes of Primary Education. The objective of PEQIP is to introduce policies and mechanisms for the improvement of the quality of primary education. PEQIP aims at raising student achievement, decreasing retention and drop-out rates and increasing enrolment in secondary education. As such PEQIP can be considered as a school improvement project aimed at enhancing school effectiveness.

During the period September 1996 till September 1997 an evaluation study was conducted to establish the effects of PEQIP. The research questions are the following:
1. Do PEQIP schools differ from non-PEQIP sc hools in student achievement?
2. Which factors explain differences between PEQIP and non-PEQIP schools in student achievement?

The PEQIP activities were centralised in four main components: (1) Teacher Professional Development, (2) Educational Management (including Community Participation), (3) Books and Learning Materials, and (4) Evaluation and Monitoring. These four components refer to the school variables which consistently affect student achievement in developing countries (Levin & Lockheed 1993).

Teacher Professional Development focused on training for tutors and subject matter specialists in new teaching strategies for the subjects Bahasa Indonesia (the official Indonesian language), mathematics, science and social studies. After being trained they

were supposed to transfer their newly acquired knowledge and skills to teachers. Principals, supervisors and education managers at subdistrict, district and provincial level were trained in order to learn how to support the implementation of new teaching strategies in classrooms.

The Management component was directed at training of supervisors, educational managers, and principals. Topics in the training for educational managers and supervisors were the need of participants to raise the quality of primary education participation (using parents for funding, increasing the educational role of parents), the organisation of clusters of schools to support the needs of teachers and principals, monitoring, evaluation and supervision, procurement and budgeting. The main topic in the training for principals was school and cluster management.

The component Books and Learning Materials included some training activities. When it became clear that teachers did not have enough time and knowledge to produce learning materials (as initially they were supposed to do), the focus shifted mainly to the development, organisation and use of libraries, newsletter production and to a better distribution system of teaching and learning material. Student books and teacher guides were provided to the schools as well as science kits, reading kits, globes and other teaching aids for the main subjects. Also items of sports, art and music equipment were provided in some areas. *Inti* schools (central cluster schools) have been provided with video cameras, television sets and overhead projectors.

The Evaluation and Monitoring Component consisted of providing training to supervisors and principals. During the training supervisors and principals learned about recognising the scope of monitoring and evaluation needs, ways of obtaining and using information to satisfy these needs, skills in finding and evaluating information, the use of information obtained by monitoring and evaluation for planning and how to benefit from community participation.

Although Community Participation was included in the Management Component, it was treated as a separate component by the project and by the consultants. Training focused on school fundraising and contributions of services and supplies to meet school needs and on community programs to enforce compulsory education, school attendance, and evening study hours. Training was directed at educational managers, supervisors, principals and community leaders.

4.2 The Research Model

The design of the study was based upon a research model utilising past experiences in educational effectiveness research in Western and developing societies, as described in the sections 2 and 3. As point of departure served the comprehensive model for educational effectiveness as presented in Figure 1. The final model for the evaluation study of PEQIP is presented in Figure 2. The factors at different levels of the educational system are not expected to influence the outcomes directly, instead they are supposed to provide conditions for the effectiveness at the school and classroom level. The same holds for the PEQIP interventions. These are certainly supposed to influence outcomes by mediating other factors. The interventions will be carried out first. They have to be implemented at the cluster and school level and ultimately they have to reach the classroom level. The model takes the PEQIP inputs as starting point and arrives at the student outcomes by mediating factors.

Figure 2: Research Model for the Evaluation of PEQIP
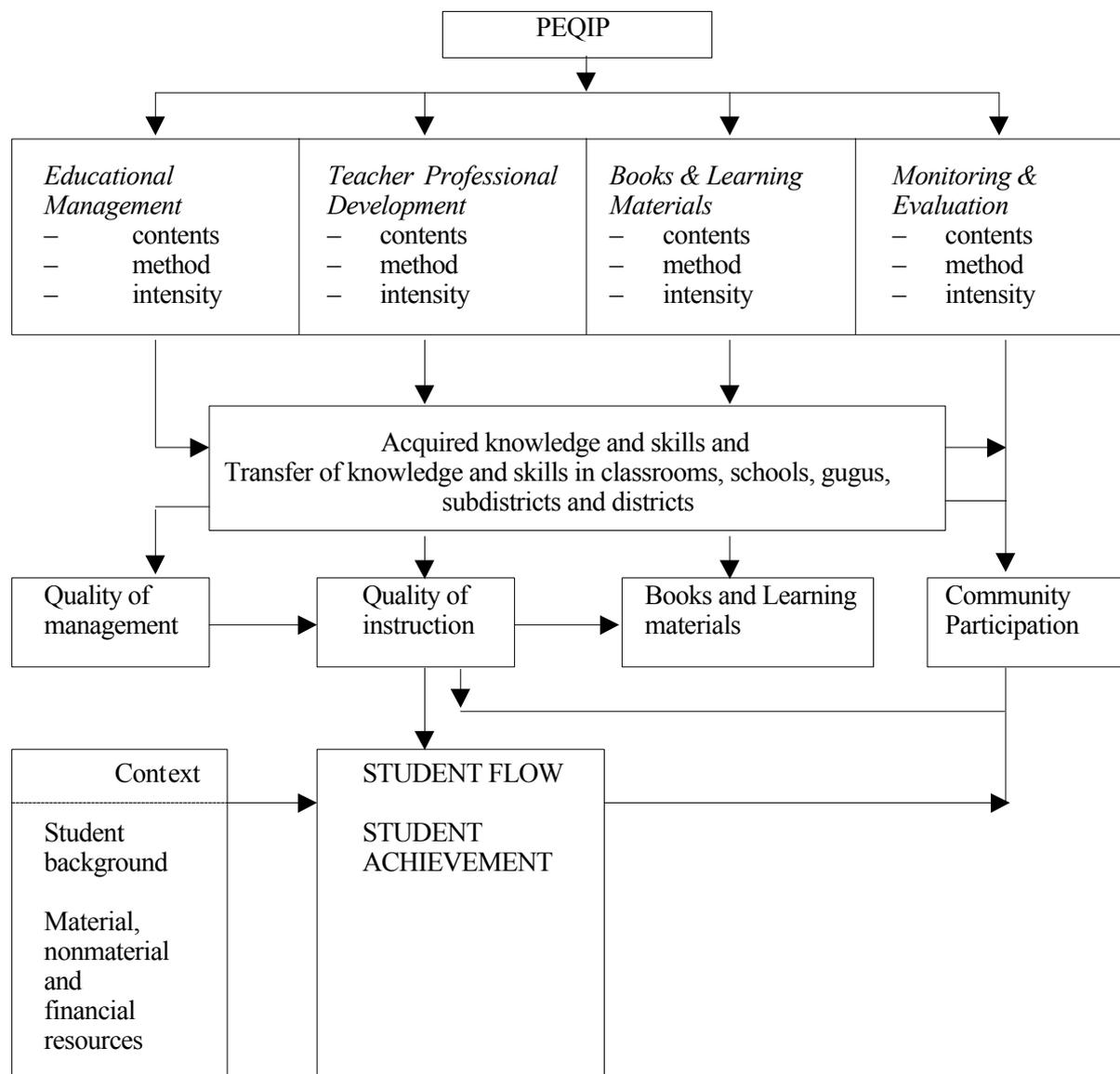
Sample

The study was conducted in the provinces Aceh and Sulawesi Utara. Within each province 27 PEQIP schools were selected from all three subdistricts participating in PEQIP.

In each of the subdistricts three clusters of schools were participating in PEQIP. For the study all *Inti* schools (central cluster school) in these clusters were selected, as well as every third and sixth *Imbas schools* (satellite school within school cluster). So in each province 9 Inti schools and 18 Imbas schools were participating. In each province also control schools were selected, 14 in each province.

Instruments

For the study, standardised tests related to the national curriculum were developed for Bahasa Indonesia, Mathematics and Science in grade 6. The data were used as pre- and posttests.

Indicators for student background characteristics were intelligence, socio-economic status (SES), home language, gender, distance between home and school, and kindergarten attendance.

For measuring the variables with respect to the context, resources, inputs and implementation, a variety of instruments were developed. The quality of instruction in classrooms for Bahasa Indonesia, mathematics and science was observed twice by the Indonesian researchers and once by the supervisors. The observation instrument consisted of a low inference part (a coding sheet on which the observer coded every minute what was happening during the lesson) and a high inference part (a rating scale consisting of 34 statements about classroom organisation, aspects of teaching and classroom environment).

Analysis of data

The data were analysed with the multilevel program VARCL, separately for Bahasa Indonesia, mathematics and science. Firstly, an unconditional model was specified in order to establish the proportions of variance to be explained at the pupil and school level. Then, the following models were specified. Model (1) with pupil variables; Model (2) with pupil and school context variables. Model (3) with category of school (contrast Inti vs Control; contrast Imbas vs Control).

Results

The results are presented in Table 3. Pupil and school context variables explained the largest part of the variance between schools. Including both dummy variables for category of school in the final model reduced the unexplained variance with only 1 or 2 percent. This addition only improved significantly the fit of the model for science. The difference between schools in science achievement are explained by PEQIP.

In Table 4 the unstandardised regression coefficients and standard errors of pupil and school context variables and category of school per subject for model 3 are presented. With respect to the Category of school the table confirms the data on Table 4. The effects are only significant for science. In the further analyses it turned out that this was only the case for Sulawesi Utara. In this province also the effects for Bahasa Indonesia were significant.

Table 3: Proportions of Variance at Pupil and School Level, Per Subject

	Model 0	Model 1	Model 2	Model 3
Mathematics				
Pupil level	54%	40%	40%	40%
School level	46%	25%	23%	23%
Total	100%	65%	63%	63%
Deviance	11911.98	11330.79*	11324.84	11323.63
Bahasa Indonesia				
Pupil level	61%	45%	45%	45%
School level	39%	26%	20%	20%
Total	100%	71%	65%	65%
Deviance	10462.35	10326.46*	9887.61*	9887.47
Science				
Pupil level	59%	53%	53%	53%
School level	41%	27%	17%	14%
Total	100%	80%	70%	67%
Deviance	10972.96	10773.98*	10704.02*	10691.53*

* Significant improvement of model fit ($p < .05$).

Model 0 = model without school and pupil variables
Model 1 = model with pupil variables intelligence, SES and pretest score (except for science)
Model 2 = model with pupil variables   and school context variables location, % low SES pupils and % pupils with Kindergarten
Model 3 = model with pupil and school context variables and category of school

Table 4: Regression Coefficients and Standard Errors of Pupil and School Context Variables and Category of School, Per Subject (Model 3)

	Mathematics	Bah. Indo.	Science
Grand Mean	5.3	11.8	9.7
Pretest score	.52(.02)*	.38(.02)*	n.a.
Intelligence	.16(.02)*	.05(.01)*	.22(.02)*
SES	.51(.14)*	.24(.09)*	.62(.12)*
% Kindergarten	.02(.01)*	.01(.01)	.02(.01)*
Location	-.12(.68)	-.05(.41)	-1.40(.42)*
% Low SES pupils	-.01(.01)	-.02(.01)*	-.02(.01)*
Cat. of school (Inti)	.91(1.0)	.18(.61)	2.10(.62)*
Cat. of school (Imbas)	1.20(1.2)	.26(.76)	2.30(.77)*

* $p < .05$

In order to explain the effects of PEQIP the following analyses were carried out. Firstly, we selected the school and classroom variables that, according to F-tests and χ^2-tests, differed significantly ($p < .05$) between PEQIP and control schools. These variables may be considered as indicators of the implementation of the PEQIP inputs in the practice of schools and classrooms. These variables were entered in the model with pupil and school context variables in two steps. In the first step, the school level variables were entered (model 4). In the second step, the classroom variables were entered (model 5). Subsequently, all school and classroom variables that were significant in the models 4 and 5, were entered at the same time (model 6). In the final model (model 7) Category of school was entered again. In Table 5 the results of the last two models are presented.

Table 5:   Results Multilevel Analyses with School and
             Classroom Variables, Per Sub ject

	Math.	Bah.Ind.	Science
Grand Mean	21.9	18.6	.60
Perc. Unexplained Variance			
Pupil level	40	45	53
School level	12	10	9
Total	52	55	62
Improvement model fit [1]	47.1**	46.0**	43.1**
Regression coefficients:			
Student variables			
Pretest	.51(.03)*	.38(.02)*	n.a.
Intelligence	.16(.02)*	.05(.01)*	.22(.02)*
SES	.50(.14)*	.26(.09)*	.63(.12)*
School context			
% Kindergarten	.04(.02)*	.02(.01)*	.04(.01)*
%Low SES	.00(.01)	.03(.01)*	-.01(.01)
Location	.68(.84)	.03(.01)*	-.62(.56)
Province (Sulawesi)	-4.6(1.4)*	-2.1(.85)*	-.62(.88)
Impl. Variables			
Teacher Prof. Dev.			
Experience teacher	-.15(.06)*	-.09(.03)*	.05(.04)
Homework	-.28(.29)	-.07(.16)	-.02(.18)
Time spent to subject	.01(.00)*	-.00(.00)	-.01(.00)*
Presentation of content	-.03(.07)	-.04(.05)	-.4(.02)*
Pupils working	-.02(.06)	-.04(.05)	-.06(.02)*
Other activities	-.16(.08)*	-.08(.07)	.01(.05)
Active learning time	.00(.02)	.00(.01)	-.00(.01)
Quality of instruction	-1.8(1.9)	-1.7(1.3)	-.24(.13)
Innovative teaching	-.14(.18)	1.9(.98)*	1.9(.97)*
Frequency of testing	-.48(.73)	-.44(.43)	.11(.56)
Use of test results	.01(.56)	.31(.34)	-.51(.41)
Pupils' attention	.54(1.7)	.36(.96)	3.4(2.5)
Questioning	1.7(.88)*	.28(.46)	.37(.53)
Comprehensive questions	.14(.62)	-.03(.33)	.44(.45)
Monitoring work	.72(.80)	-.33(.39)	.06(.57)
Grouping of pupils	.46(.73)	-.20(.47)	.03(.34)
Management/Evaluation			
Observations in classrooms	.06(.07)	.12(.05)*	.09(.07)
Evaluation of teachers	2.5(1.2)*	.10(.85)	-2.5(1.0)*
Evaluation school quality	2.3(1.6)	.25(.95)	2.8(1.0)*
Books and learning mat.			
Availability of equipment	.16(.18)	.06(.10)	-.09(.12)
Availability of student books	-.01(.01)	.00(.00)	.01(.00)*
Community participation			
Homework regulations	-1.6(1.1)	-.23(.60)	.16(.71)
Voluntary work of parents	.34(.21)	.33(.12)*	.07(.15)
Educ. Involvement of parents	-2.5(2.9)	-1.6(1.8)	-.42(2.4)
Help with homework	2.9(1.4)*	1.3(.89)	1.3(1.1)
Signing homework	1.6(2.5)	.02(.96)	.04(1.3)
Financial contribution	-.02(.05)	.00(.03)	.03(.03)

[1] Difference compared to the Deviance of model 2, the model with pupil and school context variables.
* p < .05; ** p < .01

Starting with Bahasa Indonesia, the results of model 6 show that, after taking into account pupil and school context variables as well as school and classroom variables, the unexplained variance at school level has reduced to 6 percent. All school and classroom variables, with the exception of the frequency of homework and signing homework by parents, are significant contributors. The effects of the experience of the teaching staff and of the teacher of grade 6 are negative as well as the effects of parental involvement. The results of model 7 show that if taking into account Category of school, the unexplained variance at school level is reduced with another 2 percent. The model fit has improved significantly as well. However, the regression coefficients of Inti and Imbas

schools show that the effects of both categories are not significant.

For science, in model 6 only the variable innovative teaching has a significant positive effect. Adding category of school to the model reduces the unexplained variance at school level, but does not significantly improve the model fit. The effect of innovative teaching is not significant anymore and again there is a significant positive effect of Inti schools.

4.3 Conclusions

With respect to the first research question addressed in this article, we may conclude that overall PEQIP schools only differed from non-PEQIP schools in student achievement in Bahasa Indonesia and science for Sulawesi Utara, after taking into account differences in student background characteristics, like intelligence, SES, kindergarten attendance and after taking into account the location of the schools and the composition of their student population.

The effects are not strong. Inti schools score 2 points higher on   a Bahasa Indonesia test of 40 items and almost 3 points on a science test of 40 items, compared to control schools. This is about half a standard deviation. Also from a cost-effectiveness point of view these results are quite disappointing (Creemers & Werf forthcoming). The small effects may be explained by the fact that, as appeared from our descriptive data, the implementation of PEQIP in the practice of schools was not always successful. PEQIP schools did not differ much from non-PEQIP schools with respect to all indicators of the implementation of the PEQIP components.

The fact that PEQIP did not have any effects in Aceh may be explained by the way the local authorities in this province treated the schools during PEQIP. They not only supported the PEQIP schools, but were also strongly involved in all schools in the province, thus in the control schools as well. They provided the same books and learning materials to the non-PEQIP schools as the PEQIP schools received from the project management. In Sulawesi Utara this was not the case. In that province only PEQIP schools were supported and provided with materials. The strong involvement of the local authorities in schools in Aceh probably also is the reason why student achievement in this province is substantially higher than in Sulawesi Utara (and than in other provinces also, as appeared from other output data). From an educational effectiveness point of view this is an interesting finding. It supports the assumption in the multilevel models that factors at higher levels also are important conditions for the effectiveness of schools (Creemers 1994; Scheerens 1992; Tsang & Wheeler 1993).

Explaining the effects of PEQIP in Sulawesi Utara with variables indicating the implementation of the PEQIP components shows some remarkable results. Firstly, the effects only can be explained for Bahasa Indonesia. After taking into account school and classroom variables the effects of category of school are not significant anymore. The effects of PEQIP on science achievement could not be explained by school and classroom variables. Only one variable – innovative teaching – indicating the implementation of the Teacher Professional Development component, explains the between-school variance in science achievements, but this effect disappears if category of school is added to the model. So with respect to science, we have to conclude that there is an effect of PEQIP, but it is impossible to explain this effect. An explanation for this finding may be that the variable innovative teaching is very strongly related to the category of schools. In fact only in Inti schools innovative teaching was observed in classrooms during science lessons.

The effects of PEQIP on student achievement in Bahasa Indonesia are explained by only a few variables that are related to the PEQIP interventions. We may conclude that there are some effects of the PEQIP components Teacher Professional Development, Community Participation and Evaluation and Monitoring. The effects of the first component, indicated by innovative teaching, is the strongest. This is a variable at classroom level and refers to grouping of pupils, using a variety of teaching methods, asking comprehensive questions to pupils, stimulating interaction between pupils and keeping pupils actively involved during the lessons. This finding clearly supports the assumption of educational effectiveness models that variables at classroom level are the most important for improving student achievement and that these variables are not only relevant in Western countries but also in developing countries (Creemers 1994; Stringfield & Slavin 1992). It also confirms the assumption that an active approach of teaching in countries where traditionally mainly passive learning takes place is a successful tool to improve student achievement (Lockheed 1993). The finding of this study that innovative teaching, although it could not explain the effects of PEQIP, is related to science achievement also, gives additional support to this conclusion.

Another variable that contributes to the explanation of the effects is the frequency of observations in classrooms by the principal. This variable can be considered as an indicator of Evaluation and Monitoring. Apparently the activities of principals support good teaching practices at classroom level, which again confirms the assumption in the educational effectiveness models that school level variables are conditional for effective teaching in classrooms.

Also the amount of voluntary work conducted by parents, indicating Community Participation, had a positive effect on student achievement. This variable can be considered as a contribution of local resources to the school and as an indicator of the responsibilities of the parents to support and reinforce the education of their children. From community participation however also negative effects are apparent, like involvement of parents in educational matters and signing homework by parents. These findings support the idea that community participation is an important concept to take into account in theoretical models of educational effectiveness in developing countries (Levin & Lockheed 1993), but more research is needed to establish which variations of community participation are effective and which are not.

No effects of resources, books and learning materials were found in this study. With respect to resources and learning materials the results are in agreement with those of other studies in Western and developing countries (Hanushek 1986; Harbison & Hanushek 1992; Lockheed 1993). There are several explanations for the absence of effects of student books and teacher guides. Firstly, there is not much variation in availability of books and guides between schools in Indonesia. Secondly, the variation that exists is strongly related to location and composition of the student population of the school. Taking into account these context variables the effects of books and guides disappear. Thirdly, not the availability of books and guides matters but the quality and the frequency and way of using these in the teaching-learning process (Riddell 1997). Just like in other developing countries, it was observed that the frequency of using books in classrooms was very low. Apparently the provision of materials does not guarantee their use (Fuller & Snyder 1989).

Finally, two effects of immaterial resources were found. The experience of the teaching staff and of the teacher in grade 6 appeared to have negative effects. These effects may be explained by the level of teacher education which is in Indonesia strongly related to age and experience. Older teachers generally have a lower level of teacher

training, because they were educated before specific teacher qualifications were demanded. As may be expected, higher qualified teachers have better subject matter knowledge. Research in developing countries including Indonesia has proved frequently that teachers' subject matter knowledge does have strong positive effects on student achievement (Fuller 1987; Suryadi 1992). So, subject matter knowledge of teachers also is a concept to take into account in theoretical models of educational effectiveness in developing countries.

Summarising, we may conclude that programs like PEQIP do make sense to improve the quality of primary education. Aiming at improvements in the teaching process at classroom level is the most promising strategy to realise higher achievements. Improvement of management at school level and effective involvement of parents in the school might support the processes at classroom level.

5   Lessons Learned

Educational effectiveness research and improvement has a short history, although it is based on theory and research with respect to education in the past. This research is mainly *within* countries, especially Western countries. During the last decades major progress was made in different areas: in theory, in research methodology, and in improvement and intervention projects. It seems that this knowledge base and research may bring benefits for education, for educational research and educational effectiveness in developing as well as industrialised societies. Educational effectiveness research provides a useful instrument to evaluate interventions and improvement pro   jects, but also the knowledge base on educational effectiveness can be used for educational improvement projects. At least at the classroom level, it turns out that the same factors contribute to the effects of education in terms of student outcomes. This is supported by other comparative studies in educational effectiveness, like the International School Effectiveness Research Project (ISERP). It seems that the same factors at classroom level induce better learning and better learning outcomes by students. However, the same processes at the classroom/group level, the teaching learning processes, can be or maybe have to be introduced by different instruments, factors at the other levels of the educational system. Sometimes, different conditions at the school cluster and other system levels are directed to achieve the same processes at the teaching and learning level (classroom). The evaluation of the PEQIP study also provides that kind of information that can be used in the further development of contextual, multilevel, multi-factor theory about educational effectiveness. This kind of research within different cultural contexts comparing with other cultures are extremely important for the further elaboration and understanding of educational effect. The results point at two major issues: (1) the differences in context and the differences at levels above the classroom; (2) the communalities between different countries about what happens at the teaching/learning level. In this way the results contribute to the further development of core theories on educational effectiveness.

These insights into the generalisability of factors about teaching and learning, the differences at the several other levels, although related to the broader context, e.g. monitoring, evaluation and leadership, again stress the point that in improvement pro   jects this kind of information can be useful for the design of further interventions.

In this respect it is also important to look at the costs of interventions related to the effect sizes they can achieve (Creemers & Werf forthcoming).

We agree that there is a common knowledge base with respect to factors that are important for effectiveness of teaching and learning processes and with respect to the

different effects factors can have at the other levels of the system, in order to provide the conditions for the teaching-learning situation.

It becomes clear that in international consultancy and research knowledge should be available at least about (1) the content of intervention (about educational effectiveness and the way education can be improved using the available knowledge); (2) procedural and technical knowledge (i.e., how this kind of interventions might be designed and evaluated); and (3) contextual knowledge (how to apply educational effectiveness knowledge to national and local circumstances). Next to the competency of transferring knowledge to professionals, the two first measured competencies, content knowledge and procedural knowledge, might be expected from international experts. The third mentioned competency, contextual knowledge, and procedural knowledge, how to transform knowledge to local needs, might be expected from local experts. However, the latter group gradually may acquire the first two mentioned competencies as well. That will increase the cost effectiveness of educational interventions.

This implies that the best educational specialists, researchers, and developers should be contracted for international consultancy and support. A part of the intervention program (terms of references) should be the transfer of knowledge to local experts, who might be included more and more in international networks of educational expertise and research.

Bibliography

Bashir, S. (1994). *Public versus private in primary education: A comparison of school effectiveness and costs in Tamil Nadu.* Unpubl. Ph.D. dissertation, Faculty of Economics, London School of Economics.

Brookover, W.B., Beady, C., Flood, P. & Schweitzer, J. (1979). *School systems and student achievement: schools make a difference.* New York: Praeger.

Coleman, J.S., Campbell, E.Q., Hobson, C.F., McPartland, J., Mood, A.M., Weinfeld, F.D. & York, R.L. (1966). *Equality of Educational Opportunity.* Washington, D.C.: US Government Printing Office.

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

Creemers, B.P.M. & Scheerens, J. (1994). Developments in the educational effectiveness research programme. *International Journal of Educational Research, 21*, 125–140.

Creemers, B.P.M. & Werf, M.P.C. van der. (forthcoming). Economic viewpoints in educational effectiveness: Cost-effectiveness analysis of an educational improvement project. *School Effectiveness and School Improvement.*

Edmonds, R.R. (1979). Effective schools for the urban poor. *Educational Leadership, 37* (1), 15–27.

Fuller, B. (1987). What factors raise achievement in the Third World? *Review of Educational Research, 57*, 255–292.

Fuller, B. & Heyneman, S.P. (1989). Third World school quality. Current collapse, future potential. *Educational Researcher, 18* (2), 12–19.

Fuller, B. & Snyder, C. (1989). *Vocal teachers, silent pupils? Live in Botswana classrooms.* Paper presented at the Annual Meeting of the Comparative and International Education Society, Boston, Massachusetts.

Gray, J. & Wilcox, B. (1995). *Good school, bad school.* Buckingham: Open University Press.

Hanushek, E.A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*, 1141–1177.

Harbison, R.W. & Hanushek, E.A. (1992). *Educational performance of the poor: Lessons from rural northeast Brazil.* New York: Oxford University Press.

Hargreaves, D.H. (1995). School culture, school effectiveness and school improvement. *School Effectiveness and School Improvement, 6*, 23–46.

Heyneman, S.P. & Loxley, W.A. (1983). The effect of primary-school quality on academic achievement across twenty-nine high- and low-income countries. *The American Journal of Sociology, 88* (6), 1162–1194.

Kaluge, L. (1998). *Some factors related to educational attainment in Indonesian primary schools.* Unpubl. Doctoral dissertation, University of London, Institute of Education.

Levin, H.M. & Lockheed, M.E. (Eds.). (1993). *Effective schools in developing countries.* London: Falmer.

Levine, D.U. & Lezotte, L.W. (1990). *Unusually effective schools: a review and analysis of research and practice.* Madison (USA): National Center for Effective Schools Research and Development.

Lockheed, M.E. (1993). The condition of primary education in developing countries. In H.M. Levin & M.E. Lockheed (Eds.), *Effective schools in developing countries* (pp. 20–41). London: Falmer.

Lockheed, M., Fuller, B. & Nyirongo, R. (1989). Family effects on students' achievement in Thailand and Malawi. *Sociology of Education, 62*, 239–256.

Lockheed, M. & Longford, N. (1991). School effects on mathematics achievement gain in Thailand. In S. Raudenbush & D. Willms (Eds.), *Schools, classrooms and pupils* (pp. 131–148). San Diego: Academic.

Ralph, J.H. & Fennessey, J. (1983). Science or reform: some questions about the effective schools model. *Phi Delta Kappan, 64* (10), 689–694.

Raudenbush, S., Kidchanapanish, S. & Kang, S. (1991). The effects of preprimary access and quality on educational achievement in Thailand. *Comparative Education Review, 35* (2), 255–273.

Reynolds, D., Creemers, B.P.M., Nesselrodt, P.S., Schaffer, E.C., Stringfield, S. & Teddlie, C. (Eds.). (1994). *Advances in School Effectiveness Research and Practice.* Oxford: Pergamon.

Riddell, A.R. (1989). An alternative approach to the study of school effectiveness in Third World countries. *Comparative Education Review, 33* (4), 481–497.

Riddell, A.R. (1997). Assessing designs for school effectiveness research and school improvement in developing countries. *Comparative Education Review, 41* (2), 178–204.

Rutter, M., Maughan, B., Mortimore, P. & Ouston, J. (1979). *Fifteen Thousand Hours.* London: Open Books.

Sammons, P., Hillman, J. & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research.* London: OFSTED.

Scheerens, J. (1992). *Effective Schooling: Research, Theory and Practice.* London: Cassell.

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

Scheerens, J. & Creemers, B.P.M. (1989). Conceptualizing school effectiveness. *International Journal of Educational Research, 13*, 691–706.

Singh, S., Saxena, R.R., Jain, V.K., Arora, O.P., Rath, K.B., Gupta, J.K. & Gupta, M.K. (1995). Effect of school policies and practices on students achievement. In K. Kumar (Ed.), *School effectiveness and learning achievement at primary stage: International perspective* (pp. 143–166). New Delhi: National Council of Educational Research and Training.

Slavin, R.E. (1987). A theory of school and classroom organization. *Educational Psychologist, 22* (2), 89–108.

Slavin, R.E. (1996). *Education for All.* Lisse: Swets & Zeitlinger.

Stringfield, S. (1994). A model of elementary school effects. In D. Reynolds, B.P.M. Creemers, P.S. Nesselrodt, E.C. Schaffer, S. Stringfield & C. Teddlie (Eds.), *Advances in School Effectiveness Research and Practice* (pp. 153–187). Oxford: Pergamon.

Stringfield, S.C. & Slavin, R.E. (1992). A hierarchical longitudinal model for elementary school effects. In B.P.M. Creemers & G.J. Reezigt (Eds.), *Evaluation of effectiveness* (pp. 35–69). Groningen: ICO.

Suryadi, A. (1992). *Improving the educational quality of primary education.* Jakarta: Office of Education and Culture Research and Development.

Tsang, M.C. & Wheeler, C. (1993). Local initiatives and their implications for a multi-level approach to school improvement in Thailand. In H.M. Levin & M.E. Lockheed (Eds.), *Effective schools in developing countries* (pp. 108–130). London: Falmer.

Werf, M.P.C. van der. (1995). *The educational priority policy in the Netherland. Context, implementation and outcoumes.* The Hague: SVO.