CHAPTER 3

ALTERNATIVE MODELS OF SCHOOL EFFECTIVENESS PUT TO THE TEST

ROEL J. BOSKER and JAAP SCHEERENS*

University of Twente, Division of Educational Administration, Faculty of Educational Science and Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands

Abstract

The term "educational effectiveness" designates causal models of educational outcomes that may or may not contain school-level variables. The term "school effectiveness models" is used in the more restricted sense of outcome-oriented models that explicitly contain school-level variables. These models are categorized to a context-input-process-output structure, are multi-level, recognize causal chains, and sometimes include feedback loops. In this chapter the advantages of comprehensive vs. partial school effectiveness models are discussed. Furthermore alternative interpretations of the idea that higher-level variables some way facilitate lower-level variables in the production of educational outcomes are given. The debate concerns: (i) additive vs. interactive models, (ii) direct vs. indirect models, (iii) contextual vs. genuine effects, (iv) additive vs. synergetic interpretations, and (v) recursive vs. non-recursive models. In this chapter the results of empirical research are presented to test the first four completing interpretations. The results indicate that the indirect model is the least productive. The synergetic model looks promising, but more simple interactive models account for almost the same percentage of variance in students outcomes. The actual problem might be that the most likely school effectiveness models are indirect, non-recursive, contextual, synergetic, and interactive at the same time. To avoid complexities in future research decompositions are advocated. And therefore it is contended that organizational variables of interest first be examined on their possible alternative interpretation of cross-level facilitation, before formulating the model.

Introduction: The Emergent Global School Effectiveness Model

Despite the rather inductive nature of the various types of educational effectiveness research and the enduring criticism of a-theorism, quite a few school effectiveness models are to be found in the literature. The term "educational effectiveness" designates

*Authors listed in alphabetical order.
causal models of educational outcomes that may or may not contain school-level variables. The term "school effectiveness models" is used in the more restricted sense of outcome-oriented models that explicitly contain school-level variables. We shall adhere to this terminology although a strong case could be made for referring to these models by their real name: causal models of educational attainment. In actual fact, on closer inspection, several of the more recent "comprehensive school effectiveness models" (Scheerens, 1990; Stringfield & Slavin, 1992; Levine & Lezotte, 1990) are really more like conceptual maps, in which, broadly speaking, the most promising variables to have come out of educational effectiveness research are more or less ordered according to an input-process–output framework. When the interrelationships between the various categories of variables are specified in more detail (as, for instance, the models of De Corte & Lowyck, 1983; Reynolds & Walberg, 1990; Creemers, 1991) the ordered summaries of variables are more correctly referred to as models. The general characteristics of these school effectiveness models are the following:

1. (As was already stated in the above) variables are categorized according to an input-process–outcome and context structure;
2. the models incorporate a multi-level structure, usually at pupil, classroom and school-level, sometimes even extending to school-environment level;
3. the models also recognize causal chains, i.e., intermediate causal variables that reflect the influence of certain exogenous conditions while at the same time affecting certain other variables;
4. in some cases the models also include non-recursive relationships (feedback loops) implying self-regulating causal mechanisms.

The most appealing substantive connotation of these models is the image of the production of educational outcomes, where the school organization is seen as a system of nested layers (Barr & Dreeben, 1983).

It should be noted that the emphasis on the specification of particular categories of variables among the available models of this type varies widely according to the disciplinary background of the author(s). Educational economists, interested in educational production functions, focus on material inputs and processes allocating teachers and pupils to classes (Monk, 1992). Educational sociologists have a particular interest in contextual variables (e.g., classroom composition). Educational psychologists examine teacher characteristics and instructional learning conditions in great detail, while investigators that operate in the field of educational administration are most interested in organization and management conditions and the school’s environment.

Within the nested layers image of school functioning, the most pervasive view on the cross-level influences is the notion that, in the production of educational outcomes, higher level conditions somehow facilitate conditions at lower levels. Despite the common features of the available models, outlined above, there is still a great deal of uncertainty surrounding these models. We therefore prefer to speak of "an emergent global school effectiveness model". One source of uncertainty is the lack of consistency in the research findings that corroborate the models. Despite the differences in disciplinary background already mentioned there is some consensus, on the kind of input and process variables that are important, although some relationships have received considerably more empirical support than others (cf. Scheerens, 1990). The other source of uncertainty, which will be addressed in this chapter, concerns
the interpretation and formal specification of the cross-level interrelationships within the models. The notion of facilitation by higher level variables is still rather vague. In an attempt to clarify some of the globality and vagueness of current models, an analysis of alternative interpretations of this general idea, as well as alternative formal specifications, will be given. Formal modeling is also seen as an important means to bridge the gap between conceptual models and future empirical research. First, however, the question of the usefulness of comprehensive vs. more partial models of educational effectiveness will be briefly addressed.

Comprehensive vs. Partial Educational Effectiveness Models

The most efficient strategy for building multi-level school effectiveness models appears to be an "inside-out" approach, that is, to first model micro-processes that are closest to the actual learning processes of individual students and then to consider facilitating conditions one level up. This would mean that representations of learning and decision-processes by individual students should be at the core of educational effectiveness models.

The sensibility of such an approach can be demonstrated by considering developments in instructional psychology. Most current comprehensive school effectiveness models draw heavily on the well-known Carroll-model (Carroll, 1989), emphasizing instructional time, motivation to learn and quality of instruction, which in later elaborations was specified in terms of a structured approach and optimization of opportunity to learn (i.e., the correspondence between educational objectives and subject-matter actually taught).

More recent developments in instructional psychology emphasize procedural knowledge (e.g., learning to learn), a more "interactive" view of motivation to learn, and the importance of affective and social aspects in the learning situation (cf. Simons, 1990; Boekaerts, 1991; Pieters, 1992). Clearly this latter view on learning and instruction has important consequences for the type of teaching-repertoires and the use of instructional media that are thought to be conducive to learning.

In consideration of modeling teaching and learning behavior, the addition of conditions at higher levels of school functioning would make comprehensive models extremely complex both with respect to "internal processes" and behavior and interactions and feedback loops. The more so because school functioning at the level of classroom organization and school management can likewise be depicted as situationally-determined and full of interactions. In our opinion, three strategies can be used to deal with this seemingly prohibitive degree of complexity. The first strategy focuses on the essential conditions at each level, and calls for parsimony in the selection of variables and hypothetical interrelationships, while taking a certain roughness or globality for granted.

The second strategy gives up the idea of empirically testing comprehensive school effectiveness models altogether and, instead, concentrates on the development and exploration of partial educational effectiveness models. The third strategy consists of using and integrating detailed partial models in their full complexity and exploring the possibilities of systems-methodology and the techniques used in operations research.
The first strategy is the real issue for debate, bearing in mind that the second strategy is not so much an intentional strategy as a process that will go on anyway in various educational sub-disciplines, whilst the third is little more than an interesting methodological exercise. In our opinion there are two compelling reasons for continuing to work in line with the first strategy. The first argument is of an epistemological nature and rests upon the assumption that abstraction and generalization are inevitable and in fact major vehicles for the advance of scientific knowledge. Shying away from abstraction by referring to the complexities of everyday reality, leads to the trap of so-called “strong empiricism” (cf. Thomas, 1979). The second argument is, perhaps, of a more parochial “sub-discipline-tied” nature and is related to the idea of an outcome-oriented approach to the study of educational organization and management. Boyd and Crowson (1985) have criticized the educational administration research approach for its process-orientation. They consider a more outcome-oriented approach as an important aspect of what they think of as a paradigm-shift. When endorsing this latter perspective — which we do — it becomes virtually inevitable to consider intermediate processes — that is between organizational and managerial conditions on the one hand, and educational outcomes on the other. The success of working with comprehensive school effectiveness models that — of necessity — will have a certain globality and roughness — is strongly dependent on conceptual clarity and the suitability of the available techniques of analysis. Before turning to these themes in more detail, the importance of partial models for the development of more comprehensive global models should be underlined. In this respect the following fields of inquiry are thought to be of particular interest:

— ongoing work with respect to time spent on the task, learning opportunity and structured teaching (Stringfield & Slavin, 1992);
— developments in instructional psychology that draw upon cognitive psychology and motivational theory (cf. Boekaerts, 1991; Simons, 1990);
— contributions in the field of educational economics with respect to educational production functions, incentive structures and choice processes of educational “agents” (pupils, teachers, parents) (cf. Monk, 1992);
— motivational aspects of teachers’ functioning (cf. Fuller, Wood, Rapoport, & Dornbusch, 1982);
— the pivotal functional of evaluation within schools conceptualized as “learning organizations” (cf. Scheerens, 1992);
— work on school culture and educational leadership, pinpointing in particular a joint sense of achievement (cf. Creemers, 1991).

Key-Variables and Central Interrelationships within the Framework of the Nested Layers Image of School Functioning

It is beyond the scope of this chapter to give a detailed account of specific variables that should be considered within the framework of a comprehensive school effectiveness model. Nor will the important issue of measurement and data characteristics of these variables be dealt with. Instead, the focus will be on delineating the most important dimensions of variables. These dimensions might be thought of as composite indicators of the most important antecedent conditions of educational outcomes. As explained
earlier, the empirical basis for the selection of these dimensions of variables differs considerably (see Scheerens, 1990).

— At the pupil level *achievement scores* or *attainment indexes* (the outcome variables) need to be adjusted for background variables like *intelligence, cultural capital of the home* and *peer group characteristics*.

— *Student motivation to learn* should have a place within the model as a factor that is influenced by instructional variables, while at the same time being an antecedent condition of achievement and a factor that is influenced by feedback on achievement.

— The third dimension is located at classroom-level and a joint index of educational time and content covered; this dimension could be called *intensity and relevance of educational exposure*.

— The fourth dimension represents the technical quality of education. We shall refer to this dimension as *degree of structure and adaptiveness of instruction*, where we assume that both characteristics from mastery-learning type approaches and aspects of self-regulated learning and meta-cognition can be assembled within this one index.

— The fifth dimension represents *teacher quality* and could be thought of as a composite rating scale used by school leaders or inspectors.

— At school-level, the sixth dimension represents a *shared achievement-oriented mission* among staff and school-leadership.

— *Coordination among sub-units* represents a more structural factor thought to be conducive of consistency in effective classroom practices represented by several of the other dimensions.

— *Evaluation practices* (at all levels of school functioning) represents an eighth dimension which is of vital importance for the image of a school as a learning organization.

— *Quality of allocation of pupils and teachers* is a ninth dimension, which assumes that optimal combinations of teachers and groups of pupils exist.

— *Achievement incentives from the school environment* represents the tenth dimension (i.e., overt attainment targets at higher educational levels).

It should be noted that several dimensions have a meaning at more than one level of school functioning. For instance, learning time and content covered are shaped both at the level of the school curriculum and in the actual process of teaching and classroom management. Evaluation practices take place at school level (as for instance in so-called school based review) and at classroom level. At the same time individual dimensions may have several "bearers". Structure and adaptiveness, for instance, is both a characteristic of instructional media and of teacher behavior whilst achievement orientation depends on school principals and their staff. A third point worth noting with respect to this multi-level framework of variables, which is depicted in Figure 3.1, is that preferably it should be studied longitudinally, if only to have some indication of the stability (and thus: changeability) of effectiveness and effectiveness-enhancing conditions.

The three relationships of focal interest from the perspective of the nested layers image of schooling are:

— The environmental stimulation required to make schools operate effectively;
the facilitation of effective instruction by organizational arrangements (both cultural and structural);
— the guidance of learning processes by means of instruction.

Before turning to alternative interpretations of these cross-level relationships, two important features should be mentioned. First of all the magnitude of cross-level causation will tend to be small in comparison to “within level” causation. To give one example, variation in achievement between schools amounts to a maximum of only 20% of the total variation in achievement (between and within schools). If teacher autonomy is taken into consideration, the quality of teaching and classroom management will definitely depend more on teacher characteristics than on facilitating organizational conditions, and the same can be said of achievement-enhancing environmental stimuli in relatively autonomous schools. Secondly, the literature on micro economic educational production indicates that cross-level facilitation may be blocked or give rise to counter-productive results. For instance, to a certain extent, it will be up to the pupils to make effective use of an increase in the intensity and targetedness of educational exposure. Pupil motivation to fully use educational offerings may be seen in terms
Educational Effectiveness Research

The general notion of higher-level variables in some way facilitating lower-level variables in the production of educational outcomes, conceals rather divergent interpretations of these cross-level relationships.

First, relationships between conditions at higher and lower levels can take the shape of so-called contextual effects. This being the case, the aggregates of certain attributes defined at micro-level are seen as having an additional causal influence. For instance, Erbring and Young's (1979) model on endogeneous feedback predicts that, should a school have a majority of effective teachers and sufficient feedback among the staff, the minority of less effective teachers will be stimulated to become more effective. Other contextual effects that are, as a matter of course, important in educational organizations, have to do with classroom and peer group composition. To some extent contextual effects, acting as selection mechanisms, work against improving education by means of optimizing organization variables. On the other hand, it should be realized that selection and allocation decisions, subject to the distribution of authority within national educational systems, can be partly controlled by school management. Further on, we shall present a formal multi-level model from which the relative magnitude of contextual effects in comparison with genuine cross-level facilitating can be assessed.

Secondly, conditions at higher levels can act as mirrors to conditions at lower levels. As Berliner (1985, p. 143) observes: “The evidence on effective classrooms and effective schools is amazingly congruent”. Features such as achievement pressure, high expectations of pupils’ achievement, monitoring and an orderly atmosphere have been found relevant in studies of effective classrooms as well as effective schools. Congruence of factors that have meaning at different levels could be thought of as creating a consistent school culture. According to this interpretation higher level facilitation is more likely to provide a general supportive background than a purposeful and direct manipulation of lower level conditions, which is the interpretation we shall consider next.

Third, higher levels can be thought of as overt measures creating effectiveness-enhancing conditions at lower levels. The concept of instructional leadership comes under this heading, as do increasing the allocated learning time, the recruitment of “effective” teachers, the selection of teaching materials that have effectiveness-enhancing characteristics, stimulating evaluations at classroom level, keeping records of pupils’ progress, and many more.

Fourth, conditions at higher levels can serve as incentives to promote efficiency-enhancing conditions at lower levels. Higher-level conditions include rewards for teachers from their superiors for effective teaching and monetary grants from their district if schools reach certain achievement standards. Educational consumer-demand for effective schooling is another type of external incentive. In fact this view of
cross-level facilitation reflects a somewhat restricted input-output view, where school processes are treated as a black box and where standards or attainment targets are the main category of inputs.

In the fifth place, conditions at higher levels can serve as material facilities for conditions at lower levels (a more restricted case of the second "mirror" category). One example is a computerized, school-monitoring system implemented at school level to provide teachers with more sophisticated means of monitoring student progress.

Sixth, and finally, higher level conditions may serve as buffers to protect efficiency-enhancing conditions at lower levels. These high-level conditions include "maintenance" functions carried out by school directors like safe-guarding student-enrollment figures, dealing with extramural pressures, and coping with governmental regulations, and representational activities. This view implies minimal expectations of the direct influence of school management on the actual education production process, which in this scenario would be left entirely to the autonomous teachers.

In the next section some of these alternative interpretations will be formulated more precisely in terms of alternative specifications of multi-level or structural models. In a similar way, some further intricacies of the interrelationships of the nested layers image of school functioning will be described in terms of alternative specifications of formal models.

Alternative Causal Specifications

When turning to the question of alternative causal specifications within a global framework of schools as nested layers, the following competing models are distinguished:

(a) Additive vs. interactive models

According to additive models higher-level conditions are seen as increments to variables operating at the lower level; e.g., achievement-oriented policy at an administrative level higher than the school "adds to" the effects of achievement-oriented policy at school level. In the interactive models, higher-level conditions impinge on the (causally interpreted) relationship between lower-level antecedent conditions and the criterion variable; for instance, when instruction at classroom level is thought of as determining the impact of the ability and effort put into their work by individual learners. In terms of multi-level modeling, the comparison of these two interpretations involves an interest in comparing intercepts (additive model) vs. an interest in comparing slopes (interactive model).

The additive model can be written as:

\[ Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + e_{ij} \quad (3.1a) \]

\[ \beta_{0j} = \beta_{00} + \gamma Z_j + u_{0j} \quad (3.1b) \]

in which \( Y_{ij} \) represents the achievement score of pupil \( i \) in school \( j \); \( \beta_{0j} \) is the school-specific intercept; \( X_{ij} \) represents the ability of pupil \( i \) in school \( j \); \( \beta_1 \) is the regression coefficient and \( e_{ij} \) is the pupil-level error term. In Equation (3.1b) the school-specific intercept, which can be interpreted as the mean school achievement score corrected for
ability, is modeled as a function of the Grand Mean $\beta_{00}$, a school level variable $Z_j$, with $\gamma$ being the regression coefficient, and $u_{ij}$ the school level error term. The interactive model is somewhat different:

$$Y_{ij} = \beta_1 + \beta_{1j}X_{ij} + e_{ij} \quad (3.2a)$$
$$\beta_{1j} = \beta_{10} + \gamma Z_j + u_{1j} \quad (3.2b)$$

The Equations (3.2a) and (3.2b) differ from the first set of equations in that no school-specific intercepts are specified, but instead school-specific regression coefficients $\beta_{ij}$ for the regression of achievement ($Y_{ij}$) on ability ($X_{ij}$) are incorporated in the model. These school-specific regression coefficients are then modeled as a function of an overall regression coefficient $\beta_{10}$, a school variable $Z_j$, with regression coefficient $\gamma$, and a school-level error term $u_{1j}$, which expresses the school-specific deviation from the overall regression of achievement on ability. Combining Equations (3.2a) and (3.2b) by mere substitution shows that a term $\gamma Z_j X_{ij}$ is included in the model, the cross-level interaction:

$$Y_{ij} = \beta_0 + \beta_{10}X_{ij} + \gamma Z_j X_{ij} + X_{ij}u_{ij} + e_{ij} \quad (3.2c)$$

In Figures 3.2 and 3.3 a graphic representation of the models is given.

Examples of tests of additive vs. interactive multi-level models of educational effectiveness are presented by Gamoran (1991) and Bosker, Kremers, and Lugthart (1990). Gamoran found that, in general, the additive model might be preferred when studying educational effects. Content coverage and the quality of instructional discourse were instructional factors that added to achievement. Nevertheless, he also found evidence that initially high-achieving pupils especially benefited from the amount of curricular content covered. Bosker, Kremers, and Lugthart did research on organizational level conditions for effective instruction. They came up with the conclusion that effective instruction (measured as actual instruction time) is related
to achievement only in those cases where the degree of formalization of school-rules is below average. Instruction effects on achievement, although present in an additive sense, are, to a certain degree, contingent on school organization. Moreover, girls especially benefited from effective instruction. Aitkin and Zuzovsky (1992), calling for a new paradigm in educational research, argue that all educational models should be interactive until proven to be empirically invalid. Their case is strong, since the additive model is a special case of the interactive model. This observation also leads to the conclusion that the additive model is scientifically more powerful, since it is more parsimonious than the interactive model.

(b) Contextual vs. "genuine" multi-level effects

A basic challenge to the nested layers approach on school functioning is the thesis that school effectiveness is largely determined by selection mechanisms (effective schools are schools that attract good students, good teachers and good administrators). "Higher level causation", according to this competing perspective, would be largely determined by the contextual effects of aggregates (for instance: weak pupils do better in classes where average achievement is higher). Issues of contextual vs. "genuine" multi-level effects can be settled by including both types of variables in multi-level models and by examining the relative magnitude of regression coefficients.

The formalization of these models is quite straightforward by including $X_j$ (the average ability of pupils in school $j$) in model (3.1b). In addition to Equation (3.1a) we then also have:

$$
\beta_0 + \beta_1 z_j + \gamma_2 X_j + u_{ij}
$$

The test of genuine school effects then concerns $Y_2 \neq 0$.

A graphic representation of the contextual effects model is presented in Figure 3.4.

![Contextual or genuine school effects](image)

Figure 3.4. Contextual or genuine school effects.

Examples of these kind of contextual vs. "genuine" effects are presented in Bryk and Raudenbush (1992). They compared catholic and public schools, and demonstrated that the contextual effect of socio-economic status was twice as strong as the individual effect. But more importantly, the sector-effect (catholic vs. public), although still
Educational Effectiveness Research

significant, was shown to be far smaller once the contextual effects were taken into account.

(c) Indirect vs. direct causal effects

Conditions that are "more than one level up" with respect to educational achievement can either be seen as direct causes of achievement or as indirectly influencing achievement via intermediate levels. It should be noted that this incidence of competing causal models cannot simply be settled by comparing different specifications of the usual LISREL-type or by path-analytic models. Instead, we would need multi-level path-analytic techniques, which are in the process of development. In the absence of these models one could assess direct and indirect effects as follows:

\[ Y_{ijk} = \beta_{0jk} + e_{ijk} \]  

(3.4a)

We now have a three-level structure in which the achievement score of pupil \( i \) in class \( j \) of school \( k \) is modeled as a function of a class-specific intercept and a pupil-level error term \( e_{ijk} \). In addition to this:

\[ \beta_{0jk} = \beta_{00k} + \gamma T_{jk} + u_{0jk} \]  

(3.4b)

in which the class-specific intercept is modeled as a function of a school-specific intercept \( (\beta_{00k}) \), an instruction variable \( T_{jk} \) (for teacher \( j \) in school \( k \)) with regression coefficient \( \gamma \), and a class-specific error term \( u_{0jk} \) and finally:

\[ \beta_{00k} = h_{00} + \phi Z_{k} + v_{00k} \]  

(3.4c)

in which the school-specific intercept \( (\beta_{000}) \) is modeled as a function of a Grand Mean \( \beta_{000} \), a school level variable \( Z_{k} \) with regression coefficient \( \phi \) and a school level error term \( v_{00k} \). The indirect vs. direct effects can be tested by assessing that \( \phi \) is zero, while it differs from zero when \( T_{jk} \) is deleted as a predictor from the model. Next to this we can model \( T_{jk} \) as a function of \( Z_{k} \):

\[ T_{jk} = \psi_{0k} + w_{jk} \]  

(3.4d)

\[ \psi_{0k} = \psi_{00} + \delta Z_{k} + \nu_{0k} \]  

(3.4e)

In Equation (3.4d) the teacher variable is modeled as a function of the mean of the scores of teachers in school \( k \) \( (\psi_{0k}) \) and some teacher-level residual, whereas in Equation (3.4e) this mean is predicted by a school-level variable \( Z_{k} \) with regression coefficient \( \delta \).

Equations (3.4a) to (3.4e) form a set of multi-level structural equations, that can be estimated in the same way as ordinary path models. A graphic representation of the indirect model is presented in Figure 3.5.

One might view the example of Bryk and Raudenbush (1992) presented earlier as a means of demonstrating that the sector-effects in American high schools are mediated by contextual effects. A clear example of the path model approach is given by Kreft (1992). She tries to find empirical evidence for the hypothesis that in more heterogeneous groups pupils achieve less because they get less attention from the teacher. The interesting case of
her example is that the combination of the interactive with the indirect model eventually solves the riddle: homogeneous groups outperform heterogeneous groups, but the effect is counteracted by the indirect negative effects mediated by attention.

(d) Additive vs. synergetic interpretations

School effectiveness researchers, confronted with very low correlations between their antecedent conditions and achievement, have sometimes consoled themselves with the idea that the joint effect of several variables, which, when taken individually appear to be of marginal influence, would, in combination “do the trick”. The question is whether this magic of the whole being more than the sum of its parts is amenable to more precise and formal specification. When confronted with a set of school predictor variables one might investigate the synergetic interpretation by allowing for higher order interactions in the model. We would then have a complex interactive model of educational effectiveness. But the number of interactions potentially of interest grows exponentially with the number of predictors available. In research practice, one therefore constructs ideal types, by means of cluster analysis on the school level predictor variables. The cluster analysis then searches for groups of schools that are
as different from each other as possible, while within each group the schools are as much alike as possible. The theoretical reason for constructing such a typology is found in Mintzberg’s configuration hypothesis (Mintzberg, 1979): Organizations are effective only if they succeed in finding a consistent pattern of structuring.

A graphic representation of the synergetic effect is presented in Figure 3.6.

Empirical evidence that supports this approach is presented by Bosker (1990): Whereas no single organizational variable is linked to educational attainment of pupils, a configuration of cohesive, transparent, and goal-oriented characteristics proves to be the most successful organizational structure.

(e) Recursive vs. non-recursive models

Negative correlations between variables that are thought to be effectiveness-enhancing and achievement are no exception in school-effectiveness research. The inherent ambiguity in correlational research allows then for the interpretation that (e.g.) instructional processes are adapted to achievement levels. In fact, it is not at all implausible that several interrelationships among key-variables of school-effectiveness models are in fact non-recursive.

The idea of non-recursive systems is visually depicted in Figure 3.7.

![Figure 3.7. The recursive model.](image)

Questions about the recursiveness or non-recursiveness of certain interrelationships within school effectiveness models can be tackled in three ways:

— By means of experimental research;
— by means of alternative path-analytical models;
— by means of system-dynamic models.

De Vos (1989) presents a theoretical model that shows some non-recursive features: Individual achievement contributes to the mean group achievement, which in its turn affects the standard the teacher sets for the class. The discrepancy between individual achievements and the standard set by the teacher affects the learning gain to be made. In a following cycle this process is repeated, and so on. Empirical evidence of non-recursive relations in the field of school organizations is virtually non-existent, while at the same time the non-recursive interpretation seems all too plausible. Bosker and Guldemond
(1994) simulated De Vos' model in a hierarchical system-dynamics model for secondary schools, and were able to explain to a certain degree why efficiency, equality and effectiveness are almost irreconcilable in education. What is needed, however, to give some empirical basis to the recursive models, is longitudinal research at school level: examining school changes and their potential causes over a specific period of time.

System-dynamic approaches highlight the important question as to which exogeneous factors can "break" repetitive cycles or feedback loops. Conceptually, this issue concerns the primary "levers" of school effectiveness. The idea of meta-feedback which originates from the image of the learning organization can also be tackled by means of this methodology.

An Empirical Test of the Alternative Models

Description of the Data-Set

In the above, examples have been given of alternative models of school effectiveness as tested by several researchers. In the sequel to this chapter, a test of the competing models on one data-set will be described. The non-recursive vs. recursive models debate, for which longitudinal data are a prerequisite, cannot be settled, since this data-set is cross-sectional at the school-level. The data were gathered in 1989 as part of a Dutch national assessment of performance in mathematics, biology, the Dutch language and the English language and non-cognitive skills. The instructional behavior of teachers and the schools' organizational features were also measured. The original data-set contains data on approximately 400 schools, 1,200 teachers and 10,000 students in the third year of secondary education (details are given in Bosker, Haanstra, Lugthart, & Roeders, 1992; Kremers, 1990). In the Dutch educational system, when pupils reach the age of twelve, a choice is made as to which type of secondary education will be followed. The following main forms can be distinguished, ranked in order of decreasing difficulty:

1. VWO: pre-university education, duration 6 years;
2. HAVO: higher general education, duration 5 years;
3. MAVO: intermediate general education, duration 4 years;
4. LBO: junior vocational education, duration 4 years.

These four streams may be interpreted as curriculum tracks, with little mobility between them, in terms of upgrading midway through the course between them, except for graduated pupils (LBO certificate gives access to MAVO grade 4 and so on). The data to be used cover 680 pupils — that had not repeated a year in secondary education (insufficient marks results in staying down a year) — in the higher general (2) and the pre-university track (1). The variables for which data on these pupils are available are:

1. Sex (0: girls; 1: boys);
2. socio-economic status (a composite index based on the educational and occupational level of both parents);
3. mathematics achievement (a test developed by the Dutch National Institute for Educational Measurement).
The following data are available on the 43 mathematics teachers of these pupils:

1. Adaptive instruction (a scale as to the degree of individualized instruction);
2. Gathering of evaluative information (the number of different ways in which student progress is assessed and evaluated);
3. Use of evaluative information (the number of different ways the information is used; e.g., individual feedback or diagnosis, or group feedback or diagnosis, etc.);
4. Reduction of teacher autonomy (a scale containing the number of different fields in which teachers as opposed to functionaries at intermediate levels in the school organization have decisional discretion over instructional and learning issues);
5. Pressure to achieve (a scale measuring the degree to which teachers stress the importance of progress in the core curriculum);
6. Working conditions (a scale assessing teacher-satisfaction with the school, his or her colleagues and the students);
7. Content coverage (the degree to which the items of the mathematical achievement-test are covered by the curriculum);
8. Opportunity to learn (the percentage of the available time actually used for instruction and learning).

Moreover, six variables are available on the 30 schools in which these teachers work:

1. Curricular track (higher general track: 0; pre-university track: 1);
2. Team consensus (the views of the principal on team-consensus with respect to various educational topics);

<table>
<thead>
<tr>
<th>Table 3.1</th>
<th>Descriptive Statistics and Reliability Estimates (Cronbach's α)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Pupil-level variables</td>
<td></td>
</tr>
<tr>
<td>Mathematics achievement</td>
<td>-0.07</td>
</tr>
<tr>
<td>Sex</td>
<td>0.55</td>
</tr>
<tr>
<td>Socio-economic status</td>
<td>4.32</td>
</tr>
<tr>
<td>Teacher-level variables</td>
<td></td>
</tr>
<tr>
<td>Adaptive instruction</td>
<td>14.86</td>
</tr>
<tr>
<td>Gathering of evaluative information</td>
<td>9.72</td>
</tr>
<tr>
<td>Use of evaluative information</td>
<td>22.94</td>
</tr>
<tr>
<td>Reduction of autonomy</td>
<td>14.36</td>
</tr>
<tr>
<td>Pressure to achieve</td>
<td>26.24</td>
</tr>
<tr>
<td>Working conditions</td>
<td>59.80</td>
</tr>
<tr>
<td>Content coverage</td>
<td>0.79</td>
</tr>
<tr>
<td>Opportunity to learn</td>
<td>87.14</td>
</tr>
<tr>
<td>School-level variables</td>
<td></td>
</tr>
<tr>
<td>Curricular track</td>
<td>0.37</td>
</tr>
<tr>
<td>Team consensus</td>
<td>31.50</td>
</tr>
<tr>
<td>Evaluative practices</td>
<td>13.55</td>
</tr>
<tr>
<td>Standardization of rules</td>
<td>23.70</td>
</tr>
<tr>
<td>Mean(ses)</td>
<td>3.91</td>
</tr>
<tr>
<td>Mean(sex)</td>
<td>0.54</td>
</tr>
</tbody>
</table>
(3) evaluative practices (a composite scale for the number of arrangements made at school level to support teachers in their evaluative behavior, e.g., the availability of a student-monitoring system);

(4) standardization of rules (a scale comprising a number of occasions on which teachers have to act according to school rules, e.g., with respect to disruptive behavior of students);

(5) mean socio-economic status of the student population;

(6) proportion of boys in the student population.

Statistics on these variables (including a reliability index for the scales: Cronbach's α) are presented in Table 3.1.

Additive vs. Interactive Models

The first test concerns the issue of main effects vs. interaction effects. Since there are so many variables we opt for a very modest strategy of analysis: univariate effects of school and instructional variables will be tested for the main effects model, and only first order cross-level interactions will be investigated (i.e., the effect of one school-level variable, one instructional variable and their interaction). All analyses are done using a three-level hierarchical model in which students are nested within teachers, and teachers within schools (Bryk & Raudenbush, 1992).

Of the student level covariates only sex appeared to have a significant effect on

<table>
<thead>
<tr>
<th>School</th>
<th>Curricular track</th>
<th>Team consensus</th>
<th>Evaluative practices</th>
<th>Standardization of rules</th>
<th>Total univariate effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive instruction</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4.61%</td>
<td>-</td>
</tr>
<tr>
<td>Gathering of evaluative information</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Use of evaluative information</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reduction of autonomy</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pressure to achieve</td>
<td>-</td>
<td>10.58%</td>
<td>9.85%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Working conditions</td>
<td>-</td>
<td>-</td>
<td>8.26%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Content coverage</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.57%</td>
</tr>
<tr>
<td>Opportunity to learn</td>
<td>-</td>
<td>12.72%</td>
<td>5.89%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total univariate effect</td>
<td>64.72%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
mathematical achievement. The results of the following analyses should therefore be interpreted as effects of teacher variables and school variables on mathematics in so far as controlled for sex. Table 3.2 contains the results of the analyses.

Of the school variables, only the structural characteristic curricular track appears to have an (impressive) effect. It accounts for more than 64% of the variation between teachers and schools. None of the other school-level variables has a significant effect. Of the 8 instructional variables, only content coverage has a (modest) effect. It accounts for almost 3% of the variation in mathematics achievement between teacher and schools. There are some indications that the interactive model has more explanatory power than the main effects model: 6 interaction effects are detected. Their interpretation is, that the effect of instructional variables on achievement is contingent on school organizational conditions (and, of course, vice versa). Pressure to achieve, for example, has less effect on achievement when there is more team consensus. The three variables together, i.e., pressure to achieve, team consensus and their interaction, account for almost 11% of the variation in mathematical achievement between teachers and schools.

**Genuine vs. Contextual School Effects**

The next test is concerned with the antithesis of genuine vs. contextual school effects. The socio-economic status of the composition of the pupil body appears to have a (medium) effect on achievement: it accounts for 27.77% of the variation between teachers and schools. Using the above-mentioned results, the structural effect of curricular track (that, in its turn, is positively related to the compositional variable), is to be judged as more important, since it accounted for more than 60% of the variance. But mean socio-economic status, on the other hand, accounts for more variation than the other three school-level variables.

**Table 3.3**

School Effects on Teacher Behavior (the Percentages Refer to Variance in the Teacher Variable Accounted for by the School-Level Variable)

<table>
<thead>
<tr>
<th>School Teacher</th>
<th>Curricular track</th>
<th>Team consensus</th>
<th>Evaluative practices</th>
<th>Standardization of rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive instruction</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Gathering of evaluative information</td>
<td>13.90%</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Use of evaluative information</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Reduction of autonomy</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Pressure to achieve</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Working conditions</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Content coverage</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Opportunity to learn</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
Direct vs. Indirect Effects

The next test involves the comparison of the direct vs. indirect effects model. Since we have already presented the results for the main effects direct model, the next table (Table 3.3) only contains data on the effects of school organizational variables on instructional variables.

From Table 3.3, it can readily be deduced that the indirect model is only partially supported by the data: There are only two significant effects of school-level variables on teacher-level variables. But, as we know from the main-effects analysis, the effects of curricular track and standardization of rules on mathematical achievement are not mediated by the gathering of evaluative information nor by the reduction of teacher autonomy, since these two latter variables appeared not to have an effect on achievement. The effect of standardization of rules on reduction of autonomy is negative, which is of course peculiar, since both variables restrict teacher autonomy.

The Synergetic Hypothesis

The last alternative model to be tested concerns the synergetic hypothesis. According to this interpretation, complex higher order interactions between the variables could provide an answer. Certain configurations have to be constructed, before this model can be tested. This is done by cluster analysis on the variables that are transformed to standard normal scores, in order to ensure that when constructing the configuration all variables are equally weighted. The method used was Ward's, and the distance measure the squared Euclidian. Two sets of configurations were constructed: one set on the school and one set on the teacher-variables. The results of the school-level analyses, from which the structural variable curricular track was excluded, are presented in Table 3.4.

<table>
<thead>
<tr>
<th>Table 3.4 Configurations of Organizational Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Config. 1</td>
</tr>
<tr>
<td>n = 7</td>
</tr>
<tr>
<td>Team consensus</td>
</tr>
<tr>
<td>Evaluative practices</td>
</tr>
<tr>
<td>Standardization of rules</td>
</tr>
</tbody>
</table>

The +, −, and 0 signs indicate whether the schools belonging to a certain configuration are scoring significantly above, below or equal to the population mean.

Substantively, these configurations could be interpreted as restrictively “cooperative” (1), “bureaucratic” (2), “evaluation centered” (3), and “laissez faire” (4) educational arrangements. The configurations are not as clearcut as the synergetic interpretation: There is not one cluster of schools that scores positively on all the variables involved. Empirically speaking the best possible combination of the variables is found in the second configuration, where positive scores are found for evaluative practices and
standardization of rules; the worst possible combination is found in the two schools that make up the fourth cluster. This will be chosen as the contrast group in the analysis to test the configuration hypothesis. The effect of the school configurations on residual mathematical achievement is given in Table 3.5.

The first three configurations do better than the last one, although the effect of the third when compared to the fourth is not significant. The explanatory power of this classification is moderate: it accounts for 15.78% of the variation in residual mathematical achievement at the teacher and school level. If compared to the main effects model, the synergetic model indeed seems to be more productive than the separate school-organizational variables in terms of predictive power.

The next analysis tests synergetic effects of instructional variables. The results of the preceding cluster analysis are presented in Table 3.6.

The classification of the 43 teachers into four configurations seems somewhat more successful than the classification of the schools. Especially the contrast between the last, mainly positive-scoring, configuration and the third, mainly negative-scoring configuration is quite clear. Substantively the interpretation of the four configurations is difficult with respect to the first and second configuration, both of which are positive on the evaluation and “exposure to subject matter” dimension. The third configuration could be termed as “laissez faire”, whereas the last configuration reflects several characteristics of the classical image of the effective school with a lack of emphasis on evaluation as the clear outsider to this image. The third configuration will be chosen as the contrast group in testing the synergetic model at teacher-level. The results of the teacher-level synergetic model are presented in Table 3.7.
The third configuration does indeed have the worst effects on mathematical achievement, although the difference with the effects of the fourth is not significant. Although the synergetic model once again appears to be more productive than the main-effects model with isolated instructional variables, its explanatory power is modest: It accounts for 11.70% of the variation between teachers and schools. The interpretation of the results, however, is difficult: Teachers of the first configuration outperform the other teachers, even those of the fourth configuration, which (with the exception of evaluative emphasis) reflects most features of the "classical" effective schools model.

Conclusion

In the foregoing, we have argued that there is still a considerable discrepancy between theoretical models of educational effectiveness and research presented in this field. The empirical research often fails in explicitly formulating propositions as to how school-level factors might influence pupil attainment. In addition to this, the causal ordering of variables is somewhat problematic. From the empirical example given above, no clearcut conclusion can be drawn as to which model is best in terms of predictive power. The indirect model appears to be the least productive. But, on the other hand, the school-level variables that were included in the data-set were selected with the objective of explaining achievement directly, rather than indirectly, through teaching behavior. The synergetic model looks promising, but the more simple interactive model accounts for almost the same percentage of variation. A word of warning for those who would opt for one of the models on the basis of the empirical results presented above: The analyses suffer to some degree from chance capitalization, the synergetic model in particular. Moreover, the synergetic model actually requires a more elaborate theory: What should the configurations look like? In addition to this, there are problems with non-recursion as indicated by the following phenomena:

(1) The curricular track and the socio-economic status of the student body are positively related: Do potentially successful students choose potentially successful schools?

(2) Standardization of rules and reduction of teacher autonomy are negatively related: Should standardization compensate for autonomy and can rules be of a more loosely-structured nature if the organization is more structured?

(3) Content coverage and opportunity to learn are negatively covarying in two successful instructional configurations: Can teachers that succeed in advancing through the curriculum at a greater speed (possibly by giving the pupils more homework) give a loose to classroom activities?
The actual problem might in fact be that the most likely model might be indirect, non-recursive, contextual, synergetic, and interactive at the same time. This is of course a conclusion that does not make life easier for empirical school effectiveness researchers, if only because of the heavy data-requirements implied (large-scale, longitudinal studies). As far as the complexities of analysis are concerned, certain “decompositions” could be employed. We therefore contend that the organizational variables of interest first be examined on their possible alternative interpretation of cross-level facilitation, before formulating the model. “Mirroring”, for instance, would imply a recursive model, with the effect of school-level mirrors being mediated by instructional images of the mirror. “Buffers”, to give another example, might well be modeled in an interactive way: teaching quality might only be important in an orderly organized school.

As these examples show, the poly-interpretability of global and vague conceptual models may give rise to interesting competing causal specifications amenable to statistical analysis and empirical verification.

References


