ROUTLEDGE

Methodological Advances in Educational Effectiveness Research

QUANTITATIVE METHODOLOGY SERIES

Bert P. M. Creemers | Leonidas Kyriakides | Pam Sammons

Methodological Advances in Educational Effectiveness Research

Methodological Advances in Educational Effectiveness Research is an important new work by some of the leading researchers in the field of Educational Effectiveness Research (EER). This book provides a state-of-the-art snapshot of the methodology of EER now and clearly demonstrates the way it is applied in both research and evaluation. It shows how developments in the research methodology area, such as the use of multilevel modelling approaches to analyse nested data, have promoted the knowledge base of educational effectiveness. But at the same time, as the authors show, the knowledge base of educational effectiveness and the attempt to establish theoretical models do paradoxically challenge the development of methodologically appropriate studies, including ways of analysing data.

This book guides readers though the effective and appropriate use in educational effectiveness of:

- longitudinal studies
- experimental studies
- mixed research methods
- meta-analyses of effectiveness studies
- using Item Response Theory to measure outcomes and factors
- using Generalizability Theory to test the quality of data
- multilevel modelling
- Structural Equation Modelling techniques.

The authors draw in the expertise of scholars from around the world to show the mathematical background of each technique, current and future applications and specific examples of applying each orientation to help the readers design their own effectiveness studies using specific methodological tools.

Bert P.M. Creemers is Professor and Director of the Centre for Learning and Teaching, University of Groningen.

Leonidas Kyriakides is Associate Professor of Educational Research and Evaluation, Department of Education, University of Cyprus.

Pam Sammons is Professor of Education, Department of Education, University of Oxford.

Quantitative Methodology series
Series Editor: George A. Marcoulides

This series presents methodological techniques to investigators and students. The goal is to provide an understanding and working knowledge of each method with a minimum of mathematical derivations, while retaining the clarity and rigor of quantitative research. Each volume focuses on a specific method – for example factor analysis, multilevel analysis, structural equation modelling, providing an up-to-date snapshot of current methodological practice in qualitative studies.

The authors and contributors draw upon the expertise of scholars from around the world to show the diversity and demands of quantitative research, providing current and future applications, as well as specific examples to help readers design their own studies using methodological tools.

Previous titles in the series include:

Modern Methods for Business Research
George A. Marcoulides

Latent Variable and Latent Structure Models
George A. Marcoulides and Irini Moustaki

Multilevel Analysis: techniques and applications
Joop Hox

Studying Educational and Social Policy: theoretical concepts and research methods
Ronald Heck

New Developments in Categorical Data Analysis for the Social and Behavioral Sciences
L. Andries Van der Ark, Marcel Croon and Klaas Sijtsma

An Introduction to Latent Variable Growth Curve Modeling: concepts, issues, and applications (second edition)
Susan Duncan and Lisa Strycker

An Introduction to Multilevel Modeling Techniques (second edition)
Ronald Heck and Scott Thomas

Applying Generalizability Theory Using EduG
Jean Cardinet, Sandra Johnson and Gianreto Pini

Multilevel and Longitudinal Modeling with PASW/SPSS
Ronald Heck, Scott Thomas and Lynn Tabata

Multilevel Analysis: techniques and applications (second edition)
Joop Hox

Methodological Advances in Educational Effectiveness Research

Bert P.M. Creemers,
Leonidas Kyriakides and
Pam Sammons

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2010
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Avenue, New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group, an informa business

This edition published in the Taylor & Francis e-Library, 2010.

To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Creemers, Bert P.M.
Methodological advances in educational effectiveness research/Bert P.M. Creemers, Leonidas Kyriakides, and Pam Sammons. – 1st ed.
p. cm.
Includes bibliographical references and index.
1. Education – Research – Methodology. I. Kyriakides, Leonidas. II. Sammons, Pam. III. Title.
LB1028.C736 2010
370.7′2–dc22 2009048160

ISBN 0-203-85100-5 Master e-book ISBN

ISBN10: 0–415–48175–9 (hbk)
ISBN10: 0–415–48176–7 (pbk)
ISBN10: 0–203–85100–5 (ebk)

ISBN13: 978–0–415–48175–5 (hbk)
ISBN13: 978–0–415–48176–2 (pbk)
ISBN13: 978–0–203–85100–5 (ebk)

Contents

Contributors

Authors of the book

Bert P.M. Creemers, Centre for Learning and Teaching, University of Groningen. b.p.m.creemers@rug.nl

Leonidas Kyriakides, Department of Education, University of Cyprus. kyriakid@ucy.ac.cy

Pam Sammons, Department of Education, University of Oxford. pamela.sammons@education.ox.ac.uk

Authors of/contributors to Part B of the book

Jan-Eric Gustafsson, University of Gothenburg, Sweden. jan-eric.gustafsson@ped.gu.se

Hans Luyten, University of Twente, Department of Educational Organisation and Management. j.w.luyten@gw.utwente.nl

George A. Marcoulides, Graduate School of Education, University of California, Riverside. georgem@ucr.edu

Robert E. Slavin, Johns Hopkins University and University of York. rslavin@successforall.org

Charles Teddlie, Department of Educational Theory, Policy and Practice, Louisiana State University. edtedd@lsu.edu

Norman Verhelst, CITO, the Netherlands. norman.verhelst@cito.nl

Figures

Tables

Preface

Compared to other areas of research in the social sciences and even in education, Educational Effectiveness Research (EER) has had a relatively short history of about 40 years. This short history is, however, already fairly well documented and in most cases the authors have given recommendations for the development of EER in the future. However, although there have been many reviews and several handbooks focusing on EER, there has been little systematic attention paid to documenting and analysing methodological advances. In this book, therefore, we seek to provide an overview of the methodological development of EER and make a plea for a more theoretical orientation in educational effectiveness, for more experimental and longitudinal research and for further rigorous evaluation of improvement projects in schools and classrooms.

We were already convinced – and after reading the contributions to this publication our conviction has been strengthened – that EER makes progress through the advancements in the research methodology of, and the statistical techniques used in, the analysis of relevant hierarchically structured data sets. However, in our view it is not only the quality of the research design, the instruments and the statistical analysis that promote the knowledge base of educational effectiveness but also the development of better theories about the complex nature of educational effectiveness. In our opinion, the theoretical development of educational effectiveness both asked for and also promoted methodological advances, and the further development of educational effectiveness theory is thus facilitated by various methodological and technical advances that have been made over the last two decades or so. These advances have enabled researchers to test the various components of theories of educational effectiveness and to evaluate improvement practices, which in turn have pushed our thinking and theory development further.

In this publication we concentrate on describing and explaining the main methodological advances that we believe will help readers to improve the quality of future EER studies. Thus, we do not restrict ourselves to a simple overview of different research tools, but instead we choose to present them from the perspective of the development and testing of a theoretical body of knowledge about effectiveness. For this reason we provide the readers in the first part of

the book with an analysis of the background to EER, an account of the advances in research and theory made in the past and an analysis of the mutual relationship between methodological advances and progress in theory development. Special attention is given to two issues related to our theoretical orientation in educational effectiveness: first, the difficulties in demonstrating causal relations between the various context, input, output and process factors of interest in EER and the most fruitful avenues for identifying causality, and second, an analysis of the implications of this orientation for research design. The second issue is concerned with the purpose of EER, to have an impact on educational practice in classes, schools and on policy and ways to promote this impact and evaluate it.

In the second part of the book – the main core of the publication – we provide readers with a state-of-the-art overview of advances in the EER methodology with respect to design measurement theories and data analysis. Experts in each of the areas (in most cases in collaboration with the editors) provide the background to each of the specific approaches of particular interest, present a comprehensive overview and give examples of the use of the approach and/or technique in specific EER studies to illustrate their application. After this presentation of advances made in research design, measurement and data analysis, we return in the third part of the book to a discussion of the main implications for EER. We draw conclusions about the implications of such methodological advancements for the future of EER, especially with respect to the prospect for improved development and testing of theories about educational effectiveness, which can have an impact on practice and policy. In order to promote excellence in research we provide readers with an instrument (which we regard as analogous to a conceptual road map or guide) that will support them in making decisions in the development of their own research plans and in the implementation of those plans.

In writing this book we received support from many colleagues, policymakers, practitioners and our families. We would like to mention some of them, especially the following. Our colleague Dr James Hall provided us with constructive feedback of the early draft of the book and challenged us to sharpen our arguments. Our friend Ioannis Yiannakis, who is currently an inspector and used to be a teacher for many years, gave us comments that helped us identify the extent to which our work could contribute in the improvement of policy and practice. We also thank him for helping us to draft figures of the proposed conceptual map for conducting methodologically appropriate effectiveness studies. The research assistants on our team and especially our PhD students gave us comments from the perspective of young researchers in the field of EER. Evi Charalambous helped us in the production of the manuscript and supported us in the process of linguistic editing. Finally, our universities were supportive in facilitating our academic efforts to write the book.

As we have already mentioned, in producing this volume our main purpose is to promote future high-quality research in educational effectiveness. We see the current advancements in methodology and statistical techniques as great

opportunities to improve the quality of EER. Moreover, the relationship can be seen as reciprocal because, as we show, some of the key advancements have been promoted and fostered by research in the educational effectiveness field. We welcome comments, criticism and contributions to its further development and anticipate that new knowledge will be produced by readers with different perspectives in education and research. We hope that readers, and especially students, will be challenged to conduct high-quality research in this area in order to contribute to the growing EER evidence base and to enhance current understanding and the use of EER in the improvement of educational practice.

Part A

The state of the art of Educational Effectiveness Research

Challenges for research methodology

Chapter 1

Background to Educational Effectiveness Research

Introduction

The essential difference between this book and other books on research methodology is a focus on methodological advances in a specific area within research on education, namely Educational Effectiveness Research (EER). Thus, this introductory chapter offers a background to EER and helps readers recognize the importance of the progress that has been made in research methodology for this field. This chapter also seeks to enable EER researchers to identify ways of making use of advanced research methods that will further promote the development of this field. At the same time, this chapter also provides readers who are more generally interested in research methodology with a brief account of the background to EER in order to help them understand the specific context in which advanced research methods can be applied and so contribute to the ongoing development of this field.

In this introductory chapter we give a short outline of the history of EER, identify its essential characteristics and illustrate the strengths and weaknesses of this field. The progress made in modelling educational effectiveness is presented and major research questions are identified that have implications for the choice of most appropriate research methodology. It is shown that developments in research methodology, such as in multilevel modelling to analyse nested data, have promoted the knowledge base of educational effectiveness. Further, this chapter outlines attempts by researchers to establish theoretical models and describes how the complexity of educational effectiveness has provided challenges to the development of methodologically appropriate studies, including ways of analysing data. Finally, the aims of this book and its structure are then outlined.

Educational Effectiveness Research can be seen as an overarching theme that links together a conglomerate of research in different areas, including research on teacher behaviour and its impacts; curriculum; student grouping procedures; school organization; and educational policy. The main research question underlying EER is the identification and investigation of which factors in the teaching, curriculum and learning environments (operating at different levels, such as the classroom, the school, and levels above the school) can directly or indirectly explain measured differences (variations) in the outcomes of students.

Further, such research frequently takes into account the influence of other important background characteristics, such as student ability, socio-economic status (SES) and prior attainment. Thus, EER attempts to establish and test theories that explain why and how some schools and teachers are more effective than others in promoting better outcomes for students. However, it is also important to note that the three terms – school effectiveness, teacher effectiveness and educational effectiveness – are used inconsistently in the literature and that these are themselves interrelated. In this book, 'school effectiveness' is taken to mean the impact that school-wide factors, such as a school policy for teaching, school climate and the 'mission' of a school, have on students' cognitive and affective performance. On the other hand, 'teacher effectiveness' is taken to mean the impact that classroom factors have on student performance, and includes teacher behaviour, teacher expectations, classroom organization and use of classroom resources.

Teddlie (1994) argued that most teacher effectiveness studies have been concerned only with the processes that occur within classrooms, to the exclusion of school-wide factors, whereas most school effectiveness studies have involved phenomena that occur throughout the school with little emphasis on particular teaching behaviours within individual classrooms. Only a few EER studies have attempted to examine both school and classroom effectiveness simultaneously (Mortimore *et al.* 1988; Teddlie and Stringfield 1993), although this weakness has begun to be addressed in recent studies (de Jong *et al.* 2004; Kyriakides 2005; Reynolds *et al.* 2002; Opdenakker and Van Damme 2000). The attempts to deal with both teacher and school influences can be seen as a significant development in EER, since joint studies on school and teacher effectiveness reveal that neither level can be adequately studied without considering the other (Reynolds *et al.* 2002). In this context, we are using the term educational effectiveness rather than teacher and/or school effectiveness to emphasize the importance of conducting joint school and teacher effectiveness research, which can help us identify interactions between the school, classroom and student levels and their contributions in explaining variation in student outcomes, both cognitive and non-cognitive. Finally, it is important to note that EER also refers to the functioning of the educational system as a whole, and this research can, therefore, also be used to support the development and testing of different models of effectiveness (Creemers 1994; Creemers and Kyriakides 2008; Scheerens 1992; Stringfield and Slavin 1992). In turn, these models of effectiveness ultimately attempt to explain why educational systems and their subcomponents perform differently, with the aim of providing relevant evidence for policymakers.

History of Educational Effectiveness Research

The origins of EER largely stem from reactions to seminal works on equality of opportunity in education that were conducted in the United States and undertaken by Coleman *et al.* (1966) and Jencks *et al.* (1972). These two

innovative studies from two different disciplinary backgrounds – sociology and psychology, respectively – drew very similar conclusions in relation to the amount of variance in student outcomes that can be explained by educational factors. Although the studies did not suggest schooling was unimportant, the differences in student outcomes that were attributable to attending one school rather than another were modest. However, these studies were also criticized for failing to measure the educational variables that were of the most relevance (Madaus *et al.* 1979). Nevertheless, it is important to note that these two studies both claimed that, after taking into consideration the influence of student background characteristics such as ability and family background (for example, race and SES), only a small proportion of the variation in student achievement could be attributed to the school or educational factors. This pessimistic sense of not knowing what, if anything, education could contribute to reducing inequality in educational outcomes and in society as a whole was also fed by the apparent failure of large-scale educational compensatory programmes, such as 'Headstart' and 'Follow Through' conducted in the United States, which were based on the idea that education in pre-school/schools would help compensate for initial differences between students. Similarly disappointing results have since also been reported for the effects of compensatory programmes that have been conducted in other countries (Driessen and Mulder 1999; MacDonald 1991; Schon 1971; Taggart and Sammons 1999; Sammons *et al.* 2003).

The first two school effectiveness studies that were independently undertaken by Edmonds (1979) in the United States and Rutter *et al.* (1979) in England during the 1970s were concerned with examining evidence and making an argument about the potential power of schooling to make a difference in the life chances of students. This was an optimistic point of view, because many studies published in that period had shown that teachers, schools and maybe even education in general had failed to make much of a difference. The early appearance of these two independent research projects in different countries that asked similar questions and drew, to a certain extent, on similar quantitative methodologies demonstrated the potential for establishing a scientific domain dealing with effectiveness in education (Kyriakides 2006). Thus, the publications by Brookover *et al.* (1979) and Rutter *et al.* (1979) were followed by numerous studies in different countries on educational effectiveness and the development of international interest and collaboration through the creation of the International Congress for School Effectiveness and Improvement (ICSEI) in 1990 (Teddlie and Reynolds 2000). Looking at the history of EER, we see four sequential phases in the field, which address different types of research questions and promote the theoretical development of EER.

- **First phase: a focus on the size of school effects. Establishing the field by showing that 'school matters'.**
  During the early 1980s, the studies that were conducted attempted to show that there were differences in the impact that particular teachers and schools

have on student outcomes. This research showed how important it is for students to have effective teachers and schools, and that school and teacher effects tend to be larger for disadvantaged groups (Scheerens and Bosker 1997).

- **Second phase: a focus on the characteristics/correlates of effectiveness. Searching for factors associated with better student outcomes.**
  In the late 1980s and early 1990s, researchers in the area of EER were mainly concerned with identifying factors that were associated with student outcomes. These studies resulted in a list of factors that were treated as characteristics of effective teachers and schools (Levine and Lezotte 1990; Sammons *et al.* 1995; Scheerens and Bosker 1997).
- **Third phase: modelling educational effectiveness. The development of theoretical models that show why specific factors are important in explaining variation in student outcomes.**
  By the late 1990s and early 2000s several integrated models of educational effectiveness (Creemers 1994; Scheerens 1992; Stringfield and Slavin 1992) had been developed. These models sought to explain why factors that operate at different levels are associated with student outcomes, and these models guided not only the theoretical development of EER but also the design of empirical studies within this field (Kyriakides *et al.* 2000; de Jong *et al.* 2004).
- **Fourth phase: focus on complexity. A more detailed analysis of the complex nature of educational effectiveness that developed further links with the study of school improvement.**
  A gradual movement from the third to fourth phase was observed particularly after 2000. This featured a focus on change over time and addressed issues such as consistency, stability, differential effectiveness and departmental effects. Researchers increasingly gave attention to the study of complexity in education and pointed to the fact that the theoretical models of the third phase had not emphasized the *dynamic* perspective of education nor had they paid sufficient attention to the differential character of some factors (Creemers and Kyriakides 2006). Moreover, this gradual movement also saw an interest develop in investigating the question of changes in the effectiveness of schools, rather than exploring the extent of stability in effectiveness (Kyriakides and Creemers 2009). The move away from seeing effectiveness as an essentially stable characteristic of different schools or teachers to one that varies across years, and may differ for different student outcomes or in relation to different student groups, places change at the heart of EER. As such, the field became increasingly linked with the growth of larger scale, systematic investigations of the long-term effect of teachers and schools (Kyriakides *et al.* 2009; Pustjens *et al.* 2004). As a consequence, EER is seeing a growth in interest concerning the processes of school improvement, which is leading to the use of new theoreties such as the dynamic model (Creemers and Kyriakides 2008). Such developments

also point to the value of building links with other research areas such as organizational change in educational administration.

Although these four phases are analysed in more detail below, it is argued that EER has gradually evolved from focusing on a single issue (that is, schools and education 'matter') into a more intellectually sophisticated and mature area within the educational sciences. This has come about through EER trying to explain the complex and essentially dynamic nature of educational effectiveness and educational change. In this book, it is stressed that the above theoretical developments in the field were greatly facilitated by developments in research methodology that were increasingly adopted by EER in order to understand the complexity of the links between educational processes and outcomes. As a result, during these four phases one can also observe further parallel development in the methods used within EER. For example, there has been a movement from outlier studies (during phase one) to cohort studies (phases two and three) and most recently to longitudinal and experimental studies (phases three and four). Advanced techniques in analysing the data of effectiveness studies can be especially identified after the mid-1980s when multilevel modelling techniques (Goldstein 1995) were systematically employed. Therefore, it can be assumed that further developments in the field will continue to involve close links between theoretical and methodological advances.

First phase: establishing the field by showing that school matters

The studies from the first phase of EER were mainly concerned with revealing that teachers and schools differ among themselves in their impact on student performance and thereby that student progress in part depends on who is their teacher and which school they attend. The extent to which schools differ was the next question raised by researchers in the field, with a more precise version of this question being how much schools differ in terms of student outcomes when they are more or less equal in terms of the innate abilities and socio-economic background of their students (using statistical controls for variations in student intake characteristics). EER aimed to make fair comparisons between teachers and schools in order to assess the impact of schooling on student achievement that could be uniquely attributed to, for example, teacher A or school X, rather teacher B or school Y. Such research was enabled through the use of multilevel models that allowed for clustering of the data at the teacher (class) level and at the school level, which enabled more precise estimates of variation between schools and the identification of individual school effects (through residual estimates and their associated confidence limits). By the end of this phase, a clear message about the important role of teachers and schools had emerged from a large number of studies conducted in various countries and

these provided a strong argument against critics who had argued that teachers and schools did not matter for student outcomes (Scheerens and Bosker 1997; Teddlie and Reynolds 2000). However, the issue of educational effectiveness did not end by assessing the differences between schools and teachers in terms of their effectiveness. Rather, this was merely a prelude to exploring what matters in schools.

Second phase: searching for factors associated with student outcomes

The main research question of the second phase of EER attempted to identify those factors that help to explain differences in the effectiveness of schools. The results of studies conducted during this phase produced lists of correlates that were associated with better student achievement and which were treated as key effectiveness factors. One of the first of these was concerned with the so-called 'five-factor model' (Edmonds 1979). These five correlates of educational achievements were:

- strong educational leadership;
- high expectations of student achievement;
- an emphasis on basic skills;
- a safe and orderly climate;
- frequent evaluation of student progress.

This initial model has since been criticized on methodological (Ralph and Fennessey 1983) and also conceptual grounds (Scheerens and Creemers 1989). However, more refined models of educational effectiveness were also developed from this (Clauset and Gaynor 1982; Duckworth 1983; Ellett and Walberg 1979; Glasman and Biniaminov 1981; Murphy *et al.* 1982; Schmuck 1980; Stringfield and Slavin 1992; Squires *et al.* 1983). These later models elaborated on the framework for a causal model of educational effectiveness as developed by Scheerens and Creemers (1989). This framework stressed the fact that various levels in education can be seen to contribute to variations in student performance. The characteristics for educational effectiveness that are found in this phase of research can also be placed (that is, be seen to operate) at different levels. However, this framework does not answer why certain characteristics correlate positively with achievement. Finally, it is also important to note that reviews of the results of the studies conducted during this phase (Levine and Lezzotte 1990; Sammons *et al.* 1995) resulted in numerous correlates for effective classrooms, schools and above-school levels (districts, states, country). Together, these studies emphasized once more the importance of further developing the relatively limited theoretical foundation of EER by including the combination of correlates into categories.

Third Phase: development of theoretical models

The third phase of EER saw researchers use several theoretical orientations to help explain why certain characteristics might contribute to educational effectiveness (Scheerens and Bosker 1997). Generally speaking, there are three perspectives within EER in this phase, which attempted to explain why and how certain characteristics contribute to educational effectiveness, and three relevant theoretical models that emerged from these approaches.

First, in order to explain variation in the effectiveness of teachers and schools, economists have focused on variables concerned with resource inputs, such as per-student expenditure. Specifically, this economic approach is focused on producing a mathematical function that reveals the relationship between the 'supply of selected purchased schooling inputs and educational outcomes controlling for the influence of various background features' (Monk 1992: 308). This function may be viewed as either linear, consisting of main effects and interaction effects, or nonlinear (Brown and Saks 1986). Thus, the associated emergence of 'education production' models (Brown and Saks 1986; Elberts and Stone 1988) were based on the assumption that increased inputs will lead to increments in outcomes. These models are mainly concerned with: (a) the selection of relevant resource inputs as the major type of selection of antecedent condition, (b) the measurement of direct effects, and (c) the use of data at only one level of aggregation (that is, either at micro [for example, student] level or aggregated [for example, school] level).

The second model to emerge from this phase of EER featured a sociological perspective and focused on factors that define the educational and family background of students, such as SES, ethnic group, gender, social-capital and peer group. This perspective examined not only student outcomes but also the extent to which schools manage to ameliorate or increase the variation in student outcomes when compared to prior achievement. Two dimensions of measuring school effectiveness emerged from this perspective and concerned the quality of schools (students reaching high outcomes) and enhancing the equity in schools (reducing the achievement gaps between advantaged and disadvantaged groups). Moreover, the sociological perspective also brought attention to school processes that emerged from organizational theories (including climate, culture and structure) and to contexts such as the concentration of disadvantaged students and the impacts of this on student outcomes and school and classroom processes.

Finally, educational psychologists in this period focused on student background factors such as 'learning aptitude' and 'motivation', and on variables measuring the learning processes that take place in classrooms. Further, an interest in identifying and understanding the features of effective instruction practice was also observed and led to a list of teacher behaviours that were positively and consistently correlated with student achievement over time. For example, Rosenshine (1983) identified general teacher factors associated with achievement, which he labelled the 'direct instruction model' of teaching, sometimes called a 'structured approach'. From this, a slightly different model

called 'active teaching', with more emphasis put on the involvement of students in the learning and teaching process, was then also developed.

However, more recent research on teaching has featured a gradual trend toward less interest in teacher behaviour and the effects of teacher and instructional behaviour, and more interest in teacher cognition and teacher thinking about their professional practice (Creemers 2008). Within EER at this time, attention was initially directed to the effects of schools; however, after the introduction of multilevel analysis and a more theoretical orientation of EER, more emphasis was put on the learning and instructional level (Teddlie and Reynolds 2000). Theoretically, it was expected that student outcomes were related to learning activities that take place mostly at the learning/instructional level. This resulted in a re-orientation, empirically and theoretically, of effectiveness research toward a more explicit focus on the processes taking place at the teaching/learning level. Factors at the classroom level or the teaching and learning level are therefore seen as the primary effectiveness factors (Creemers and Kyriakides 2008). When a better foundation for EER was sought, this was therefore also concerned with an orientation towards developing theories and models about learning in schools. These theories and models were seen as a possible bridge between learning outcomes, which are used as criteria for effectiveness, and processes at the classroom and school level.

Fourth phase: analysing in more detail the complex nature of educational effectiveness

During the fourth phase, researchers attempted to respond to a major criticism of earlier EER that was concerned with the failure of the field to contribute significantly to the establishment of strong links between research on effective factors and developmental work intended to improve the quality of education. However, a dynamic perspective on education is now being taken into account more explicitly in theoretical and empirical EER studies (Creemers and Kyriakides 2006). Thus, in this phase, teaching and learning are seen as dynamic processes that are constantly adapting to changing needs and opportunities. This thereby has seen studies investigating the process of change in schools gradually emerge (Opdenakker and Van Damme 2006), which has had implications for modelling educational effectiveness in a way that takes into account the complex nature of education (Creemers and Kyriakides 2008). Moreover, such studies have helped us look at the functioning of each effectiveness factor using a dynamic rather than an instrumental perspective. This also implies that specific developments in the methodology of EER are needed since this research does not support the traditional approach of modelling effectiveness as a static feature or attribute of schools or teachers. Modelling effectiveness or change should not be restricted to fitting conditional models in which measures of student learning outcomes (adjusted for background characteristics such as SES, gender and prior knowledge) are regressed on a set of explanatory variables. Further, measures of

change based on only one or two time points are also unreliable (Bryk and Raudenbush 1987; Goldstein 1997; Willet 1988) and so provide an inadequate basis for studying change (Bryk and Weisberg 1977; Rogosa *et al.* 1982).

Therefore, the approach promoted in the current phase of EER does not place undue emphasis on measuring the short-term outcomes of the immediate effect of schools and teachers on student achievement gains during a single school year. In contrast, this approach instead reveals a need for longitudinal research to study results of schools and classrooms and their functioning over a longer period, modelling the growth in student outcomes over at least three time points. The progress made in the way educational effectiveness is conceptualized by EER also reveals that theoretical developments in the field have been facilitated by several methodological developments, as is shown in the next section.

Methodological developments promoting theory and research on effectiveness

This section attempts to show that much of the progress in EER is due to advances in methodology made during the last 30 years. During the first phase of EER, major emphasis was given to conducting outlier studies and comparing the characteristics of more effective schools with those of less effective schools (for example, the Louisiana Study). However, the emphasis on identifying outliers was criticized both for conceptual and methodological reasons (Goldstein 1997). During the 1980s, researchers moved to the use of mainly cohort and longitudinal designs involving larger numbers of schools and students, and such studies multiplied in the 1990s. In addition, the development of hierarchical regression approaches involving multilevel modelling techniques enabled researchers in the area to take the multilevel structure of educational systems into consideration and thereby deal with the methodological weaknesses of earlier studies that used only uni-level regression analysis (for example, the *Fifteen thousand hours* study by Rutter *et al.* 1979). Early examples of studies that made use of hierarchical regression approaches include the *School matters* study (Mortimore *et al.* 1988) and the *Young children at school in the inner city* research (Tizard *et al.* 1988) in England. Gradually, the way of measuring the links between inputs, outcomes and processes became more sophisticated, particularly through the development of contextual value-added models that controlled for student level, prior attainment and background as well as contextual measures of school or class composition (Sammons *et al.* 1997). Moreover, progress in the area of developing tests and other assessments helped researchers establish better measures of achievement not only in the area of basic skills such as reading, writing and mathematics but also in higher-level cognitive outcomes, and even in the affective and social domains.

During the second and third phase of EER, a large number of reviews were conducted and their main purpose was to provide the research community and policymakers with the latest developments from the field (Creemers and Reezigt

1996; Levine and Lezotte 1990; Sammons *et al.* 1995). However, although these reviews were usually based on a collection of studies that were seen by the authors as providing good examples of research (Creemers and Reezigt 1996, Sammons *et al.* 1995), their judgements of the methodological deficiency of studies that were excluded were not necessarily systematic and were often based on implicit rather than explicit criteria. On the other hand, the reviews that were not selective resulted in a huge number of factors for which not much information about their empirical support was provided (Levine and Lezotte 1990). As a consequence, the results of these early reviews can be questioned. In this context, carrying out meta-analyses using advanced quantitative approaches can be seen as a major methodological development that promotes theoretical development in the field and enables researchers to identify generic and more specific factors the impact of which is dependent on the educational setting in which they are operating (Scheerens and Bosker 1997).

During the third and the fourth phases of EER, emphasis was given not only to searching for predictors with direct effects on student outcomes but also those with indirect effects (for example in studies of school leadership and the links with student outcomes (Silins and Mulford 2002). Moreover, the theoretical models that have been developed during this phase refer to relations among factors situated at different levels (Creemers and Kyriakides 2008; Scheerens and Bosker 1997). In this context, the development of multilevel Structural Equation Modelling (SEM) approaches (Heck and Thomas 2000; Hox 2002; Muthén 1997) enables researchers to search for indirect effects and/or test the validity of the current models of EER in relation to this assumption (de Fraine *et al.* 2007; de Maeyer *et al.* 2007).

During the fourth phase of EER, emphasis has been given to modelling the dynamic nature of effectiveness. This implies, among other things, that longitudinal studies that last for at least three years should be conducted in order not only to measure the long-term effect of schools and teachers but also to find out how changes in the functioning of factors are associated with changes in educational effectiveness (Kyriakides and Creemers 2009). For this reason, developments in advanced quantitative research methods such as the use of growth modelling techniques are to be welcomed because they can help us answer such research questions. Moreover, conducting longitudinal studies enables researchers to search for reciprocal relations that, according to current theoretical developments in the field of EER, are expected to exist. Such relations are often included in relevant statistical models and refer to the relation of student factors that are likely to change with achievement (Creemers and Kyriakides 2008). For example, advanced SEM techniques can be used to search for reciprocal relations between motivation or academic self-concept and student achievement by making use of data collected at different points of time (Marsh *et al.* 2006).

At this point, we would like to claim that a challenge for EER is to make better use of current developments and progress in research methodology and to provide empirical support for new ways of conceptualizing the dynamic nature

of educational effectiveness. At the same time, it is acknowledged that the knowledge base of educational effectiveness and its attempt to establish theoretical models challenge the development of methodologically appropriate studies and ways of analysing data. An example is the development of SEM and multilevel techniques, which help EER to search not only for linear but also for non-linear relations between effectiveness factors and student achievement. Furthermore, a challenge for EER research methodology is to establish and use advanced quantitative techniques that will identify meaningful groupings of factors (Kyriakides *et al.* 2009). Such a development would contribute to the establishment of stronger links between EER and those who are involved in school/teacher development or improvement projects that seek to promote better practice (defined as that which promotes better outcomes for students) and offers the hope that more comprehensive strategies for school improvement will eventually emerge.

Aims and outline of the book

The aims of the book

The main aim of this book is to provide an authoritative account of the history and current developments in the methodology of EER and the way EER has evolved and is being applied in research and evaluation. By doing so, we hope to promote the further development of theory and research in educational effectiveness, which in turn depends on the further development of research methodology. At the same time, this book shows that the knowledge base of educational effectiveness and its attempt to establish theoretical models offer a challenge to the development of methodologically appropriate studies, including better ways of analysing clustered data. Next to this more theoretical perspective of the book, there is also a very practical argument for the promotion of the appropriate use of advanced research techniques by researchers in the area of EER. We try to do so by giving a sufficient background for each method and cite examples of using each within the context of EER.

The nature and structure of the book

This book is organized in three parts, which feature a summary of the main points therein. In the last chapter, the main conclusions emerging from the book are then outlined and their implications for further methodological development are considered.

Part A presents an overview of the state of the art of educational effectiveness studies and pinpoints challenges for the development of EER and research methodology. Chapter 2 provides an overview of major methodological issues in EER. More specifically, this chapter refers to the main research questions prevalent in EER and how these are related to methodological issues that are

connected to both designing studies and using advanced techniques to analyse quantitative and qualitative data, including their integration in mixed-methods designs. Moreover, issues concerned with the development of appropriate instruments and the process of investigating their psychometric properties are also discussed. In Chapter 3 by contrast, we argue that searching for causality is an important issue within EER since it promotes the theoretical and methodological development of the field. We also discuss causality's meaning and how it can be addressed by looking at different orientations within the research methodology. In the last chapter of this part of the book, we discuss the impact of EER on the design and evaluation of reform policies and the establishment of strategies to improve practice. It is argued that EER can contribute to the development of theory-driven evaluation studies that will serve both policymakers and educational practice, as well as promoting further theoretical development of the field. Thus, in this chapter, we provide guidelines to readers on how to design theory-driven evaluation studies that will contribute to the establishment of an evidence-based approach in policymaking and a theory-driven approach to improving education.

In Part B, each chapter is designed to help the readers understand why each methodological orientation is important for EER, and the mathematical background of each technique is summarized to help readers understand its current and future applications. Specific examples of applying each orientation are offered to help readers design their own effectiveness studies using the methodological tools presented here. The sequence used to present the methodological tools of EER follows the main decisions that have to be taken in designing original studies. Initially, different types of original studies are presented before we then refer to the two main measurement theories that can be used to develop research instruments and test their validity. Next, advanced statistical techniques that can be used to analyse nested and longitudinal data are presented before, finally, we then promote the use of meta-analyses and secondary analyses of international studies and explain how to conduct them. Specifically, in Chapters 5, 6 and 7, we illustrate the importance of using different types of research design in EER (that is, longitudinal studies, experimental studies and mixed research methods, respectively). It is shown that each research method can address specific research questions in an appropriate way. Chapters 8, 9 and 10 refer to two different measurement theories and show how Item Response Theory and Generalizability Theory can be used for developing and testing the validity of psychometrically appropriate research instruments. The next two chapters (11 and 12) refer to advanced techniques in analysing data. Chapter 11 examines the use of multilevel modelling techniques and their application, whereas Structural Equation Modelling techniques are discussed in Chapter 12. Finally, Chapter 13 examines the importance of conducting quantitative syntheses of original studies. The importance of testing and developing theoretical models of educational effectiveness by conducting meta-analyses is also stressed. Beyond illustrating how meta-analyses can be conducted to search for generic and

differential factors of effectiveness, we also refer to the importance of conducting secondary analyses of comparative international studies such as the Programme for International Student Assessment (PISA) and the Third International Mathematics and Science Study (TIMSS). Advantages and limitations of these approaches are also discussed.

In the very last part of the book, we relate the further development of theory and research in educational effectiveness to current trends and advances in the methodology used within the social sciences. Topics for further development concerning both EER and research methodology that are adopted in such studies are identified and a conceptual map for further methodological advancements in EER is provided.

References

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

Brown, B.W. and Saks, D.H. (1986) 'Measuring the effects of instructional time on student learning: evidence from the beginning teacher evaluation study', *American Journal of Education*, 94: 480–500.

Bryk, A.S. and Raudenbush, S.W. (1987) 'Application of hierarchical linear models to assessing change', *Psychological Bulletin*, 101(1): 147–58.

Bryk, A.S. and Weisberg, H.I. (1977) 'Use of the nonequivalent control group design when subjects are growing', *Psychological Bulletin*, 84: 950–62.

Clauset, K.H. and Gaynor, A.K. (1982) 'A systems perspective on effective schools', *Educational Leadership*, 40(3): 54–9.

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

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

Creemers, B.P.M. (2008) 'The AERA handbooks of research on teaching: Implications for educational effectiveness research', *School Effectiveness and School Improvement*, 19(4): 473–7.

Creemers, B.P.M. and Kyriakides, L. (2006) 'A critical analysis of the current approaches to modelling educational effectiveness: the importance of establishing a dynamic model', *School Effectiveness and School Improvement*, 17(3): 347–66.

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

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

de Fraine, B., Van Damme, J. and Onghena, P. (2007) 'A longitudinal analysis of gender differences in academic self-concept and language achievement: A multivariate multilevel latent growth approach', *Contemporary Educational Psychology*, 32(1): 132–50.

De Jong, R., Westerhof, K.J. and Kruiter, J.H. (2004) 'Empirical evidence of a comprehensive model of school effectiveness: A multilevel study in mathematics in the 1st year of junior general education in the Netherlands', *School Effectiveness and School Improvement*, 15(1): 3–31.

De Maeyer, S., Rymenans, R., Van Petegem, P., van den Bergh, H. and Rijlaarsdam, G. (2007) 'Educational leadership and pupil achievement: The choice of a valid conceptual model to test effects in school effectiveness research', *School Effectiveness and School Improvement*, 18(2): 125–45.

Driessen, G.W.J.M. and Mulder, L.W.J. (1999) 'The enhancement of the educational opportunities of disadvantaged children', in R.J. Bosker, B.P.M. Creemers and S. Stringfield (eds) *Enhancing excellence, equity and efficiency: Evidence from evaluations of systems and schools in change*, Dordrecht: Kluwer Academic Publishers, pp. 37–64.

Duckworth, K. (1983) *Specifying determinants of teacher and principal work*, Eugene, OR: Centre for Educational Policy and Management, University of Oregon.

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

Elberts, R.W. and Stone, J.A. (1988) 'Student achievement in public schools: Do principles make a difference?', *Economics Education Review*, 7: 291–9.

Ellett, C.D. and Walberg, H.J. (1979) 'Principal competency, environment and outcomes', in H.J. Walberg (ed.) *Educational environment and effects* 140–67, Berkeley, CA: McCutchan.

Glasman, N.S. and Biniaminov, I. (1981) 'Input–output analyses of schools', *Review of Educational Research*, 51: 509–39.

Goldstein, H. (1995) *Multilevel statistical models,* London: Edward Arnold.

Goldstein, H. (1997) 'The methodology of school effectiveness research', *School Effectiveness and School Improvement*, 8(4): 369–95.

Heck, R.H. and Thomas, S.L. (2000) *An introduction to multilevel modelling techniques,* Mahwah, NJ: Lawrence Erlbaum Associates.

Hox, J. (2002) *Multilevel analysis. Techniques and applications,* Mahwah, NJ/London: Lawrence Earlbaum Associates.

Jencks, C., Smith, M., Acland, H., Bane, M.J., Cohen, D., Gintis, H., Heyns, B. and Michelson, S. (1972) *Inequality: A reassessment of the effects of family and schooling in America,* New York: Basic Books.

Kyriakides, L. (2005) 'Extending the comprehensive model of educational effectiveness by an empirical investigation', *School Effectiveness and School Improvement*, 16(2): 103–52.

Kyriakides, L. (2006) 'Using international comparative studies to develop the theoretical framework of educational effectiveness research: A secondary analysis of TIMSS 1999 data', *Educational Research and Evaluation*, 12(6): 513–34.

Kyriakides, L., Antoniou, P. and Maltezou E. (2009) 'Investigating the short- and long-term effects of secondary schools upon academic success and development', paper presented at the 90th Annual Meeting of the American Educational Research Association, San Diego, USA.

Kyriakides, L., Campbell, R.J. and Gagatsis, A. (2000) 'The significance of the classroom effect in primary schools: An application of Creemers' comprehensive model of educational effectiveness', *School Effectiveness and School Improvement*, 11(4): 501–29.

Kyriakides, L. and Creemers, B.P.M. (2009) 'Explaining stability and changes in schools: A follow-up study testing the validity of the dynamic model', paper presented at the EARLI conference, Amsterdam, The Netherlands.

Kyriakides, L., Creemers, B.P.M. and 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.

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

MacDonald, B. (1991) 'Critical introduction from innovation to reform – a framework for analysing change', in J. Rudduck (ed.) *Innovation and change: Developing involvement and understanding,* Milton Keynes: Open University Press.

Madaus, G.G., Kellagham, T., Rakow, E.A. and King, D. (1979) 'The sensitivity of measures of school effectiveness', *Harvard Educational Review*, 4: 207–30.

Marsh, H.W., Wen, Z. and Hau, K.-T. (2006) 'Structural equation models of latent interaction and quadratic effects', in G. R. Hancock and R.O. Mueller (eds) *Structural equation modelling: A second course,* Greenwich, CT: Information Age Publishing, pp. 225–65.

Monk, D.H. (1992) 'Education productivity research: An update and assessment of its role in education finance reform', *Educational Evaluation and Policy Analysis*, 14(4): 307–32.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D. and Ecob, R. (1988) *School matters: The junior years,* Somerset: Open Books.

Murphy, J.F., Weil, M., Hallinger, P. and Mitman, A. (1982) 'Academic press: Translating high expectations into school policies and classroom practices', *Educational Leadership*, 40(3): 22–6.

Muthén, B.O. (1997) 'Latent growth modelling with longitudinal and multilevel data', in A. E. Raftery (ed.) *Sociological methodology*, Boston: Blackwell, pp. 453–80.

Opdenakker, M.C. and Van Damme, J. (2000) 'Effects of schools, teaching staff and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes', *School Effectiveness and School Improvement*, 11(2): 65–196.

Opdenakker, M.C. and Van Damme, J. (2006) 'Differences between secondary schools: A study about school context, group composition, school practice, and school effects with special attention to public and Catholic schools and types of schools', *School Effectiveness and School Improvement*, 17(1): 87–117.

Pustjens, H., Van de Gaer, E., Van Damme, J. and Onghena, P. (2004) 'Effect of secondary schools on academic choices and on success in higher education', *School Effectiveness and School Improvement*, 15(3–4): 281–311.

Ralph, J.H. and Fennessey, J. (1983) 'Science or reform: Some questions about the effective schools model', *Phi Delta Kappan*, 64(10): 689–94.

Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C. and Schaffer, G. (eds) (2002) *World class schools: International perspectives on school effectiveness*, London: RoutledgeFalmer.

Rogosa, D.R., Brand, D. and Zimowski, M. (1982) 'A growth curve approach to the measurement of change', *Psychological Bulletin,* 90: 726–48.

Rosenshine, B. (1983) 'Teaching functions in instructional programs', *The Elementary School Journal*, 83(4): 335–51.

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

Sammons, P., Hillman, J. and Mortimore, P. (1995) *Key characteristics of effective schools: A review of school effectiveness research,* London: Office for Standards in Education and Institute of Education.

Sammons, P., Power, S., Elliot, K., Campbell, C., Robertson, P. and Whitty, G. (2003) *New community schools in Scotland: Final report – national evaluation of the pilot phase,* Edinburgh: Scottish Executive Education Department.

Sammons, P., Thomas, S. and Mortimore, P. (1997) *Forging links: Effective schools and effective departments,* London: Paul Chapman.

Scheerens, J. (1992) *Effective schooling: research, theory and practice,* London: Cassell.

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

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

Schmuck, R.A. (1980) 'The school-organization', in J.H. McMillan (ed.) *The social psychology of school learning*, New York: Academic Press.

Schon, D.A. (1971) *Beyond the stable state,* Harmondsworth: Penguin.

Silins, N. and Mulford, B. (2002) 'Leadership and school results', in K. Leithwood and P. Hallinger (eds), *Second international handbook of educational leadership and administration*, Norwell, MA: Kluwer Aacademic Publishers, pp. 561–612.

Squires, D.A., Hewitt, W.G. and Segars, J.K. (1983) *Effective schools and classrooms: A research based perspective*, Alexandria, VA: Association for Supervision and Curriculum Development.

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

Taggart, B. and Sammons, P. (1999) 'Evaluating the impact of raising school standards initiative', in R.J. Bosker, B.P.M. Creemers and S. Stringfield (eds) *Enhancing educational excellence, equity and efficiency: Evidence from evaluations of systems and schools in change*, Dordrecht: Kluwer Academic Publishers, pp. 137–66.

Teddlie, C. (1994) 'The integration of classroom and school process data in school effectiveness research', in D. Reynolds, B. Creemers, P.S. Nesselrodt, E.C. Shaffer, S. Stringfield and C. Teddlie (eds) *Advances in school effectiveness research and practice*, Oxford: Pergamon, pp. 111–33.

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

Teddlie, C. and Stringfield, S. (1993) *Schools make a difference. Lessons learned from a 10 year study of school effects*, New York: Teachers College Press.

Tizard, B., Blatchford, P., Burke, J., Farquhar, C. and Plewis, I. (1988) *Young children at school in the inner city*, Hove: Lawrence Erlbaum.

Willett, J.B. (1988) 'Questions and answers in the measurement of change', *Review of Research in Education,* 15: 345–422.

Chapter 2

Methodological issues in Educational Effectiveness Research

Introduction

In this chapter, we analyse the current (fourth) phase of EER, examine the main research questions that form the foci of educational effectiveness enquiry and analyse methodological issues that should be taken into account when designing studies, as well as in analysing quantitative and qualitative data. These issues are presented to show the contribution research methodology has made and can make to the development of EER, before further clarification of the methodologies themselves and how they can be used is presented in the second part of the book. As mentioned in the introductory chapter, the current phase of EER is concerned with promoting a better understanding of the complex and dynamic nature of educational effectiveness. Thus, this chapter is concerned with methodological issues that have important implications, especially for modelling effectiveness, and the next chapter discusses issues associated with the design of studies that search for causal relations. More specifically, causality is a general issue in EER, and almost all effectiveness studies deal in one way or another with hypothesized (or implied) cause-and-effect relations by searching for factors that explain (in a statistical sense) variation in student outcomes. Further, this is usually in terms of academic achievement tests or measures of some kind. However, EER also encounters frequent difficulties in claiming cause-and-effect relations due to the non-experimental nature of most EER designs. In the last chapter of the first part of this book we acknowledge that the emphasis given by the current phase of EER to the dynamic perspective of educational effectiveness draws attention to the importance of searching for predictors of the processes of school improvement. This implies that researchers should not restrict themselves when describing effective practices to only those features that can be observed in schools. Instead, EER should also contribute to the design of theory-driven improvement strategies and in developing measures of their impact on changes in the effectiveness of teachers and schools over time. Since this shift in the research agenda of EER raises specific methodological issues, it is discussed in more detail in Chapter 4, which is concerned with establishing links between EER and policy and practice.

The methodological issues that are presented in the next section of this chapter refer to the contribution of research methodology in modelling effectiveness. Within this, it is taken as essential that researchers should attempt to search for more complex relations (direct, indirect and reciprocal) between different student, classroom/teacher and school factors and change in student achievement or other outcomes. Further, it is also taken as necessary for researchers to explore potential relations among factors operating at the same or at different levels in order to describe the complex nature of educational effectiveness. As such, this section therefore discusses the methodological issues that are associated with the need to search for generic and differential factors. It is acknowledged that during the current phase of EER there is a need to make a distinction between generic and differential factors and identify those that are common across different educational settings and those that are differential or specific, where the size of their effect depends on the setting in which they are operating (Campbell *et al.* 2004; Kyriakides 2007). Thus, a methodological issue stemming from this is not only how differential effectiveness can be investigated but also how the results of such studies can be incorporated into our attempts to establish generic and more context-specific models. Finally, an argument is made that researchers should also attempt to conceptualize the dynamic nature of education, which includes the need to look at ongoing changes in the functioning of teachers and schools and how these are related to changes over time in their effectiveness.

The emphasis that is given to the dynamic perspective of educational effectiveness also implies that the stability of effects related to teachers and schools over time cannot be seen as a necessarily clear measure of the reliability of these concepts (as was assumed during the second and third phases of EER; Kyriakides and Creemers 2008a). Thus, this chapter also discusses methodological issues concerned with the research that should be conducted to search not only for short but also for long-term effects of teachers and schools. These issues are then compared with the processes that should be used in analysing data that is collected for these different types of studies.

The last two sections of this chapter refer to some ongoing methodological questions that have not yet been solved. Specifically, there is an initial discussion of the possibility of research methodology enabling researchers to use different outcomes of schooling as the criteria of effectiveness. Special emphasis is given not only to the valid measurement of these outcomes but also to how they relate to each other. This issue also brings to attention the importance of using appropriate instruments to collect data. Thus, the last section of this chapter is concerned with advances in measurement theory in order to improve the quality of instruments that are used to measure not only student outcomes but also the nature and functioning of different effectiveness factors.

The contribution of research methodology in modelling educational effectiveness

Modelling educational effectiveness was an important research issue during the last (third) and current phases of EER. Moreover, the increasingly important role of this modelling was already evident in the second phase since here it became clear that EER should look at both classroom and school factors and that thus only multilevel models were appropriate to describe educational effectiveness. Following on from this during the third phase of EER, a number of studies were conducted that tried to test the validity of the early theoretical models of EER (Driessen and Sleegers 2000; Kyriakides *et al.* 2000; Reezigt *et al.* 1999). These revealed a number of methodological issues that should be taken into account in developing the theoretical framework of EER. These issues are analysed below and were also partially products of studies that had searched for the effect of school factors and especially factors such as leadership (Leithwood and Jantzi 2006; de Maeyer *et al.* 2007; Robinson *et al.* 2008) and quantitative syntheses of both school (Witziers *et al.* 2003; Scheerens and Bosker 1997) and teacher effectiveness (Monk 1994; Seidel and Shavelson 2007). More specifically, these studies made clear that both the direct and indirect effects on student achievement of such factors should be examined.

Searching for direct and indirect effects of factors on student achievement

In order to search for the relationships between significant educational factors and student achievement, researchers have considered the use of multilevel SEM models. These allow the specification of cross-level relationships by employing multivariate multilevel modelling techniques, which allow the use of more than one dependent variable. For example, the testing of a relationship between a factor concerned with school policy on the quantity of teaching and the management of time by the teacher can be conducted by treating the school factor (that is, school policy on quantity of teaching) as an explanatory variable and treating both the classroom factor (that is, management of time) and student achievement as simultaneous dependent variables.

Searching for relations between factors operating at the same level

The methodological procedure used in multilevel modelling techniques to search for cross-level relationships can also be used to search for relationships between factors that are operating at the same level. Obviously such relations can also be investigated through path analytic models. Searching for relations among factors operating at the same level is also an issue that needs further attention for the theoretical development of EER. This is not only due to different theoretical

models referring to particular groups of factors but also because research on instructional effectiveness has developed specific teaching approaches consisting of combinations of particular teaching factors (for example, direct instruction, active teaching, new learning) that are seen as more effective than others.

Searching for nonlinear relations between factors and student achievement

Meta-analyses of the effect of some effectiveness factors upon student achievement has revealed that although they have been conceived of as having an impact on teacher or school effectiveness, the research evidence for their role remains problematic. For example, teacher subject knowledge is widely perceived as a factor that affects teacher effectiveness (Scriven 1994), but teachers' subject knowledge, regardless of how it is measured, has only rarely been correlated with student achievement in practice (Borich 1992; Darling-Hammond 2000). The explanation may be, as Monk (1994) reported, that the relationship is curvilinear. In other words, a minimal level of knowledge is necessary for teachers to be effective, but beyond a certain point, there may be a negative relation. Similar findings have been reported for the association of self-efficacy beliefs with teacher effectiveness (Schunk 1991; Stevenson *et al.* 1993) and for the impact of classroom emotional climate and teacher management upon effectiveness. These findings imply that models of educational effectiveness should acknowledge that nonlinear relations might exist and therefore that a search is needed for the optimal values of factors that are nonlinearly related with achievement (Creemers and Kyriakides 2006).

The challenge of this for modelling educational effectiveness has implications for both the design and the analysis of effectiveness studies since the investigation of nonlinear relations implies that more complex statistical techniques should be used in analysing the data and that more emphasis should be given to ensuring the quality of the measures that are used.

As far as an analysis of data is concerned, two issues need attention. Since models of educational effectiveness refer to factors operating at different levels, it is important to use multilevel modelling techniques that are able to identify the variables at the student, teacher, school, and system levels that are associated with student outcomes of interest. However, as noted above, an issue that has to be taken into account is that some variables may not be linearly related with student achievement. In the case of education, we have already considered the likely existence of inverted-U curvilinear relations since these reveal that there is an optimal point for the impact of a specific factor. After the optimal point, a flattened or negative relation with achievement can exist, and thereby the identification of the optimal point has important implications for improving educational practice. Therefore, in analyses where researchers search for inverted-U relations, the effect of both the various explanatory variables (X_i) and the

effect of the second power of these variables (that is, X_i^2 values) upon student achievement have to be identified. This approach may allow us to identify the optimal value of this predictor factor (that is, the values of X for which Y has a maximum value). Of course, other forms of nonlinear relations might also exist. In cases where more than one optimal point can be identified, a question can be raised about the efficiency of application of a certain factor after the first optimal point is reached.

The second issue needing further attention is the issue of measurement errors of variables, because these act as significant obstacles in establishing the existence of possible nonlinear relations. In turn, this implies that researchers should give more emphasis to measurement issues in order to increase the quality of the data collected. Because the investigation of nonlinear relations is based on searching to see whether the second (or even higher) power of a factor is able to explain any additional variation in student achievement, it is important to reduce measurement error as this is magnified in power calculations. In this case, instead of dealing with the measurement error of a factor, you have to deal with this error raised to the relevant power. As a consequence, it is much more difficult to identify statistically significant relations. By giving more emphasis to measurement issues, both the quality of the data collected and their statistical power (that is, reducing the type II error) would be increased. Thus, when a particular study fails to evidence nonlinear relations, this does not necessarily imply that the relationships are truly linear because the finding may instead simply be an artefact of the relatively high measurement error of the particular effectiveness factor in question.

Finally, a failure to demonstrate nonlinear relations also may be attributed to the difficulties of establishing enough variation in the functioning of some factors, especially since almost all the effectiveness studies have been conducted in a single country. Primarily, there are two alternative approaches in the search for nonlinear relations. First, experimental studies can be conducted to create enough variance in the functioning of each factor before then searching for optimal values. However, research on the impact of changes in class sizes reveals that there may be practical and ethical difficulties in attempts that concern the manipulation of school or classroom conditions. Thus attention should be given to the ecological validity of experimental studies as well as to the ethical issues associated with the experimentation (Miller 1984; Robson 1993). On the other hand, comparative studies can be conducted that allow the validity of EER models to be tested, especially when searching for the possible existence of nonlinear relations. For example, international longitudinal studies are more likely to tap the full range of variation in school and classroom quality measures and therefore also variation in potential school and classroom effects. Thus, these studies could help us identify nonlinear relations since within national studies the lack of a significant effect might be due to difficulties in identifying enough variation in either the student outcomes or, more likely, in the explanatory variables studied.