EVALUATION STUDIES

PROCESS INDICATORS OF SCHOOL FUNCTIONING: A SELECTION BASED ON THE RESEARCH LITERATURE ON SCHOOL EFFECTIVENESS

Jaap Scheerens

University of Twente, Department of Education, P.O. Box 217, 7500 AE Enschede, The Netherlands

Introduction

Educational indicators are statistics that allow for value judgements to be made about key aspects of the functioning of educational systems. To emphasize their evaluative nature, the term "performance indicator" is frequently used. Included in this definition of educational indicators are:

- The notion that we are dealing with measurable characteristics of educational systems;
- The aspiration to measure "key aspects", be it only to provide an "at a glance profile of current conditions" (Nuttal, 1989) rather than in-depth description;
- The requirement that indicators show something of the quality of schooling, which implies that indicators are statistics that have a reference point (or standard) against which value-judgements can be made.

Usually policymaking at national level is seen as the major source of application of indicators (indicator systems as policy-information systems). This view on the application of indicators should be enlarged, however, since consumers and "third parties" like private industry are also seen as users of the information that indicator systems provide. Likewise, the education system at local administrative level and even individual schools
could also use indicators to support policymaking (indicator systems as management information systems).

During the last decade various types of collections of indicators, usually referred to as indicator-systems, has been proposed and a sub-set of these have been actually used. Van Herpen (1989) gives a comprehensive overview of what he calls "conceptual models of educational indicators". For our purpose it is sufficient to discern some major developments in these various approaches to conceptualizing education indicator systems.

Economic and social indicators are the origin of educational indicators. "Social indicators of education" describe educational aspects of the population, whereas educational indicators describe the performance of the educational system (Van Herpen, 1989, p. 10). The first trend in the development of educational indicators was the transition from descriptive statistics to measuring performance, or, more generally, a shift towards statistics of evaluative importance.

When we look at developments in educational indicators at the National Center for Statistics of the US Department of Education we can discern a second trend. At first this offered descriptive statistics on the state of the educational system, including data on inputs and resources. Since 1982, "outcome" and "context" data were given a more prominent place, and in a recent proposal to redesign the education data system, "process" aspects of the functioning of educational systems were also included (Stern, 1986; Teauber, 1987). This second trend can thus be characterized as a movement towards more comprehensive indicator systems, first by adding output measures and context measures to the more traditional measurement of inputs and resources, and secondly by a growing interest in "manipulative input factors" and process-characteristics.

The third trend is somewhat related to the second one, as far as the interest in process characteristics is concerned. Traditionally indicator systems have concentrated on macro-level data, such as national illiteracy rates, the proportion of pupils that have passed their final secondary examinations, school etc. When we think of process-indicators as referring to the procedures or techniques that determine the transition of inputs into outputs, interest in process-indicators naturally leads to an interest in what goes on in schools. So, the third trend in conceptualizing indicator systems is to measure data at more than one aggregation level (national system, school, perhaps even the classroom; see Teauber, 1987; Scheerens et al., 1988).
What emerges from this brief overview of developments in the field of educational indicators is the notion that a context-input-process-output model is the best analytic scheme to systematize thinking on indicator systems. Such a model is depicted in Figure 1.

Process indicators and their specific place within educational indicator systems will be discussed in the next section. In the following sections the literature on school effectiveness will be used as a basis for the selection of specific process indicators.

Basic Questions on Process Indicators

The question arises whether process measures of school functioning do indeed conform to the definition of educational indicators that has
already been given. It is doubtful whether process measures as such can be used as a basis for judging the performance of an educational system.

Would one, for instance, be willing to accept the degree to which a school (or a school district) used a specific curriculum as a trustworthy criterion for judging educational performance? Process measures relate to outcome measures as a means to an end and thus, it could be argued, using these as performance indicators would be "goal displacement" in a new form. To put it bluntly, process indicators could lead to evaluative conclusions of the "operation succesful, patient deceased" kind. The only way out of this would be the existence of empirically supported causal models of educational performance, from which the importance of specific process measures could be deduced. Unfortunately, as we shall further demonstrate, no such established causal models exist. Generally, the variance in output measures that is accounted for by input and process measures is rather low. I believe the only legitimate way to employ process indicators is to always link them to output indicators. Process indicators then have the function of offering hypothetical explanations on why certain schools, or school systems do better than others. The notion that process indicators derive their value from their relationship with output indicators forms the basis of this paper; namely to select process indicators by examining the findings of school effectiveness research.

Process indicators generally refer to characteristics of educational systems that can be manipulated. Adding process measures therefore enhances the policy relevance of indicator systems. The question of measurement is also another point.

Generally, indicators are thought of as quantitative "low inference" measures. Process characteristics like school climate, educational leadership and opportunity to learn are rather broad characteristics; their operationalization and quantification is neither straightforward nor "low inference". Besides, elaborate procedures sometimes have to be used to collect data on these. Remaining strictly within the domain of indicators this problem can only be solved by using "proxy" measures of process characteristics, for instance by measuring instructional time during school days in a year as a proxy for the time students spend on task-related activities, and by omitting all process variables that are not amenable to this kind of approximations. Another way of tackling this problem is to use in-depth studies (surveys or observational studies) that are connected to regular indicator systems.
School Effectiveness Research and the Identification of Process Indicators

As has already been stated the most likely rationale for selecting process indicators is to choose those variables that are manipulative predictors of school output. Research literature on school effectiveness can be used as a source to identify promising process variables. Generally speaking, school effectiveness research is aimed at discovering school characteristics that are positively associated with school output, usually measured as students' achievement. Various research traditions can be subsumed under this heading, including (in)equality of education (sociological), educational production functions (economical), school improvement and effective schools, and teacher- and instructional effectiveness (psychological). Apart from these, more theoretical and analytic contributions from organizational science and micro-economic theory of public-sector organizations can also be sources of inspiration in selecting process indicators (see Cameron & Whetten, 1983 and Niskanen, 1971, respectively).

In the following sections the main findings of each of the four types of school effectiveness research will be briefly summarized and used to generate proposals for process indicators. We shall take the findings of school effectiveness research at face-value, and later on shortcomings in interpreting these results will be discussed.

Research on (In) Equality in Education and School Effects

The Coleman report (Coleman et al., 1966) on the Equal Educational Opportunity Survey should be seen as the impressive starting point for school effectiveness research. Although the major thrust of the Coleman survey was to investigate (in)equality in education, it also became quite famous for its supposed negative conclusions on the influence of school on educational achievement. Coleman et al. found that schools accounted for approximately 10% of the variance in pupil achievement, after statistical adjustments had been made for the influence of background characteristics of pupils. The significance of this finding for judging the importance of process indicators on school functioning is that all feasible process indicators one could think of would not account for more than ten percent of the variance in pupil achievement.

However, depending on the achievement measure that is used and the heterogeneity versus the homogeneity factor of school system in a particular country, later research has found a higher percentage explained
by schools. Coleman's 10% is not exceptional when considering more recent studies (cf. Purkey & Smith, 1983; Bosker & Scheerens, 1989). Other large-scale studies replicated Coleman's findings in their pessimistic conclusions on the importance of schooling as such and its possibilities to lower educational inequality (Jencks et al., 1972; Hauser, Sewell & Alwin, 1976; Thorndike, 1973). Specific school characteristics that were measured in these studies were mainly resources and material inputs (such as the age of the school building, per pupil expenditure and the number of books in the library) although some measures of teacher attitudes and instances of classroom management were used. The significance of this first generation of school effectiveness research for the issue of educational indicators can be summarized in three points:

a. School process variables account for relatively little variance in educational achievement. The educational significance of this will be given further consideration in a subsequent section.

b. Resources and "material" inputs are not very promising in explaining school output, though this would not necessarily imply that they should not be included in indicator systems, because in heterogeneous school systems in, for instance, developing countries, they might still be of great importance.

c. Pupil background characteristics such as socio-economic status or race should be used to adjust raw output measures to arrive at fair and valid performance indicators and to allow for an unbiased interpretation of the influence of process characteristics on the functioning of schools.

Research on Educational Production Functions

The economic approach to school effectiveness is concerned with the question which inputs lead to more output, also considering the cost of the inputs. Stated in more abstract terms knowledge about stable relationships between input and output variance is sought in order to specify a function that could express the effects more inputs would have on output. This school of effectiveness research is both known as input output analysis and as research on educational production functions. In fact this type of research is very similar to other types of educational effectiveness research in that the relationships between school characteristics and achievement is investigated, while adjusting for background characteristics of pupils (such as level of intelligence and socio-economic status). The characteristic that
sets this research tradition apart is the choice of a particular category of inputs that are readily expressed in monetary terms, such as teacher salary, teacher experience, teacher-pupil ratio, teacher qualifications, per pupil expenditure.

The results of this type of effectiveness research are rather disappointing. Research reviews like those of Mosteller & Moynihan (1972), Averch (1974), Glasman & Biniaminov (1981), Hanushek (1979, 1986) agree upon the inconsistency of research findings and the rather small effect of the input variables concerned. To illustrate this, Table 1, quoted from Hanushek (1986, p. 1161), shows the result of a research review that included 147 studies.

Table 1: Review of 147 Input-Output Studies; + Indicates a Positive Association, - a Negative Association of Input and Output; Quoted from Hanushek, 1986, p. 1161

<table>
<thead>
<tr>
<th>Input</th>
<th>statistical significant</th>
<th>statistical non-significant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of studies</td>
<td>total</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>n.s. +</td>
<td>unknown</td>
<td></td>
</tr>
<tr>
<td>teacher/pupil ratio</td>
<td>112</td>
<td>89</td>
</tr>
<tr>
<td>teacher qualification</td>
<td>106</td>
<td>95</td>
</tr>
<tr>
<td>teacher experience</td>
<td>109</td>
<td>69</td>
</tr>
<tr>
<td>teacher salary</td>
<td>60</td>
<td>50</td>
</tr>
<tr>
<td>per pupil expenditure</td>
<td>65</td>
<td>49</td>
</tr>
</tbody>
</table>

As Table 1 indicates, only the variable "teacher experience" shows some consistency, in that 30% of the estimated coefficients are statistically significant.

Hanushek's overall conclusion is that as yet educational expenditure is not consistently related to achievement. He suggests that it would take greater variation in inputs to expect important effects. So, for instance, in most countries where teacher salaries are strictly regulated and rather uniform, if a system of "merit pay" would be introduced the variance in teacher salaries would be expected to increase and significant effects in pupil achievement might be revealed.
Although this research tradition does not focus on "process" measures - in this respect the production functions are rather primitive in that the whole area of educational technology remains a black box - some of the input variables could be considered for inclusion in indicator systems. Judging from Hanushek's research synthesis, teacher experience would be the most likely candidate. Yet, I think, particularly when one wishes to construct educational indicators for international comparison, it would be wise to include variables like per pupil expenditure and teacher/pupil ratio, since these might show significant variance between countries.

Effective Schools Research

After the Coleman report a second wave of school effectiveness research came into being. Its pioneer studies can be seen as a reaction against Coleman's negative conclusions. As titles like "Schools can make a difference" (Brookover et al., 1979) and "School matters" (Mortimore et al., 1988) show, refuting the message of the Coleman report has been, and still is, an important motive for this more recent research. The most important characteristic that distinguishes the effective schools research from earlier school effectiveness research is that the black box of what happens within schools is opened and school variables are revealed that include school organization, school culture and educational technology. Several types of effective schools research can be distinguished.

First of all, there is a series of studies where exceptionally effective schools were identified and described as case-studies, sometimes also comparing them to ineffective schools in order to discover which of their school characteristics contribute to their superior results. Examples of these studies are those by Lezotte, Edmonds & Ratner (1974) and Weber (1971). School characteristics that were consistently associated with positive achievement were:

- A safe and orderly school climate;
- High expectations of pupils' achievement;
- Educational leadership (i.e. a school leader who is actively involved in developing and monitoring educational activities, and who is more than merely an administrator);
- Frequent evaluation of pupils' progress;
- Clear objectives concerning basic skills;
- A cooperative atmosphere among the teaching staff.
Table 2: Twelve Key Aspects of Effectiveness (Mortimore et al., 1988, p. 250)

1. **Purposeful leadership of the staff by the headteacher.**
   Key aspects: active involvement in school's work without exerting total control over the rest of the staff, record keeping.

2. **The involvement of the deputy head.**
   Key aspect: sharing of responsibilities of head and deputy.

3. **The involvement of teachers.**
   Key aspects: active involvement in curriculum planning, participation in decision making on school policy.

4. **Consistency amongst teachers.**
   Continuity in the teaching staff and consistency of teacher approach.

5. **Structured sessions.**
   Key aspect: teachers offer a strict framework within which pupils can work, yet allowing them some freedom.

6. **Intellectually challenging teaching.**
   Key aspects: asking of higher-order questions, enthusiastic approach, high expectations of pupils.

7. **Work centred environment.**
   Key aspects: much content related work and feedback - relatively little time spent on routine matters and the maintenance of work activities: a low level of noise.

8. **Limited focus within sessions.**
   A focus upon only one curriculum area in a lesson.

9. **Maximum communication between teachers and pupils.**
   A flexible approach, using a blend of individual, class and group communications - whole class teaching increased the amount of communication with all pupils in comparison to an approach where teachers devoted the majority of their time to speak with individual pupils.

10. **Record keeping.**
    Record keeping by both headteachers and teachers.

11. **Parental involvement.**
    Key aspects: help in classrooms, educational visits, attendance at meetings, parents' reading to their children, access to books at home.

12. **Positive climate.**
    Key aspects: more praise than blame, enthusiastic attitude of teachers, friendly attitude of pupils - absence of graffiti around the schools.

A second type of effective schools studies arose because of the rather surprising fact that - despite the relatively weak research basis for the above-mentioned studies - these findings were almost immediately used as a guiding principle for school improvement programs. Although few of these school improvement programs have been rigourously evaluated, those
that have been show positive program effects (e.g., Miller et al., 1985, Achilles & Lintz, 1986; McCormack-Larkin, 1985). It should be noted that these evaluations of school improvement programs show the overall effect of multiple changes in the functioning of schools, including the better educational leadership, building an achievement-oriented morale among pupils, stimulating professional attitudes and cooperation among teaching staff, applying structured teaching methods that include frequent evaluation, feedback and reinforcement.

A third type of effective schools research comprises a blending of the approaches of the earlier large scale input-output effectiveness studies with the more recent interest in process characteristics of school functioning. Studies like those of Brookover et al. (1979) and Mortimore et al. (1988) combine relatively large sample survey research with in-depth description of school processes. Moreover, the latter study is not exclusively involved with characteristics measured at school level but also includes aspects of instruction at classroom level. By way of illustration, the factors that Mortimore et al. (1988) found to be positively associated with pupil achievement are listed in Table 2.

It is quite obvious that ideally these effective school characteristics would require intensive data collection and high inference measurement. Some of them, however, can be operationalized in terms of relatively uncomplicated scales or questionnaire items. Some examples are

Educational leadership
- The amount of time headteachers spend on educational matters, as opposed to administrative and other tasks;
- Whether headteachers do or do not discuss test results on pupils' progress with teachers;
- The amount of instructional issues on the agenda of staff meetings;

Achievement oriented policy
- The amount of overt statements in official school documents that express an achievement oriented emphasis in school policy;

Orderly and safe climate
- Statistics on absenteeism, lesson drop-out and delinquency as instances of the degree of order in the school;
- Ratings of school discipline by teachers and headteachers;
Clear objectives
- Whether or not explicit school curricula, stating educational objectives and levels of achievement are available;

High expectations
- Estimates by teachers and/or headteacher of the percentage of students that will complete their secondary schooling;
- Student estimates of their further educational career;

Monitoring/evaluation of pupils' progress
- The frequency of the use of curriculum specific tests at each grade level;
- The same with respect to standardized achievement tests;
- Whether or not the school uses a (computerized) system to monitor pupils' progress at all grade levels;

Continuity and consensus among teachers
- The amount of changes in staff over a certain period;
- The presence or absence of subject-based communication platforms (secondary education);
- The degree of opportunity for communication and cooperation among staff.

Some characteristics associated with school effectiveness could be assessed by unobtrusive observations by researchers or inspectors. For instance the presence or absence of grafitti in and around the school and the noise level in school corridors. Other characteristics, mainly at classroom level, would require detailed observation and reporting by teachers and pupils. It would require quite some creativity to find proxy-measures for these effective school characteristics that would fit the requirement of easy, low-inference measurement. Some characteristics that might be readily available from school records include:

- Staff turnover within a certain period;
- Existence of a specific school curriculum;
- The presence or absence of procedures and records to monitor student progress;
- Whether or not schools keep records on examination results and the percentage of students passing on to higher education.
It would be far beyond the scope of this paper to attempt to give a comprehensive overview of the research literature on teacher effectiveness, effective classrooms, instructional effectiveness and the like. The distinguishing characteristic of this stream of educational research - compared to the various types of school effectiveness studies described earlier - is the fact that process characteristics of education are studied at the teacher or classroom level. So, when we are considering variables at this level that have been found to be associated with achievement, we are really delving into the primary process of schooling. We shall come back to the question which place these kind of micro-level variables could have when considering indicator systems.

Where a detailed overview of the relevant research literature is out of the question, the next best solution in the search for the most promising process variables is to use already existing research reviews. Fortunately, many excellent reviews are available (e.g., Kyle, 1985; Brophy & Good, 1986). Moreover, a growing amount of quantitative syntheses of research on educational productivity has become available (Walberg, 1984; Fraser et al., 1987; Kulik & Kulik, 1982).

This review literature reveals a set of categories of variables that have repeatedly been shown to be positively associated with achievement (for a more detailed account of the selection of these variables see Scheerens, 1989). These are:

*Effective learning time or "time on task"*. Specific aspects of learning time are: the duration of the schoolday, schoolweek and schoolyear, whether or not pupils get homework assignments, the amount of the official duration of lessons that is actually spent on task-related work, absenteeism, drop-out of lessons and reallocation of the total time that is available for instruction over school subjects. Issues of school discipline are also related to the amount of time that is effectively available for instruction.

Although interpreting the results of increased effective learning is straightforward, two points should be noted. Firstly, it is obvious that extending the official school hours must at some point become counter-productive. Secondly, moderate increases in learning time have yielded only moderate effects on achievement (cf. Levin, 1988; Walberg, 1984).
Structured or "direct" teaching.

The general idea of structured or "direct" teaching is the application of frequent interventions to support the learning process. Examples of this include: stating educational objectives clearly, dividing the total subject matter that must be learned into relatively small units, providing a well-planned sequence of these units, providing many opportunities for pupils to do exercises, giving cues and hints, frequent questioning and testing to monitor progress, and giving feedback.

Mastery learning is a didactic approach in which most of these principles are represented. Although direct teaching has been demonstrated to be particularly effective in primary education and for the teaching of basic skills, it has also been shown to work in secondary education and in the teaching of higher order cognitive skills, though in a somewhat modified form - larger steps in subject matter presentation, more initiative for pupils (cf. Doyle, 1985; Collins & Stevens, 1982). Adaptive instruction, i.e. adapting instruction to pupil characteristics, in matters like pace and way of presentation, can be seen as a more individualized use of structured teaching.

Opportunity to learn or "content covered". The essence of "opportunity to learn" is the correspondence between the subject matter that has been taught and the content of the tests that are used to measure achievement. As is to be expected, pupils do better when the subject matter is covered by test-items.

Teacher attitudes and expectations. As was shown in older literature on teacher research, the enthusiastic attitude of teachers is important. From more recent studies where effective teaching is compared to less effective teaching in inner-city schools, it seems important that teachers remain optimistic about the capabilities of their student; nothing seems worse than a defeatist attitude. This attitudinal factor of teacher functioning directly corresponds to the high expectations variable known from school effectiveness literature.

Enhancing student motivation. According to Walberg's (1984) research synthesis, the variable most strongly related to achievement is reinforcement. This variable is closely related to structured teaching, where frequent monitoring of progress and feedback are important. It appears that praise, or positive feedback, works far better than punishment (see Brophy & Good 1986, for a more detailed analysis).
The alterable curriculum of the home. This category of variables - as it was named by Walberg (1984) - covers for the ways in which the home situation of students can affect school performance. Positive instances of this category include parental interest in what children do at school, reading to children at home and moderate television viewing.

Excursion: Towards a More Comprehensive Conceptualization of School Effectiveness

Before going into the question of whether these instructional variables are amenable to relatively easy low inference measurement, and more generally, what part they might play in indicator systems, an attempt will be made to integrate the results of the various types of educational effectiveness studies into one general model. As already stated, the rationale that is being followed to identify and select useful process indicators is to look for those process-variables that are regarded as manipulative predictors of educational achievement.

Thus, the development of a conceptual model of school effectiveness is of direct relevance to identifying prospective process indicators. Some of the building blocks of a comprehensive model of school effectiveness have already been presented: a context-input-process-output-analytic framework and the most important results from various types of school effectiveness research. To these three "ingredients" must be added:

a. A multi-level framework which is most useful in specifying the input-process-output-context elements of school effectiveness models;
b. A perspective to analyse the relationship between context- and process variables known as "contingency theory";
c. A perspective for the relationship between school-level organizational and managerial conditions and process characteristics of school functioning at the micro (i.e. classroom) level.

Multi-Level Framework

For the purpose of developing a conceptual model of school effectiveness it is most useful to think of the measurement of the various types of variables at the lowest level of aggregation at which they can be defined. Besides conceptual clarity this has the advantage that available statistical techniques for multi-level analysis yield the most precise results.
This implies that the context-input-process-output framework from which we start would need at least four levels:

- The measurement of output (achievement/attainment) and important background variables (SES, previous educational achievement, intelligence) at student level;
- The measurement of instructional process variables at teacher/classroom level;
- The measurement of organizational, curriculum and managerial process variables at school level;
- The measurement of material and financial inputs, also at school level;
- The measurement of context variables either at school level (i.e. school size) or regional level.

Contingency Perspective

The general view of contingency theory is that it depends upon contextual characteristics whether or not specific organizational structures or managerial processes will be effective (e.g. Mintzberg, 1979). At first glance contingency theory seems at odds with school effectiveness research, which is concerned with a set of school characteristics that is very robust in predicting effectiveness. On closer analysis of the research and literature on school effectiveness it becomes clear that claims regarding the generalization of effective predictors across contexts have only partly survived empirical tests (Firestone & Herriott, 1982; Teddlie et al., 1987; Scheerens, Nanninga & Pelgrum, 1989). In fact, including contextual variables like student-body composition, school type, or national educational context can be seen as a relatively new and very interesting development in school effectiveness research. Some studies even try to establish interactions between various contextual conditions and performance (i.e. Friedkin & Necochea, 1988, who investigated the interactional effect of school size and SES student-body composition on performance).

Two applications of contingency thinking seem to be particularly relevant to the subject of school effectiveness. First, in organizational theory the notion of effectiveness as it is used in research literature on school effectiveness (i.e. productivity) is sometimes seen as just one particular type of effectiveness. Alternative effectiveness criteria are resource acquisition, stability and control in the functioning of the organization and cohesion and
morale among the organization's members (Cameron & Whetten, 1983; Faerman & Quinn, 1985). Depending on contingency factors such as the degree of environmental uncertainty the organization is faced with, or the stage of its development, the emphasis on each of these types of effectiveness criteria might shift. For instance, when a school is faced with important drops in enrolment, it is quite understandable that relatively more energy will be put into acquiring more pupils. And when a school, for whatever reason, has to merge with another one, a lot of attention will have to be given to reaching stability and control in the new organization.

As school effectiveness research shows, schools differ in the degree to which achievement is emphasized in school policy and teacher expectations. Moreover, achievement orientation is generally found to be positively related to actual achievement. The organizational conceptualization that recognizes effectiveness criteria other than productivity and explains their relative influence in a particular setting by referring to contextual conditions, helps in answering the question why some schools are more achievement oriented than others.

The second instance of contingency thinking relevant to school effectiveness is recognizing the importance of external incentives on achievement oriented school policy.

The political will of a school to achieve is perhaps the most essential condition for actual school effectiveness. When higher administrative bodies, consumers, or other stake-holders also emphasize achievement or even reward schools for high achievement and "punish" others for low achievement, this political will can even be seen as a malleable factor. In this respect a range of measures including output finance of schools, "privatizing" schools, deregulation, voucher systems and publishing schools' performance in local newspapers is relevant. Micro-economic theories on the efficiency of public sector organizations (i.e. Niskanen, 1971, Breton & Wintrobe, 1982) explain how these mechanisms operate. For our purpose we can put all these external mechanisms for stimulating schools to be effective under the heading achievement incentives in the local and the larger school environment. At the local level we could look at the presence or absence of achievement standards for schools, clear objectives, evaluation systems and public records on school achievement. When comparing national educational systems one could measure the presence or absence of assessment projects, the degree to which the inspectorate employs output evaluation, the development of consumerism on education, whether or not output finance is used, etc.
Meso-Micro Relationships

In our earlier review of different types of educational effectiveness research we discussed research which focuses on process characteristics at the school level and instructional effectiveness research where processes at the teacher or classroom level are of central interest. Both research schools have yielded a list of the most promising process characteristics, i.e. school and instructional characteristics respectively. The relationship between these two categories is an important aspect of our envisaged comprehensive model of school effectiveness. The most straightforward way of seeing this meso (school level) - micro (classroom level) relationship is to assume that meso-level conditions facilitate micro-level conditions. This implies that instructional processes are seen as the most direct determinants of school learning and achievement, and that organizational and curricular conditions at school level are thought of as more indirect conditions of educational achievement.

When we examine the actual research outcomes on meso and micro conditions of educational achievement more closely, it is evident that some important variables are meaningful at both the school and teacher/classroom level:

- Structured teaching at classroom level can be stimulated by means of explicit curricular policy at school level (e.g. by using school development plans);
- High expectations of student performance is essentially a variable defined at the teacher level, though its aggregate, an achievement oriented school policy, may be taken as a whole that exceeds the sum of its parts;
- Order in classrooms will be enhanced by an orderly atmosphere in other parts of the school building;
- Frequent monitoring of pupils' progress will usually take place at classroom level, though this evaluation may be a result of a school evaluation policy and will benefit from instruments at this level, such as computerized school evaluation or a management information system;
- Opportunity to learn can be defined at the classroom level, but can also be seen as being enhanced by a school curriculum that is closely linked to the educational objectives that determine the contents of achievement tests.
Apart from these factors that can be defined at both school and classroom level, we can discern a second class of conditions at school level that facilitate effective instruction at classroom level. Their successful operation is dependent on the organization's superstructure (e.g., management, coordination structures) protecting the core production process against disturbances and external uncertainties (Thompson, 1967). Instructional leadership, the degree of collaborative planning and collegial relationships plus an active policy in recruiting students and acquiring resources are examples of the latter type of conditions that have received some support in research literature on school effectiveness.

Synthesis

The ingredients for a comprehensive model of school effectiveness are in place. To summarize we have:

- An analytic systems model recognizing context, input, process and output variables;

- A multi-level framework discerning pupil-, classroom-, school and environmental characteristics;

- Perspectives to view the interrelationships between variables defined at different levels, most notably contingency theory and organizational conditions that facilitate the schools' primary process;

- Substantive findings from different types of educational effectiveness research.

Figure 2 gives a schematic summary of the model.

The hypotheses that have been stated on the interrelationship of the various categories of variables form the extra dimension that this model offers to a mere listing of the most promising variables. Although empirical testing of this integral model, and developing the theoretical explanations of these hypothetical relationships is clearly beyond the scope of indicator development, its present tentative formulation might nevertheless be helpful in selecting those process variables that are most relevant in exploring the causes of achievement differences between schools.
Process Indicators

Context
- achievement stimulants from higher administration levels
- development of educational consumerism
- "co-variables" like school size, student-body composition, school category, urban/rural

Inputs
- teacher experience
- per pupil expenditure
- parent support

Process
- school level
  - degree of achievement oriented policy
  - educational leadership
  - consensus, cooperative planning of teachers
  - quality of school curricula in terms of content covered, and formal structure
  - orderly atmosphere

Classroom level
- time-on-task (including homework)
- structured teaching
- opportunity to learn
- high expectations of pupils' progress
- degree of evaluation and monitoring of pupils' progress
- reinforcement

Outputs
- student achievement, adjusted for:
  - previous achievement
  - intelligence
  - SES

Figure 2: Integral Model of School Effectiveness

Some Problems in Interpreting the Results of School Effectiveness Research

Empirical Basis

The research base that supports the list of factors that is presented in Figure 2 varies considerably among the individual variables. Some variables have been shown to have a positive association with achievement in many
studies, others, like for instance per pupil expenditure, have been investigated many times but the results show no consistent pattern. Some other variables only have a relatively weak research base and their inclusion in the model is more a matter of conjecture than solid empirical evidence.

Table 3 gives an overview of the degree to which each of the factors in Figure 2 has been empirically supported.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Strong empirical basis</th>
<th>Moderate empirical basis</th>
<th>An, as yet, weak empirical basis</th>
<th>Mostly conjecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental incentives</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumerism/parent involvement</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Teacher experience</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per pupil expenditure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achievement oriented policy,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high expectations</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational leadership</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Consensus, cooperative planning</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Quality of curricula</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluative potential</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orderly climate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured teaching</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time on task</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opportunity to learn</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Reinforcement</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>
As appears from Table 3 we are still far from an empirically supported comprehensive school effectiveness model. In my opinion this should not preclude the use of the available research evidence as a reasonable desideratum to identify prospective process indicators. At the same time we should realize that the available research base still leaves much uncertainty as to the status of certain factors, which also implies that other reviewers of the literature will probably come up with somewhat varying lists of most important factors (see for example Purkey & Smith, 1983; Bossert, 1988).

Narrow Focus?

Although the use of educational attainment data (e.g. the proportion of an age cohort that passes the final examination of secondary school) is not at all uncommon in school effectiveness research, most effectiveness studies have used achievement test data on a limited set of subjects (language and mathematics).

Since national educational systems, school districts and individual schools may differ in the degree to which they value these particular educational outputs, they will also look differently upon the usefulness of school effectiveness models that rest upon this narrow definition of school output. A counter-argument to this possible source of normative relativity of school effectiveness research findings is that we are dealing here with basic educational objectives that appear in tests and examinations exactly because they are thought of as very important. Yet, it seems wise to keep an open mind to differences in normative contexts in which indicator systems, inspired by school effectiveness models, are to be applied.

"Small" Effects

A problem in interpreting the results of school effectiveness research is in how far the sizes of malleable educational characteristics are to be seen as educationally significant. It is common knowledge that these process characteristics generally explain relatively little variance in pupil achievement. Before going into this, it should be noted first and foremost, that we are always dealing with relative school effects, that is, the degree to which schools differ among themselves in average achievement.

Defensive research titles like: "Schools can make a difference" and "School matters" seem to have been used as weapons against the straw-man argument that it does not matter whether or not children go to school. In a
sample of about 30 school effectiveness studies (Scheerens, 1989) the average variance between schools was about 12% of the total (i.e. pupil-level) variance in achievement (range: 5% - 40%). (All figures on school effect sizes that are given in this section are based on achievement scores that have been adjusted for pupils' background characteristics). This implies that specific school-level characteristics can only account for a certain proportion of the variance between schools, which means that these specific school-level characteristics may not explain more than, say, 4% of pupil-level variance. This tiny effect would - in a situation where variance between schools is 12% - still mean that the variable in question explained one-third of all that specific school factors could explain in this particular situation.

In interpreting school effects we should bear in mind that up to 75% of the variance in pupil achievement can be accounted for by factors like intelligence and social economic factors which are to a large extend beyond the reach of educative manipulation. This means that the total direct influence of schooling is relatively small as far as a pupil's overall achievement level is concerned.

Interpreting school effects in their proper perspective could benefit from more insightful ways of expressing effect sizes (see Rutter, 1983; Bosker & Scheerens, 1989). A general approach to accomplishing this would be to attach some kind of societal value to score levels on the output variable. Purkey & Smith (1983) give an example of this by translating a difference of two-thirds of a standard deviation in achievement between highest scoring and lowest scoring schools into a certain time interval (in their case one year) that expresses the degree the average pupil of the highest scoring school is ahead of the average pupil of the lowest scoring school.

The conclusion on the "small effects" in school effectiveness research is that we should use more specific expressions of effects than merely "percentage of total variance explained", and by doing so we may discover that so-called "small effects" may still be of educational importance.

Conceptual and Methodological Problems in School Effectiveness Research

Conceptual problems in interpreting the results of school effectiveness research are closely related to the correlational nature of most studies and the "inductive" approach in building school effectiveness models. So, for instance, it is hard to say whether a variable like "high
expectations of pupils' progress" is to be seen as cause or effect of high achievement, when we have not experimentally manipulated the expectations variable. More generally, we may have some ideas now on what variables work in education, yet, we still have little knowledge on the causal mechanisms that explain the correlations. The links with more general explanatory principles or theories are still relatively weakly developed (cf. Scheerens & Stoel, 1988).

Methodological weakness of school effectiveness research has been thoroughly described in various review articles (e.g., Ralph & Fennessey, 1983). "Small samples", "insufficient adjustment for important background characteristics of schooling", "reactive research arrangements" and "improper use of analysis techniques" are the main points of criticism.

Although these conceptual and methodological problems are certainly one more reason to treat the results of school effectiveness research cautiously, some counter-arguments should be considered too. As to the methodological criticism, it should be remarked that results have shown a considerable robustness in the face of varying research settings and research approaches. An argument for the emergent school effectiveness model that is hard to neglect - though it is not very scientific - is its intuitive appeal. Some of the basic principles, like more success in specific kinds of school output when these outputs are actively strived for and are an explicit policy of the school, and better results when more time is invested, simply make sense. Finally, some attempts have been made to link the findings of school effectiveness research to more general explanatory principles like learning theory, X-efficiency, and certain conceptualizations from organization theory (Scheerens & Stoel, 1988).

Despite all cautionary notes that have been made in this section it still seems permissible to use the research literature on school effectiveness as a best guess for proposing likely candidates for process indicators. Perhaps the most important suggestion that these points of criticism have yielded is to consider carefully the usefulness of specific variables within the particular (national or local) educational context.

Multiple Context, Multiple Level and Multiple Measurement Applications of Process Indicators

As stated in the introductory section, indicator systems can be used in several administrative contexts: by national educational policymakers; by
officers at local level, by managers at school level and even by teachers at the classroom level.

A second distinction that has been used in this paper is the aggregation level at which basic variables, from which indicators can be computed, are measured. In our comprehensive model of school effectiveness we discerned measures at student, teacher/class, school and school-context level. Although we use the same levels as reference points for both distinctions there is no compelling one-to-one correspondence between the two, since the highest administrative level may wish to use measures at any level of aggregation.

A third dimension that can be used for classifying indicators refers to the nature of measurements. Whether we are dealing with high or low inference measurements (e.g., the number of books in a school library vs. educational leadership), whether or not measures are readily available from existing data sources and whether we can use standardized quantitative scales or must rely on relatively unstructured methods (such as participant observation or "open" interviews). The nature of measures on which educational indicators rely depends on whether indicator systems are exclusively thought of as "closed" information systems of longitudinal data or as "looser" structures that also allow for the inclusion of specific evaluative studies (such as the international comparisons of educational achievement conducted by the International Association for the Evaluation of Educational Achievement [IEA], or in-depth studies that could be conducted by audit commissions or the inspectorate. When we are looking at educational indicator systems from the perspective of multiple administrative levels, multiple levels of aggregation of basic data and multiple types of measurement, we are clearly using a broad interpretation of the term "educational indicator". The common core of all these applications still conforms to the definition that was stated in the ideology section: measurement of key aspects of educational systems that are of evaluative relevance.

According to the three distinctions that were discussed, we could use a three-dimensional framework to classify types of indicator applications (see Figure 3).
In order to determine the role of process indicators within this framework of applications some of the most likely combinations of scale points on the three dimensions of Figure 3 will be examined.

Determining the Condition of Education at the National Level

First, we consider decision-making at the national level supported by summary statistics at high levels of aggregation using low inference measures and existing data bases. This is the more traditional type of application of educational indicators.

Descriptive statistics on, for instance, enrolments and financial data are likely to be included in this type of application of indicators, as are statistics on educational productivity in terms of, for instance, proportion of age cohorts that pass final examinations. Could one conceive of a place for process indicators in this kind of macro-level indicator system? The aim of including process indicators would be to gain insight into possible causes for high or low educational productivity in a particular period. The difficulty with process variables (see the list that is contained in Figure 2) is, of course, that most of them are neither easily measurable nor readily available from existing data sources. Some proxy variables that might be considered as a basis for macro-level process indicators are:

- number of schooldays per year for a particular school type in a particular year;
- depending on formal regulations which require schools to register lesson "drop-out", the percentage of lessons that - for one reason or another - was not given for a particular school type, during a specific period.
These two variables could be used for calculating an indicator of net-time for tuition. The policy relevance of such an indicator would depend on the variability of this indicator over time. It could also play a role in international comparisons of the functioning of national educational systems.

One could also think of constructing a macro-level indicator for the evaluative potential of educational systems. In this respect it would be relevant whether or not a country has a national assessment project and whether or not the results of assessment are fed back to individual schools. At the national level this indicator could be of relevance because of possible changes over time, for instance, when a national assessment project is still in a developmental phase (as is the case in the Netherlands).

A final possibility for a proxy-macro-process indicator could be in the area of educational leadership, for instance, by using a measure of the total effort (time, money, enrolments) of management training courses directed at stimulating this. It should be noted, however, that such an indicator would require rather detailed content-analysis of existing management training courses and therefore would probably not fit the requirement of easy measurement.

The conclusion is that process indicators only have a limited place within pure macro level indicator systems. If, however, one would consider patching up these indicator systems by means of in-depth studies of process variables that could be linked to regular data streams, there is a tremendous increase in possibilities. In-depth studies could take the form of comparative international surveys, specific research projects at the national level, or audits by expert committies or the inspectorate. In these studies all process variables mentioned in Figure 2 could be included, be it by means of scales and questionnaire items or by means of more open check-lists for observation or content-analysis.

School Monitoring at the District Level

District-level monitoring of schools could benefit from systematic data collection as a basis for a limited set of indicators. For such management information systems, as for the previously discussed application of indicators, low inference measures would be preferable. The most likely level of aggregation of the data would be school level, though teacher- and student level data might also be used. Financial and output data (both in terms of overall productivity and student achievement) are the most
important data categories for such monitoring systems. Examples of process variables that could be included are:

- Whether or not schools use explicit achievement standards (to be deduced from school curricula, development plans or other official documents);
- Whether or not school leaders refer to achievement records in official documents and brochures intended for the school community;
- The amount of non-teaching time that is spent on collaborative planning of curricula and lessons by the staff;
- Number of students expelled from school, in a particular school year;
- Whether or not headteachers formally evaluate staff on a regular basis;
- The frequency of the use of achievement tests at all grade levels;
- Figures on pupil absenteeism and lesson drop-out.

The above variables are assumed to be available from administrative school records or assessed by means of relatively brief encounters with headteachers. The process variables could be enlarged if research-like data collection was added at regular intervals.

School Self-Evaluation

One might say that the two applications of indicator systems discussed in the above are accountability oriented rather than improvement oriented. In my opinion, the distinction between evaluating for accountability vs. evaluating for improvement should not be drawn too sharp, since there is an important common element in both: the element of learning by means of empirical test and feedback of information to relevant actors. However, the third kind of application of educational indicators that will be referred to here, is most strongly associated with improvement (cf. Hopkins & Leask, 1989). Although output indicators deserve a central place in self-evaluation procedures like in all other applications of educational indicators, in procedures for self-evaluation or school-based review, process variables are of particular relevance. Since organizational learning is the key motive for using self-evaluation schedules for school improvement, one will need rather detailed information on those processes that might explain disappointing results and at the same time offer handles for improvement. The list of process variables, based on the results of school effectiveness results, might be used to review existing schemes and check-lists for
school-based review (for an overview of these check-lists, see Hopkins, 1987).

One might conceive of integrating these macro-, meso- and micro-level applications of indicators. In fact, the US educational data redesign project (Teaubert, 1987) does exactly this. In this way, detailed information is collected at the lowest unit (the classroom), and part of this information is aggregated to be used at the next level up, and so on. Although such an integrated multiple context, multiple level and multiple measurement indicator systems have important advantages, it is a very ambitious endeavour. One could also conceive of looser coupling of macro-meso and micro applications, for instance by exchanging instruments between the various contexts of application.

Conclusion

Process indicators are a somewhat difficult category among other, more established types of indicators. The main difficulty is that process indicators usually require rather complicated procedures of data collection and measurement and thus do not fit the requirement of readily available data for computing indicators. Yet, there is an important motive to try and include process information in indicator systems. Like data on resources and inputs, process data provide background material that is helpful in making sense out of mere performance data on the functioning of educational systems.

In this paper the literature on school and instructional effectiveness has been used to suggest likely candidates for process indicators. At the same time the scope for discussing the application of process indicators was enlarged by also considering indicator applications at administrative levels below that of national governments. In Table 4 the list of variables that is proposed here as a basis for the development of process indicators is compared to other lists of process indicators that have been proposed.

As appears from Table 4, there is considerable agreement between the proposal that is made here and those by Oakes, Benveniste and Teaubert. The proposals by Unesco and Windham are stated in more general terms. The one factor that is included in all of the proposals summarized in Table 4 is instructional time.
Table 4: Comparison of Sets of Process Indicators

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement stimulants</td>
<td>Allocation of resources</td>
<td>Instructional organization</td>
</tr>
<tr>
<td>Achievement oriented</td>
<td>Retention &amp; progression rates</td>
<td>Alternative technologies</td>
</tr>
<tr>
<td>policy</td>
<td></td>
<td>Use of teacher and student time</td>
</tr>
<tr>
<td>Educational leadership</td>
<td>Teacher/hours per pupil per year</td>
<td></td>
</tr>
<tr>
<td>Teachers' cooperative</td>
<td>Cost and management</td>
<td></td>
</tr>
<tr>
<td>planning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of curriculum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluative potential</td>
<td>Teuber, 1987</td>
<td></td>
</tr>
<tr>
<td>Orderly climate</td>
<td>Instructional leadership</td>
<td></td>
</tr>
<tr>
<td>Time on task</td>
<td>curricular</td>
<td></td>
</tr>
<tr>
<td>Structured teaching</td>
<td>type of instruction (whole class, small group, etc.)</td>
<td></td>
</tr>
<tr>
<td>Opportunity to learn</td>
<td>Time on task</td>
<td></td>
</tr>
<tr>
<td>High expectations</td>
<td>School climate</td>
<td></td>
</tr>
<tr>
<td>Monitoring progress</td>
<td>Influence of peer group</td>
<td></td>
</tr>
<tr>
<td>Reinforcement</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Benveniste, 1987</th>
<th>Oakes, 1987</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher time (teaching/non-teaching)</td>
<td>Access to knowledge (e.g. instructional time)</td>
</tr>
<tr>
<td>Student learning time:</td>
<td>Press for achievement (e.g. graduation requirements)</td>
</tr>
<tr>
<td>- course enrolment</td>
<td>Professional conditions for teaching (e.g. time spent on collaborative planning)</td>
</tr>
<tr>
<td>- turnover rates</td>
<td></td>
</tr>
<tr>
<td>- pupil/teacher ratios</td>
<td></td>
</tr>
<tr>
<td>- school day activities</td>
<td></td>
</tr>
<tr>
<td>- length of school year</td>
<td></td>
</tr>
<tr>
<td>- out of school learning time</td>
<td></td>
</tr>
<tr>
<td>Order and consistency:</td>
<td></td>
</tr>
<tr>
<td>- truancy, absenteeism, vandalism, disruptions</td>
<td></td>
</tr>
<tr>
<td>- student turnover</td>
<td></td>
</tr>
<tr>
<td>- student cooperative behaviour</td>
<td></td>
</tr>
</tbody>
</table>

The categories of variables that distinguish the proposal that is made here from the others are: achievement stimulants from the school environment, structured teaching and the evaluative potential of the school.
References


The Author

JAAP SCHEERENS is a professor of education at the University of Twente. He is presently chairman of the educational administration division and director of the Centre for Applied Research in Education. His research interests are in the field of organizational effectiveness and educational evaluation.