Theories on educational effectiveness and ineffectiveness

Jaap Scheerens*

Department of Educational Organisation and Management, University of Twente, Enschede, The Netherlands

Following Snow's (1973) description of an "inductive" process of theory formation, this article addresses the organization of the knowledge base on school effectiveness. A multilevel presentation stimulated the conceptualization of educational effectiveness as an integration of system-level, school-level, and classroom-level factors. Next steps in theory formation, based on the formation of broader constructs and eclectic use of available theories, are considered in a confrontation with different appreciations of the strength of the current knowledge base. It is concluded that a range of different interpretations of the rationality paradigm (synoptic planning, contingency theory, creating market mechanisms, and cybernetics) does rather well in explaining both positive and negative outcomes. Alternative theories like loose coupling and self-organizing enlarge the scope of potentially effectiveness-enhancing factors as well as variables and mechanisms associated with ineffectiveness.

Keywords: educational effectiveness; theory formation; rationality paradigm; loose coupling; complexity theory

Introduction

Why bother about theory on educational effectiveness? The introductory note in the Society for Research on Educational Effectiveness (SREE) 2011 conference program explains the purpose of theory rather well:

> In addition to informing current practice and policy, research in education should support the development of explanatory and predictive theories of educational processes and mechanisms. Education research must answer questions about why, how, under what circumstances, and for whom, education practices and policies affect individual outcomes. Without an evidence-based theory of educational processes and mechanisms, pragmatic evidence of effectiveness may not be generalizable to new settings or different populations. (https://www.sree.org/conferences/2011/)

In this article, the state of affairs with respect to theory formation on educational effectiveness will be assessed, and confronted with the empirical knowledge base, as a first step in a more theory-driven reconstruction of the field. Although effectiveness is seen in a multilevel framework, which integrates system, school, and teaching effectiveness, the emphasis is on the school level and organizational theory.

*Email: j.scheerens@utwente.nl

What does theorizing about educational effectiveness mean?

In his seminal paper about theory construction for research on teaching, Snow (1973) discusses theory construction as a gradual process, evolving from formative hypotheses about empirical regularities to axiomatic theories. The sequence is inductive, the process starts with empirically verified facts and hypotheses and generalizations being developed from these elements. Subsequent stages are *formative hypotheses, elementism* (reducing the definition of variables to the most elementary units possible), *descriptive theories and taxonomies, conceptual theories and constructs* (including procedures for construct validation), *broken axiomatic theories* (this may involve eclecticism, in other words, borrowing from several more established theories), and, finally, as the highest form, *axiomatic theory,* described as having a set of primitives with the help of which all its remaining concepts can be deduced and all the remaining statements can be derived as consequences (Snow, 1973, p. 83).

As far as school effectiveness is concerned, its *formative hypotheses* had both a more research-based and a practical background. The scientific basis for its formative hypotheses arose as a reaction to the outcomes of the well-known Coleman report (Coleman et al., 1966), taking up the challenge that schools did matter. The practical basis was enhancing the quality of schooling, particularly for disadvantaged students. School effectiveness research and its implementation branch of school improvement has retained this dual basis, on the one hand inquiry oriented and using scientific methods, on the other hand a movement on furthering quality and equity in education. The effectiveness concept depends on establishing means-goals associations, which can be seen as formally analyzable as cause (means) and effects (attained goals) analysis. The fact that effectiveness can be placed as an important facet of educational quality (Scheerens, Luyten, & Van Ravens, 2011) underlines the normative context of the work and, as far as the research approach is concerned, makes for a close resemblance to evaluation research.

Elementism, the next higher up phase of theory development, according to Snow (1973), is about the development of key concepts and instruments for the field of study. This appears to be a relatively underdeveloped area in school effectiveness research, featuring few established instruments (exceptions are perhaps the instrument development on educational leadership [Hallinger, 1984] and recent work within the framework of the dynamic model by Creemers & Kyriakides, 2008).

Descriptive theories and taxonomies have been presented for integrated multilevel models of educational effectiveness since the early 1990s (Creemers, 1994; Scheerens, 1992; Stringfield & Slavin, 1992). More recently, the dynamic model of educational effectiveness by Creemers and Kyriakides (2008) has become a source of inspiration for empirical research. In the next section, a further analysis of the components and relationships within such descriptive models will be given, based on the work carried out for the conceptual framework of the Programme for International Student Assessment (PISA) 2009 (Scheerens, 2007).

Conceptual theories and constructs bring operational variables on a higher level of abstraction. When carrying out meta-analyses, one has to make decisions about uniting more specific variables under a more general label, as opposed to using several more partial and specific factors; for example, to unite transformational and instructional leadership under one general leadership label (Scheerens, 2012). Another example is "focused instruction" as an overarching construct for constructivist and direct teaching (Louis, Dretzke, & Wahlstrom, 2010). Doing so might be motivated by the finding that both strategies are about equally effective (Cobern, Schuster, & Adams, 2010; Scheerens,

Luyten, Steen, & Luyten-de Thouars, 2007). Using more abstract overarching constructs could be seen as one of the possible answers to plurifinality, the fact that two distinct measures appear to be equally effective, assuming that they are just specific instances of a more general approach, and suggesting a more parsimonious underlying principle. In the case of focused teaching, this might be a clear consciousness of applying a mixture of more structured teaching and independent learning methods in a teaching session.

At the next level up, Snow (1973) distinguishes *broken axiomatic theories.* In this paper, only one facet of this level of theory development will be discussed, namely *eclecticism,* as this is the only facet that seems to apply to educational effectiveness research so far. In a subsequent section, results of a recent empirical review of the use of theory in school effectiveness research will be summarized (Nordenbo et al., 2009; Scheerens, 2013b).

As Snow (1973) concludes with respect to research on teaching, so can we conclude for educational effectiveness that the stage of *axiomatic theory* has not been reached.

When discussing the process of theory and model development, two other terms may come up, those of meta-theories and paradigms. According to Snow (1973), meta-theories are concerned with the development, investigation, or description of theory itself. Examples are specific methodologies (e.g., randomized field trials) or families of theories. The term paradigm has an even broader scope, and is used more retrospectively in the sense that it is only applied to the analysis and evaluation of theories after their construction. In the main section of this paper, the rationality paradigm will be used as a family of relevant theories, and it will be placed next to an orientation that (according to the current author) has not yet developed into a well-articulated paradigm but unites ideas about loose coupling and self-organization.

The conceptual structure of educational effectiveness as a hierarchical system

In Figure 1, influence across levels is indicated by the dotted arrows that run from higher levels to lower levels. Such across-level relationships can be interpreted in terms of control, facilitation, and buffering from a higher level directed at the core process at the next lower level. Depicting education in this way and qualifying the overall system as hierarchical and loosely coupled has the following implications:

- Lower level core processes are seen as being contextualized and controlled by higher levels (the vertical aspect).
- Despite this notion of higher level control, lower levels are seen as having considerable discretion over their core processes, in other words, considerable autonomy. This is the idea of loose coupling between hierarchical levels, sometimes expressed in more prescriptive terms, like "subsidiarity"; a maxim which states that lower level autonomy should be maximized up to the point beyond which it would become counterproductive. Put differently, this approach would imply that what can be reasonably accomplished at a lower level should not be carried out by a higher level.

The degree of higher level control versus lower level autonomy is an issue of central importance at all levels. At system level, it is about effective patterns of functional decentralization, which means that, perhaps dependent on the larger context, certain patterns of centralization in some functional domain (e.g., the curriculum) and decentralization in another domain (e.g., financial management) work best. At school level, it is about the degree of participative decision making, or "distributed leadership", and at classroom level it refers to the balance between strongly structured didactic approaches
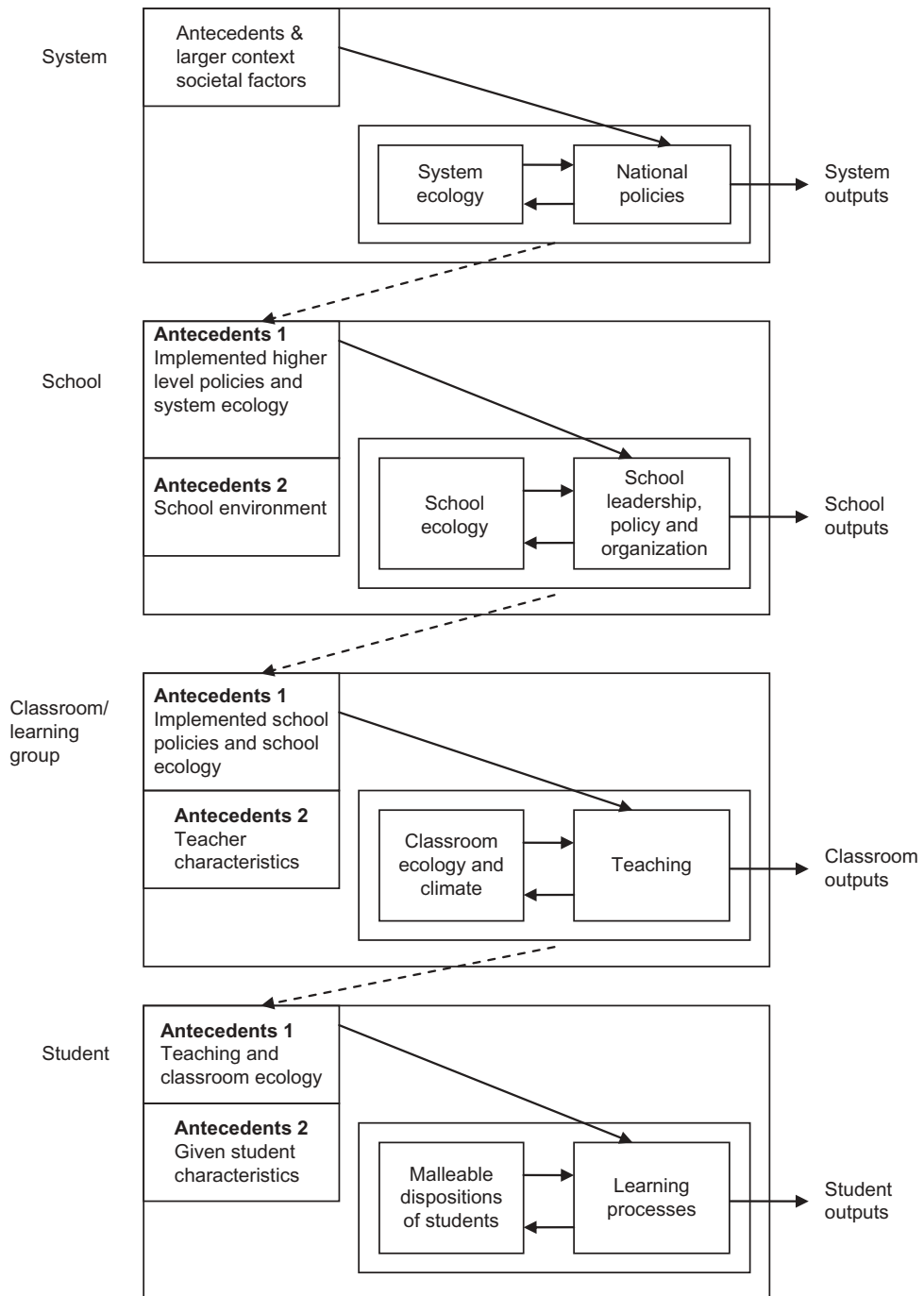
Figure 1.   Integrated multilevel model of education.
Note: The dotted arrows from one system level to the next represent across-level influences;
feedback loops are assumed to run from outcomes at each level to the box containing ecological
conditions and active policies at each object level and from lower to higher levels, but these are not
shown, to avoid a too complex pattern of arrows (for a detailed description, see text).

and more open teaching and learning situations that are expected to invite self-regulated learning. Structure versus independence is a red line that dominates policy and research agendas in education. A second key element in the representation in Figure 1 is the identification of ecological conditions as a separate class of conditions influencing educational performance. This is done by giving a more explicit place to partially controllable composition effects, and their interaction with more directly malleable variables, such as the school climate. The recognition of this kind of contextual conditions emphasizes the partiality of direct control in education, and in this way underlines the loose coupling between the hierarchical levels, but at the same time focuses the attention on a qualitatively different strand of control measures, namely, those of selection, admission, grouping, and matching of teachers and subgroups of students, as well as on cultural aspects associated with student and teacher body composition.

It is interesting that, as a thematic report on PISA showed (Luyten et al., 2005), in some countries the interaction between average socioeconomic status (SES) and "good" teaching conditions is much stronger than in other countries; implying that these latter countries do better in creating equitable conditions in education. Figure 2, from Scheerens (2007), illustrates how this empty framework can be used as a basis for categorizing variables that have been addressed in empirical research, in this case, school effectiveness research.

The way theory is being used in an eclectic way in school effectiveness research

This section refers to a recent study by the author on the use of theory in educational effectiveness research (Scheerens, 2013c). This work was part of a larger review of school

School antecedents	School ecology	School leadership, policies and organization
Implemented higher level policies • accountability and evaluation demands • experienced school autonomy *External school environment* • affluence of the school neighborhood	- student composition (e.g., school average SES) - percentage of students from immigrant background - percentage of students with a special education profile - teacher composition (e.g., the average qualification, experience, and locus of control scores of the teachers) - stability of teaching staff over the last 5 years - school climate x school composition interaction - level of school material resourcing - level of extracurricular activities of the school	- leadership focus, specifically the degree of instruction-oriented leadership - achievement orientation/ high expectations - teaching time - quality of school curriculum, opportunity to learn - coordination, cooperation, and consensus among staff - a safe and orderly climate, supporting positively experienced interactions - parental involvement - opportunities for professional development of teachers - school admission policies - evaluation, feedback, and monitoring - degree to which teachers participate in decision making

Figure 2.   Variables in a model of school functioning.

Table 1. Overview of studies in which more established theory was used, from Scheerens (2013c).

Reference	Theory	Country
Coates (2003)	Micro-economic theory	USA
Griffith (2003)	Quinn and Rohrbaugh model	USA
Hofman, Hofman, Guldemond, and Dijkstra (1996)	Coleman's functional community theory	Netherlands
Hoy, Tarter, and Bliss (1990)	Parson's social systems' theory	USA
Kyriakides, Campbell, and Gagatsis (2000)	Creemers comprehensive model	Cyprus
Kyriakides and Creemers (2008)	Dynamic model of educational effectiveness	Cyprus
Kyriakides and Tsangaridou (2008)	Creemers comprehensive model	Cyprus
Reezigt, Guldemond, and Creemers (1999)	Carroll model, Creemers model	Netherlands
Stringfield, Reynolds, and Schaffer (2008)	Schools as high reliability organizations	USA/UK
Tarter and Hoy (2004)	Bolman and Deal and Hoy and Miskell as theoretical bases	USA
Van der Werf (1997)	Creemers' comprehensive model	Netherlands

effectiveness research (Nordenbo et al., 2009). Only the main results will be referred to. The main conclusion based on an international review of 109 school effectiveness research studies was that only 6 could be seen as theory driven. This number could be, somewhat arbitrarily, raised to 11, by including those studies that were based on models that made reference to specific broader conceptual principles. These 11 studies are listed in Table 1.

A striking outcome was the fact that of the 11 more or less theory-driven studies 5 are based on the models by Creemers (1994) and Creemers and Kyriakides (2008). The overall conclusion from this study was that only a small majority of school effectiveness studies was more or less driven by existing theories.

Before attempting a systematic approach in discussing implications from two "meta-theories", the rationality paradigm, on the one hand, and a second paradigm for which the phrase "transformative teleology" (Stacey, Griffin, & Shaw, 2000) might be used, a brief excursion to the educational effectiveness knowledge base will be made. This excursion prompts an approach that looks at positive as well as negative instances of educational effectiveness, or rather effectiveness-enhancing conditions next to effectiveness-deteriorating conditions.

Cognitive dissonance about the knowledge base

More recently, several meta-analyses of educational effectiveness research studies have been carried out. In Table 2, the effect sizes for a selection of school-level and teaching-level variables are summarized from meta-analyses by Scheerens et al. (2007), Seidel and Shavelson (2007), Creemers and Kyriakides (2008), and Hattie (2009). It should be noted that the evidence base in the work of John Hattie is much more expansive, as it is a synthesis of no less than about 800 individual meta-analyses. In Table 2, the coefficients in bold are effect sizes, computed as the standardized mean between a treated and a control group, coefficient *d*. The other meta-analyses have used Fisher's *z*, a measure that is comparable to a product moment correlation coefficient. The two coefficients can be converted to one another, where *d* has roughly twice the size of the correlation coefficient.

Table 2. Results from recent meta-analyses (coefficients are based on the Fisher *z* transformation of correlations; as Hattie presents effect sizes in terms of *d*, these are indicated in bold).

School-level variables

	Scheerens et al. (2007)	Hattie (2009)	Creemers and Kyriakides (2008)
Consensus & Cohesion	.02	–	.16
Orderly climate	.13	**.34**	.12
Monitoring & evaluation	.06	**.64**	.18
Curriculum/opportunity to learn	.15	–	.15
Homework	.07	**.30**	–
Effective Learning Time	.15	**.34**	–
Parental involvement	.09	**.50**	–
Achievement orientation	.14	–	–
Educational leadership	.05	**.36**	.07
Differentiation	.02	**.18**	–

Teaching-level variables

	Scheerens et al. (2007)	Hattie (2009)	Seidel and Shavelson (2007)
Time and opportunity to learn	.08	**.34**	.03
Classroom management	.10	**.52**	.00
Structured teaching	.09	**.60**	.02
Teaching learning strategies	.22	**.70**	.22
Feedback & monitoring	.07	**.66**	.01

Even if the coefficients in the tables would be thus converted, it would still be clear that the coefficients reported by Hattie are considerably higher than those from the other meta-analyses.

According to the established conventions (Cohen, 1988), Hattie's (2009) results would be interpreted as medium-sized effects, while those found by the other authors cited as low to negligible.[1]

In recent reviews, educational effectiveness researchers take for granted that the results of educational effectiveness and school improvement research provide a solid knowledge base (Hopkins, Stringfield, Harris, Stoll, & Mackay, 2011; Muijs et al., 2011; Reynolds et al., 2011). Hattie's results are impressive. Instead of worrying that "nothing works", he expresses himself as being concerned that "everything works" (Hattie, 2009).

On the other hand, when looking closely at the research evidence, one usually finds large differences in mean effect sizes between meta-analyses. In recently conducted meta-analyses on educational leadership and instruction time, we found very small effect sizes (Scheerens, 2012; Scheerens, 2013a).

Effect sizes that may be small, large percentages of unexplained variance, most of the variance explained by individual or aggregate student background characteristics, little generalizability of the established set of malleable factors across countries, internationally relatively small changes in performance results *and* in the malleable factors that are supposed to explain them, all these considerations urge for an expedition to explore the "dark side of the moon" in educational effectiveness. When it comes to theory, this would involve looking for mechanisms that might explain not only effectiveness but also ineffectiveness.

In the remaining part of this paper, two families of theories will be discussed; the first consists of interpretation of the rationality paradigm, the second of "alternative theories", based on principles of loose coupling and self-organization.

The rationality paradigm and its fit to the educational effectiveness knowledge base

The very concept of educational effectiveness is based on the rational idea of optimal goal attainment. The factors "that work" can be seen as effective means to reach educational goals. This basic idea can be formalized by describing education as a contextualized production process, using the well-known context-input-process-output model, which is at the basis of the educational effectiveness models discussed in an earlier section. Educational effectiveness research is particularly interested in input and process indicators that are associated with outcome indicators that represent "realized goals". Pointing at this fundamental connection of the concept of educational effectiveness with the rationality paradigm is only the first step in using it as an explanatory basis in addressing the question "why" certain factors appear to work. The second, and for our purposes more important, step is that different interpretations of the rationality paradigm indicate different explanatory mechanisms.

The basic principles of the rationality paradigm are the following: Behavior is oriented towards preferred end states (such as realized goals and personal well-being), and optimal choice is made between alternative ways to reach the goals.

Different interpretations of the rationality paradigm are: synoptic planning, contingency theory, cybernetics, and public choice theory. Each of these interpretations emphasizes certain key processes, but has imperatives for organizational structuring as well.

Synoptic planning and bureaucratic structuring

The ideal of "synoptic" planning is to conceptualize a broad spectrum of long-term goals and possible means to attain these goals. Scientific knowledge about instrumental relationships is thought to play an important role in the selection of alternatives. Campbell's (1969) notion of "reforms as experiments" combines a rational planning approach to social (e.g., educational) innovation with the scientific approach of (quasi-) experimentation.

A modern interpretation is the concept of high reliability organizations. This concept has been successfully applied to schools (cf. Bellamy, 2011; Reynolds, Stringfield, & Schaffer, 2006; Stringfield, Reynolds, & Schaffer, 2011).

Contingency theory

In organizational theory, contingency theory has as its central thesis that the effectiveness of organizations depends on certain more basic and contextual conditions. This is often expressed by saying that there is no universal best way to organize; success depends on a good fit between internal organizational characteristics, and between internal arrangements and environmental conditions (cf. Kieser & Kubicek, 1977; Lawrence & Lorsch, 1967; Mintzberg, 1979). Contingency theory can be seen as an extension of rational planning and structuring, since the ideal of optimizing remains, be it conditional on contextual conditions, which, by the way, complicates the analysis considerably (Kickert, 1979). The internal alignment of organizational conditions in contingency theory is known as the "configuration hypotheses". An application in educational effectiveness

and school improvement are comprehensive school improvement programs (Borman, Hewes, Overman, & Brown, 2003; Borman, Carter, Aladjem, & LeFloch, 2004). Such programs combine an "evidence based" rational planning approach to implementation, with a coordinated set of levers for improvement, such as teaching strategies, curricular emphases, leadership, and cooperation.

Cybernetics

The third interpretation of the rationality paradigm can be metaphorically labeled as "cybernetics". Cybernetics is described as the transdisciplinary approach for exploring regulatory systems. The key mechanism consists of a sequence of evaluation, feedback, and corrective action; which can be thought of as a cycle.

The practice of using evaluative information on organizational functioning as a basis for corrective or improvement-oriented action can be seen as a less demanding kind of regulation than proactive synoptic planning. In the former case, planning is likely to have a more "step by step", incremental orientation, and "goals" or expectations get the function of standards for interpreting evaluative information. The discrepancy between actual achievement and expectations creates the dynamics that could eventually lead to more effectiveness.

Public choice theory

A central assumption in the synoptic planning and bureaucracy interpretation of the rationality paradigm is that organizations act as integrated purposeful units. Individual efforts are expected to be jointly directed at the attainment of organizational goals. In the so-called political image of organizations (Morgan, 1986, Chapter 6), this assumption is rejected, emphasizing that "organizational goals may be rational for some people's interests, but not for others" (Morgan, 1986, p. 195). The fact that educational organizations consist of relatively autonomous professionals, and loosely coupled subsystems, is seen as a general condition that stimulates political behavior of the members of the organization.

In public choice theory (Niskanen, 1971), the lack of effective control from democratically elected bodies over public sector organizations marks these organizations as being particularly prone to inefficient behavior, essentially caused by the leeway that is given to managers and officers to pursue their own goals besides serving their organization's primary mission.[2] Creating competition and market mechanisms are seen as the remedy to overcome these problems. The alleged superiority of private over public schools is the most supportive piece of empirical effectiveness research for the claims of public choice theory, although the significance of the results in question is much debated (Goldhaber, 2000; Gorard, Fitz, & Taylor, 2001; Organisation for Economic Co-operation and Development [OECD], 2005; Scheerens, 1992). A similar observation could be made with respect to decentralization and enhanced school autonomy as a lever of educational performance (Luyten et al., 2005; Scheerens & Maslowski, 2008).

An overview of the four different interpretations of the rationality paradigm and their key mechanisms is given in Figure 3.

As the overview shows, these four interpretations of the rationality paradigm and their corresponding key mechanisms can be matched with important levers of educational performance such as proactive planning approaches, highly formalized organizational structures like schools as high reliability organizations, comprehensive school reform

Interpretation of the rationality paradigm	Mechanism
Synoptic planning	Proactive structuring
Contingency theory	Fit
Cybernetics	Evaluation and feedback
Public choice theory	Market mechanisms; school competition, choice

Figure 3.   Interpretations of the rationality paradigm.

projects, all kinds of applications of evaluation and assessment, including accountability policies and data use, and measures to make schools more autonomous and private and to stimulate free school choice. As shown on earlier occasions (Scheerens, 1997), matching can also be carried out in a more detailed way, by relating key variables reflecting effectiveness-enhancing conditions to each of the rationality perspectives (see Figure 4)[3].

What rational models have to say about educational ineffectiveness

Empirical studies suggest that the antipodes of effectiveness-enhancing factors are associated with school failure. Stringfield (1998) mentions "lack of academic focus", "academic periods starting late and ending early", "bureaucratic leadership", and "lack of teacher assessment" as characteristics of failing schools.

These examples show negative effects of "too little" of the factors that are associated with rationality interpretations. At the same time, there are extreme implementations of rationality models, which might be globally indicated as "too much" of them.

In the concept of schools as high reliability organizations, there is a thin line between an acceptable degree of standardization and rigorously monitored "standardized operating procedures SOP's" as the favored focus of quality management systems (ISO and others). These latter practices cause a lot of "red tape" and might lead to goal displacement; teaching to the SOP would seem to be worse than teaching to the test!

In the domain of evaluation and feedback, a whole literature exists about the negative side effects of high-stakes testing and intensive external school inspection, like tunnel vision, all kinds of strategic behavior, and administrative burden (cf. Ehren, 2007). In less contested areas as school self-evaluation and the image of schools as learning organizations, resistance, immunization against potential criticism, and barriers to organizational learning have been noted as well. As part of their theory of "organizational learning", Argyris and Schön (1978) recognize limits to organizational learning. These limits reside in behavioral patterns of the members of the organization. They describe these as "shared strategies in individual theories in use". As examples, they mention considerations like:

- Let buried failures lie.
- Keep your views of sensitive issues private; enforce the taboo against their public discussion.
- Do not surface and test differences in views of organizational problems.
- Avoid seeing the whole picture; allow maps of the problem to remain scattered, vague, ambiguous (pp. 39−40).

They also mention that such strategies may reflect deeper and more fundamental norms, strategies, and assumptions:

Rationality interpretation	Global intervention strategies	Variables addressed in research	Side effects and Exaggerations, associated with ineffectiveness
Synoptic planning and bureaucratic structuring	Proactive planning Evidence-based policies Schools as high reliability organizations	<u>System level</u> - national curriculum planning - infrastructure for program evaluation in education - centralized structures, limited school autonomy - school inspection <u>School level</u> - instructional leadership - consistency of practice - academic emphasis and achievement orientation - clear and fair discipline <u>Classroom level</u> - time and opportunity to learn - structured lessons - monitoring and feedback	Standardized operating procedures in teaching Goal displacement "Red tape" Lack of flexibility and innovation
Contingency theory	Comprehensive school reforms Differential effectiveness	<u>System level</u> - vertically aligned educational systems <u>School level</u> - transformational Leadership <u>Classroom level</u> - adaptive instruction - differentiation	
Cybernetics	Accountability Organizational learning Education-based meritocracy	<u>System level</u> - a broad range of accountability provisions - national assessments - school inspection - new public management: free processes control outcomes - (institutionalization of) school self-evaluation - (institutionalization of) school development planning - Facilities for continuous professional development of teachers <u>School level</u> - school self-evaluation - school performance feedback <u>Classroom level</u> - pupil monitoring systems	Negative side effects of high-stakes testing Resistance to assessment and evaluation Factors preventing organizational learning Evaluation apprehension
Public choice theory	Free school choice Privatization School autonomy Competition	<u>System level</u> - free school choice - financial and managerial school autonomy - privatization - vouchers - "high-stakes" accountability arrangements <u>School level</u> - merit pay of teachers - educational entrepreneurship <u>Classroom level</u> - stimulating extrinsic motivation - efficient class size - matching teachers and students - optimizing class size	Off-task behavior Political processes "Make work" Exaggerated managerial overhead

Figure 4.   Variables addressed in educational effectiveness research matched to four alternative interpretations of the rationality paradigm, as well as associated sources of ineffectiveness.

- Protect yourself unilaterally – by avoiding both direct interpersonal confrontation and public discussion of sensitive issues which might expose you to blame.
- Protect others unilaterally – by avoiding the testing of assumptions where that testing might evoke negative feelings, and by keeping others from exposure to blame.
- Control the situation and the task – by making up your own mind about the problem[4] and acting on your own view, by keeping your view private, and by avoiding the public inquiry which might refute your view. (p. 40)

Public choice theory and its foundation micro-economic theory have off-task behavior and selfish motives of organizational members as one of its foundational pillars. One could even say that the theory's basis is negative, in the sense of preventing dysfunctional organizational functioning next to positively oriented to rational management. Market mechanisms, including competition and effective assessment strategies, are invented to drive such dysfunctional behaviors out. There is correspondence with the phenomena mentioned by Argyris and Schön (1978) under the heading of limits to organizational learning, as micro-economic theory takes into consideration protective, defensive, and egoistic reactions by members of the organization. A summary overview of the specification of the various interpretations of the rationality model, including references to ineffectiveness, is presented in Figure 4.

Alternative theories for explaining effectiveness and ineffectiveness

The interpretations from the rationality paradigm discussed in the above go a long way in explaining the findings from educational effectiveness research. It is conceivable to propose a theory-driven research agenda, in which the effectiveness of pure forms of the models, and more partial derivatives, could be tested, and where the alternatives could be pitted against one another. In this, contingency theory is a case apart, since it could be used as a meta-theory to investigate in which situations one of the three other models would work best. Implementation failures, exaggerated interpretations, and undesired side effects would appear to be plausible explanations for ineffectiveness.

For several reasons, it is still considered useful to go beyond the analysis of positive and negative instances of the rationality paradigm and look at alternative theories. These are the following:

(1) The applications of the rationality paradigm will tend to be focused at the productivity of the operational core of the organization; and organizational structural conditions, including links to the environment; this would mean that other organizational functions, such as responsiveness to new developments in the environment, reflection on aims and goals, and cultural facets of the organization, get less emphasis.
(2) Rationality applications would seem to have a stronger preoccupation with instrumental and technological levers for improvement than with improving cultural conditions and motivational aspects of work; human relations and recruitment policies.
(3) Rationality applications are closer to proactive and retroactive planning and structuring modes of organizational functioning than to implementation.
(4) Rationality applications would tend to focus on the formal organization and have less attention for the informal organization.

Two alternative perspectives will be discussed that might provide a different outlook on educational effectiveness and ineffectiveness: schools as loosely coupled organizations and theories about self-organization. The question whether loose coupling provides an alternative perspective to interpretations of the rationality paradigm is open for debate. It is also associated with systems theory, incrementalism, and "bounded rationality" (Lindblom, 1959; Simon, 1964); still part of the rational paradigm. At the same time, it offers quite different orientations towards educational effectiveness and ineffectiveness, as well as certain prescriptive interpretations regarding organizational change. Self-organization is used as the central term to refer to complex interactions in organizations that emphasize "emergence" rather than control.

Schools as loosely coupled systems

Weick (1976) describes "loose coupling" as an "image that coupled elements are responsive, but that each event also preserves its own identity and some evidence of its physical and logical separateness" (p. 1). He goes on to say that "Loose coupling also carries connotations of impermanence, dissolvability and tacitness all of which are potentially crucial properties of the 'glue' that holds organizations together" (p. 1). In educational organizations, two of the most fundamental couplings, the one among elements of the core technology, and the other between the authority structure on the one hand and the functioning of the technical core on the other, are not particularly prominent.

More specific combinations of elements in educational organizations between which loose rather than tight coupling is likely to occur are:

- intentions and actions;
- yesterday and tomorrow;
- top and bottom;
- line and staff;
- administrators and professionals;
- several means leading to the same end (plurifinality);
- teachers and materials;
- voters and the school-board;
- parent and teacher;
- teacher and pupil.

Schools as loosely coupled organizations would seem to be at odds with educational effectiveness models that depend on "rational planning" type of mechanisms. The concept of *plurifinality* (Von Bertalanffy, 1968) alone seems to present fundamental problems for the causal implications of the very concept of educational effectiveness. Loose coupling could be read as an explanation for the relatively low and inconsistent effects for core factors like coordination and consensus, educational leadership, and evaluation and feedback. Yet, according to Weick (1976), loose coupling has certain advantages. Tight and loose coupling are to be seen as forming a continuum, where weaker couplings, like shared conceptual anticipations and retrospections, may nevertheless create a certain robustness and resilience of the organization, because they contain mutations, localized adaptations, and fewer costs of coordination (p. 14).

To the extent that the educational effectiveness research orientation and knowledge base emphasizes the rationality paradigm, loose coupling offers an explanation for its mediocre success: why some factors do not work optimally, discrepancies of effect sizes,

instability of the research findings, and basic doubts about the generalizability of the research outcomes. The analysis of implementation problems with evaluation and feedback procedures, from the perspective of loose coupling, for example, provides a conceptual basis for underutilization, misuse, and undesired side effects, in applications like school self-evaluation (Scheerens, 2004) and accountability (Ehren, 2007). To the extent that schools confirm to the properties of loosely coupled systems, low effects of leadership should not come as a surprise. Facets of loose coupling, such as no clarity of intentions, individual members of the organization pursuing different goals, and the notion that several means may lead to the same end, put question marks behind some of the basic assumptions of the educational effectiveness approach.

In his well-known article from 1976 "Educational organizations as loosely coupled systems", Weick presented an outline of a research program, based on further analytic and empirical study of "loose coupling". Elements of this program are developing conceptual tools capable of preserving loosely coupled systems, explicating which elements are available in educational organizations for coupling, development of contextual methodology, the collection of thorough, concrete descriptions of the coupling practices in actual educational organizations, specification of the core technology in educational organizations, probing empirically the functions and dysfunctions associated with loose coupling, and discovering how inhabitants make sense out of loosely coupled worlds (Weick, 1976, p. 18). It would be an interesting question for research on science, why this theory-driven research program does not seem to have made it, despite the pervasiveness and actual relevance. So far, Weick's theory on loose coupling does not seem to have had very much influence on empirical educational effectiveness research.

Self-organization

The "complexity sciences", chaos theory, dissipative structure theory, and the theory on complex adaptive systems (Stacey et al., 2000), look at dynamic interactions between the micro-elements of a system, trying to model these, or to come to grasp with regularities or rules that emerge. Studies of complexity developed in the natural sciences are presented in the work of Maturana and Varela (1980) and Prirogine (1997). Famous examples are the way molecules behave in combustion, the butterfly effect, and the discovery of patterns that are "stable and instable" at the same time, so-called fractals.

Stacey et al. (2000) describe the philosophical background of "normal science" in comparison to various interpretations of complexity theory. According to them, self-organization, when it is more than unfolding already initially given patterns, is able to create novelty and is to be seen as a different kind of causality. The mechanistic image of organizations, as implied in scientific management and the "formative teleology" of general system theory is compared to the more organic idea of "transformative teleology". According to Kauffman (1993), formative causality is self-referential: "the dynamics cause themselves as the system evolves of its own accord to the edge of chaos" (as cited in Stacey et al., 2000, p. 117).

Given our purpose to connect educational effectiveness research and its knowledge base to more established theory, particularly management theory in the broadest sense, it should be noted that, according to some interpretations of complexity theory, we have moved beyond the scientific paradigm and what Stacey et al. (2000) indicate as rationalist teleology. In fact, the efficiency perspective is considered as a far too limited criterion to indicate what organizations are aiming for. Likewise, "normal" causal analyses, or analyses focused at optimizing means–end relationships, are not

considered in these interpretations. Instead, as cited above, self-organization is seen as a different kind of causality, and the ultimate intension of organizational development is sometimes indicated in terms of survival, but also in terms of "expressing identity" (Goodwin, 1994, as cited in Stacey et al., 2000, p. 119). If educational effectiveness research and school improvement would embrace the ideas of the "transformative teleology" interpretation of complexity, this would indeed mean a paradigm shift (cf. Harris et al., 2013; Scheerens, 2013b).

Before looking at social science and educational applications, some further characteristics of the way the dynamic interactions between the micro-elements of a system are seen in applications of complex adaptive systems theory will be briefly reviewed.

Importance of initial conditions

The initial characteristics of the micro-elements that interact are considered of great importance; according to Stacey et al. (2000), the diversity in the elements is a precondition for interactions that may lead to novelty and creativity.

Patterns of stability and instability

The dynamics lead through phases of stability and instability, progressing from one "state space" of equilibrium (also indicated as "attractor") to another, culminating in what is indicated in chaos theory as a "strange attractor". Strange attractors are described in various ways, like "patterns which are repeated, but never exactly the same", and examples are given in the realm of weather prediction and heart rhythms, but no examples for social science or management applications seem to exist.

Nonlinear development

Interaction and change are seen as evolving in a nonlinear way, sometimes allowing for "jumps" and qualitative changes in the system.

Preoccupation with disorder

Creativity is associated with initial states of disorder and diversity. "Not only is the system restless, but its own restlessness allows it no rest. Restlessness about restlessness may increase restlessness", writes Luhmann (1995, p. 50). In terms of change processes, one could say that a permanent state of "unfreezing" is considered good for innovation.

Rules amidst chaos

Despite the previous point, some regularities are expected to emerge from chaos, complexity is not an aim in itself. Examples of regularities and rules that may emerge are provided in the image of "fitness landscapes", clusters of intensive "good" interactions as fitness peaks, and low intensity cluster as valleys (below, a research example will be provided in which these metaphoric concepts are made more concrete).

Non-managed dynamics

The interactive processes are thought of as autonomous, circular, and self-referential. There is no room for an objective observer or for a controller; in social systems, only a participatory role is imaginable.

Applications in the social sciences are sometimes purely metaphoric (e.g., Morgan, 1986) but appear also as researchable conjectures. Stacey et al. (2000) concentrate on the interrelationships between members of an organization, while the sociologist Luhmann (1995) also considers dynamics of structural arrangements.

With respect to organizations, the applications from complexity theory have a focus on what is often indicated as the informal organization, of relatively "hidden" properties such as the deep level of organizational culture (cf. Schein, 1985) and the hidden curriculum. Stacey et al. (2000) describe the importance of this level of functioning as follows: "… systems in organizations can only function if the members weave their day-to-day interactions with each other through and around the rules of the systems they have designed" (p. 59). Though these authors expect positive developments, particularly novelty and creativity from the free interaction between the members of the organization, they are also open to the possibility that dysfunctional results and undesired consequences may occur. Next, as far as application to organizations is concerned, the focus is on describing phenomena in the realm of implementation and enactment, as all kind of management and planning is banned from the research agenda. Phenomena like "ownership" and co-construction fit the ideas on "formative teleology" quite well. In both the theories by Luhmann (1995) and Kauffman (1993), there is attention for the issue of changing the boundaries of the organization with respect to the environment, and networking between subgroups within and outside the organization.

Before reflecting on the implications for theorizing about educational effectiveness, two research applications are briefly sketched. One looks at communication patterns among school governors and principals in the context of the implementation of the US No Child Left Behind project (Daly, Moolenaar, & Carrier, 2011). The other addresses the issue of composition effects (Scheerens, 2004, 2008).

Daly et al. (2011) used complexity theory as a basis to study "how rational assumptions undergirding current reform policies limit our understanding of how policy is enacted through complex social interactions" (p. 3). They studied the implementation of No Child Left Behind policy measures, targeted at improving consistently underperforming schools, and applied longitudinal social network modeling to illustrate how school districts could be conceptualized as complex adaptive systems. The interaction patterns were analyzed in terms of "emergence", "fitness peaks", and feedback.

The authors conclude that the study suggests that policy implementation is a complex endeavor and does not necessarily follow linear, predictable patterns as might be suggested by conventional policy assumptions (Daly et al., 2011, p. 26). As these patterns are considered to be decisive for the uptake of the reform, and may vary across schools, large-scale standardized approaches to educational issues may be inadequate in addressing local problems. Emergence of reciprocity and feedback, on the other hand, is associated with sociocultural learning, and development of "ownership" of the reform. Spontaneously developing centralization (disproportional influence of district leaders) led to disconnection between district and school leadership level and was dysfunctional to the expected growth of interactions between the two levels.

A second application of ideas from complexity science is provided in the interpretation of composition effects by Scheerens (Scheerens, 2004, 2008). Composition

effects can be seen as illustrating the importance of starting conditions at the micro-level, for organizational conditions at a higher up level. In school effectiveness research, student background conditions at individual and aggregate level are often used as control variables. Important technical progress has been made on the measurement and modeling of student background composition effects (Van Damme, De Fraine, Opdenakker, Van Landeghem, & Onghena, 2000). The way composition effects operate has facets of a non-managed process, although deliberate selection and grouping policies may be brought into play to control them. For those who like the, in my view, mystifying charm of the phrase that the whole is more than the sum of its parts, this might be applied to composition effects. A more sober statement could be that composition leads to phenomena at a higher up level with distinct "higher up level" interpretations, not immediately obvious from lower level characteristics. Scheerens (2004, 2008) suggests that a phenomenon like the school culture might be conceptualized as a composition effect based on the aggregation of personality traits of teachers; associating, for example, a culture of "high expectations" with a high proportion of teachers with an externally oriented locus of control. A final way in which composition effects reflect some of the issues that are often discussed in the realm of complexity theory are the nonlinearity and qualitative "jumps" in development and change. Some analysts, for example, have addressed the question of the proportion of special needs students and thresholds for heterogeneity being still manageable by regular teaching provisions, within the framework of inclusive education, and similar critical thresholds concerning the proportion of minority students.

In making up the balance, insights from complexity theory emphasize a number of "positive" phenomena, factors in education that could be seen as enhancing effectiveness and improvement. These could be summarized as stimulating autonomy and decentralization, exploiting and optimizing composition effects, providing space for spontaneous interaction and grouping, and, in the case of externally induced reform, a keen eye for an enactment (or mutual adaptation) perspective on implementation, and "ownership".

On the other hand, the interactions among microlevel elements of the system may also go wrong and lead to power games, the pursuit of egoistic behavior, inertia, and resistance to sensible reform measures. Throughout this paper, attention has been given to policy failures, dysfunctional implementations, and, in short, ineffectiveness. Complexity theory offers another idiom to understand particularly the dynamic process dimension of negative development, but does not provide substantive ideas and explanations about why things go wrong. To the extent that complexity theory is the antagonist of rational planning and an engineering interpretation of education change, it is often positioned as a source of explanation of why such policies do not work, or have only limited success. Such critique could be taken more or less radically. A radical solution would be to see emergence and "self-organizing" as a message to abandon all kind of planning, management, and control. A more modified view would be to think of organizations as having both controlled and self-organizing facets, with attention for both formal and informal aspects of organizing. To me, this seems the preferable and more pragmatic solution. In the application of the complexity sciences to social systems, there seems to be a blind spot as far as the reality of the formal organization is concerned. Perhaps a parallel could be made with analyses of formal and informal authority and power in organizations (Pfeffer, 1978). Even when concentrating on informal power, in this literature it is recognized that those with formal authority enter the arena with lead in their gloves.

Discussion

In this presentation, two extremes of the graded continuum of theory development by Snow (1973) have obtained most of the attention: on the one hand, the "basic" empirically supported knowledge base, consisting of average effect sizes of a fairly consistent set of effectiveness-enhancing conditions and, on the other hand, rather general or meta-theories. A middle part of the continuum, consisting of middle-range theories and verified inter-mediary effect models, has been left largely unexplored. In case a program of theory-oriented research synthesis would get off the ground, it could be an interesting job to accumulate and categorize such middle-level conceptual work in a more systematic way. The usefulness of having charted the domain of meta-theories up front might help in making such an endeavor more focused.

Such preparatory work at the level of formal models and middle-range theories would be a required intermediary step to develop concise theory-based models to drive empirical research and add to theoretical reflection on already existing models (like the Creemers and Kyriakides model). Clarifying key concepts would be a necessary part of this exercise.

Another area that deserves more attention than was possible in the current presentation is methodological reflection. Empirical investigation inspired by the alternative theories, loose coupling, and system dynamics, would require a broad range of methodological approaches, including the analysis of communication and interaction patterns by means of detailed sociometric methods and simulation studies based on dynamic models. Moreover, differences concerning interpretation of causality, in the case of "transformative teleol-ogy", might require epistemological reflection.

Scientific educational effectiveness research would imply that "refutations" should be seen as equally important as the confirmation of "conjectures", to paraphrase Popper. The idealistic urge to reform and improve education could get in the way of a neutral and balanced attention for the impact of malleable factors. In reviews of studies on the effects of after-school programs, authors like Kane (2004) and Valentine et al. (2010, p. 32) suggest that unpublished research and higher quality methods produce significantly smaller treatment differences. Favoring studies that show high impact on educational performance is a not unlikely type of publication bias. The attempt at reviewing the knowledge base on educational effectiveness in this paper, however of necessity limited in scope, points at large differences in the average effect sizes found across meta-analyses and small effects and little generalizability across countries found in international studies. It might be that the malleable conditions most frequently addressed in educational effectiveness research are on the small side, although still educationally relevant. This underlines the interest to know more about the limits of the rational techniques that dominate the educational effectiveness and the education reform agenda. The theories that were described in this paper provide indications about causes of ineffectiveness, in the sense of implementation failures and side effects, and collision between the realities of the formal and informal organization. This information might be used in various ways in educational practice and policy: by actively countering implementation failures and side effects, by fostering more realistic expectations on effects and effect sizes among practi-tioners and policy makers, and by considering alternatives levers for improvement. Weick's (1976) theory of loose coupling shows that foregoing active interference and "letting go" or resorting to socialization mechanisms could occasionally be more effective. His theory underlines the importance of the initial training of teachers and the function-ality of professional autonomy. The "Finnish approach" (Sahlberg, 2011) could be seen as

an alternative lever for educational reform and improvement, very much centered on teacher initial training, esteem, and professional motivation. Complexity theory shows an interest in the informal organization, and elements of unpredictability in the interactions of organization members and the emergence of new patterns of behavior, which could be functional or dysfunctional. Similarly, self-organization could be used as an interpretation of composition effects, and nonlinear developments caused by drastic changes in the composition of student and staff intake.

Notes

1. Cohen (1988) refers to standardized effect sizes d, which are about twice the Fisher's z coefficient; he indicates small effect sizes as about $d = .20$ ($r = .10$) and medium effect sizes as about $d = .50$ ($r + .25$).
2. A more extensive treatment of the implications of public choice theory for school effectiveness research is given in Scheerens (1992, Chapter 2).
3. More operational definitions are given in Scheerens et al. (2007).
4. Argyris and Schön (1978) apply these strategies to a situation where an organization is confronted by a significant developmental or strategic problem.

References

Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective*. Reading, MA: Addison-Wesley.
Bellamy, G. T. (2011). High reliability & leadership for educational change. In McREL, *Noteworthy perspectives: High reliability organizations in education* (pp. 24–35). Denver, CO: McREL. Retrieved from http://www.mcrel.org/~/media/Files/McREL/Homepage/Products/01_99/prod87_NW_HRO.asha
Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, *73*, 125–230.
Borman, K. M., Carter, K., Aladjem, D. K., & LeFloch, K. C. (2004). Challenges for the future of comprehensive school reform. In C. T. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 109–150). Washington, DC: National Clearinghouse for Comprehensive School Reform.
Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, *24*, 409–429.
Coates, D. (2003). Education production functions using instructional time as an input. *Education Economics*, *11*, 273–292.
Cobern, W., Schuster, D., & Adams, B. (2010, March). *Experimental comparison of inquiry and direct instruction in science*. Paper presented at the Annual Conference of the Society for Research on Educational Effectiveness, Washington, DC.
Cohen, J. (1988). *Statistical power analyses for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of education opportunity*. Washington, DC: Government Printing Office.
Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. New York, NY: Routledge.
Creemers, B. P. (1994). *The effective classroom*. London: Cassell.
Daly, A. J., Moolenaar, N. M., & Carrier, N. (2011). *Reform at the edge of chaos: Connecting complexity, social networks, and policy implementation*. Manuscript submitted for publication.
Ehren, M. C. M. (2007). Het spel der neveneffecten: Strategisch gedrag van toezichthouders en scholen [The game of side-effects: Strategic behavior of supervisors and schools]. In F. L. Leeuw, J. S. Kerseboom, & R. Elte (Eds.), *Turven, tellen, toetsen; over toezicht, inspectie, handhaving en evaluatie en hun maatschappelijke betekenis in Nederland* (pp. 195–203). Den Haag: Boom Juridische Uitgevers.
Goldhaber, D. (2000). School choice: Do we know enough? *Educational Researcher*, *29*(8), 21–22.
Goodwin, B. (1994). *How the leopard changed its spots*. London: Weidenfeld & Nicolson.

Gorard, S., Fitz, J., & Taylor, C. (2001). School choice impacts: What do we know? *Educational Researcher*, *30*(7), 18–23.

Griffith, J. (2003). Schools as organizational models: Implications for examining school effectiveness. *The Elementary School Journal*, *104*, 29–47.

Hallinger, P. (1984). *Principal instructional management rating scale*. New York, NY: Leading Development Associates.

Harris, A., Chapman, C., Muijs, D., Reynolds, D., Campbell, C., Creemers, B., … Weinstein, J. (2013). Getting lost in translation? An analysis of the international engagement of practitioners and policy-makers with the educational effectiveness research base. *School Leadership and Management*, *33*, 3–19.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Hofman, R. H., Hofman, W. H. A., Guldemond, H., & Dijkstra, A. B. (1996). Variation in effectiveness between private and public schools: The impact of school and family networks. *Educational Research and Evaluation*, *2*, 366–394.

Hopkins, D., Stringfield, S., Harris, A., Stoll, L., & Mackay, T. (2011). *School and system improvement: A narrative state-of-the-art review*. Manuscript submitted for publication.

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

Kane, T. J. (2004). *The impact of after-school programs: Interpreting the results of four recent evaluations* (Working Paper). New York, NY: WT Grant Foundation.

Kauffman, S. A. (1993). *Origins of order: Self-organization and selection in evolution*. Oxford: Oxford University Press.

Kickert, W. J. M. (1979). *Organization of decision-making: A systems-theoretical approach*. Amsterdam: North-Holland Publishing Company.

Kieser, A., & Kubicek, H. (1977). *Organisation*. Berlin: Walter De Gruyter.

Kyriakides, L., Campbell, R. J., & Gagatsis, A. (2000). The significance of the classroom effect in primary schools: An application of Creemers' comprehensive model of educational effectiveness. *School Effectiveness and School Improvement*, *11*, 501–529.

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

Kyriakides, L., & Tsangaridou, N. (2008). Towards the development of generic and differentiated models of educational effectiveness: A study on school and teacher effectiveness in physical education. *British Educational Research Journal*, *34*, 807–838.

Lawrence, P. R., & Lorsch, J. W. (1967). *Organization and environment*. Cambridge, MA: Harvard University Press.

Lindblom, C. E. (1959). The science of muddling through. *Public Administration Review*, *19*, 79–88.

Louis, K. S., Dretzke, B., & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, *21*, 315–336.

Luhmann, N. (1995). *Social systems*. Stanford, CA: Stanford University Press.

Luyten, J. W., Scheerens, J., Visscher, A. J., Maslowski, R., Witziers, B., & Steen, R. (2005). *School factors related to quality and equity: Results from PISA 2000*. Paris: OECD.

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht: D. Reidel Publishing Company.

Mintzberg, H. (1979). *The structuring of organizations*. Englewood Cliffs, NJ: Prentice Hall.

Morgan, G. (1986). *Images of organizations*. Beverly Hills, CA: Sage.

Muijs, D., Creemers, B., Kyriakides, L., Van der Werf, G., Timperley, H., & Earl, L. (2011). *State of the art – teacher effectiveness and professional learning*. Manuscript submitted for publication.

Niskanen, W. A. (1971). *Bureaucracy and representative government*. Chicago, IL: Aldine-Atherton.

Nordenbo, S. E., Elstad, E., Holm, A., Scheerens, J., Soegaard Larsen, M., Uljens, M., … Hauge, T. E. (2009). *Research mapping of input, process and learning in primary and lower secondary schools: Technical report*. Copenhagen: DPU, Aarhus University.

Organisation for Economic Co-operation and Development. (2005). *Education at a glance: OECD indicators 2005*. Paris: Author.

Pfeffer, J. (1978). *Organizational design*. Arlington Heights, IL: AHM Publishing Company.

Prirogine, I. (1997). *The end of certainty: Time, chaos and the new laws of nature*. New York, NY: The Free Press.

Reezigt, G. J., Guldemond, H., & Creemers, B. P. M. (1999). Empirical validity for a comprehensive model on educational effectiveness. *School Effectiveness and School Improvement*, *10*, 193–216.

Reynolds, D., Sammons, P., De Fraine, B., Townsend, T., Van Damme, J., Teddlie, C., & Stringfield, S. (2011). *Educational effectiveness research (EER): A state-of-the-art review*. Manuscript submitted for publication.

Reynolds, D., Stringfield, S., & Schaffer, G. (2006). The High Reliability Schools Project: Some preliminary results and analyses. In A. Harris & J. H. Chrispeels (Eds.), *Improving schools and educational systems* (pp. 56–76). Abingdon: Routledge.

Sahlberg, P. (2011). *Finnish lessons*. New York, NY: Teachers College Press.

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

Scheerens, J. (1997). Conceptual models and theory-embedded principles on effective schooling. *School Effectiveness and School Improvement*, *8*, 269–310.

Scheerens, J. (2004). The evaluation culture. *Studies in Educational Evaluation*, *30*, 105–124.

Scheerens, J. (2007, March). *Conceptual framework for the PISA 2009 context questionnaires and thematic reports*. Paper presented to the PISA Governing Board, Oslo, Norway.

Scheerens, J. (2008). L'istruzione intesa come sistema adattativo complesso: Implicazioni per gli studi sull'efficacia educativa [Education as a complex adaptive system: Implications for the study of educational effectiveness]. In F. Abbona, G. Del Re, & G. Monaco (Eds.), *Complessità dinamica dei processi educativi* (pp. 136–171). Milano: FrancoAngeli.

Scheerens, J. (2012). *School leadership effects revisited: Review and meta-analysis of empirical studies*. Dordrecht: Springer.

Scheerens, J. (2013a). *Productive time in education: A review of the effectiveness of teaching time at school, homework and extended time outside school hours*. Manuscript submitted for publication.

Scheerens, J. (2013b). The ripples and waves of educational effectiveness research: Some comments to "Getting lost in translation". *School Leadership & Management*, *33*, 20–25.

Scheerens, J. (2013c). The use of theory in school effectiveness research revisited. *School Effectiveness and School Improvement*, *24*, 1–38.

Scheerens, J., Luyten, H., Steen, R., & Luyten-de Thouars, Y. (2007). *Review and meta-analyses of school and teaching effectiveness*. Enschede: University of Twente, Department of Educational Organisation and Management.

Scheerens, J., Luyten, H., & Van Ravens, J. (Eds.). (2011). *Perspectives on educational quality: Illustrative outcomes on primary and secondary education in the Netherlands* (Springer Briefs in Education). Dordrecht: Springer.

Scheerens, J., & Maslowski, R. (2008). Autonomie des établissements scolaires: Des moyens à la recherche d'un objectif? [School autonomy: Means looking for a goal?]. *Revue francaise de pedagogie*, *164*, 27–36.

Schein, E. H. (1985). *Organizational culture and leadership: A dynamic view*. San Francisco, CA: Jossey-Bass.

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

Simon, H. A. (1964). *Administrative behavior*. New York, NY: Macmillan.

Snow, R. E. (1973). Theory construction for research on teaching. In R. M. W. Travers (Ed.), *Handbook of research on teaching* (pp. 77–112). Chicago, IL: Rand McNally.

Stacey, R. D., Griffin, D., & Shaw, P. (2000). *Complexity and management: Fad or radical challenge to systems thinking?* New York, NY: Routledge.

Stringfield, S. (1998). An anatomy of ineffectiveness. In L. Stoll & K. Myers (Eds.), *No quick fixes: Perspectives on schools in difficulties* (pp. 209–221). London: Falmer.

Stringfield, S., Reynolds, D., & Schaffer, E. C. (2008). Improving secondary students' academic achievement through a focus on reform reliability: 4- and 9-year findings from the High Reliability Schools project. *School Effectiveness and School Improvement*, *19*, 409–428.

Stringfield, S. C., Reynolds, D., & Schaffer, E. C. (2011). Toward highly reliable, high-quality public schooling. In McREL, *Noteworthy perspectives: High reliability organizations in*

education (pp. 6–23). Denver, CO: McREL. Retrieved from http://www.mcrel.org/~/media/Files/McREL/Homepage/Products/01_99/prod87_NW_HRO.asha

Stringfield, S. C., & Slavin, R. E. (1992). A hierarchical longitudinal model for elementary school effects. In B. P. M. Creemers & G. J. Reezigt (Eds.), *Evaluation of educational effectiveness* (pp. 35–39). Groningen: ICO.

Tarter, C. J., & Hoy, W. K. (2004). A systems approach to quality in elementary schools: A theoretical and empirical analysis. *Journal of Educational Administration, 42*, 539–554.

Valentine, J. C., Cooper, H., Patall, E. A., Tyson, D., & Robinson, J. C. (2010). A method for evaluating research syntheses: The quality, conclusions, and consensus of 12 syntheses of the effects of after-school programs. *Research Synthesis Methods, 1*, 20–38.

Van Damme, J., De Fraine B., Opdenakker, M.-C., Van Landeghem, G., & Onghena, P. (2000, September). *Effects of schools and classes upon mathematics and language achievement: The importance of group composition.* Paper presented at the European Conference on Educational Research, Edinburgh.

Van der Werf, G. (1997). Differences in school and instruction characteristics between high-, average-, and low-effective schools. *School Effectiveness and School Improvement, 8*, 430–448.

Von Bertalanffy, L. (1968). *General systems theory.* New York, NY: George Braziller.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*, 1–19.