Critical review

School effectiveness research: a review of criticisms and some proposals to address them.

by **Andres Sandoval-Hernandez** (A.Sandoval@bath.ac.uk)
University of Bath, Department of Education

Contextualization

This literature review is part of research concerned with identifying, theoretically justifying and testing through multi-level (ML) models, the variables and inequity patterns of lower-secondary education in Mexico. The approach is one which sees developments in school effectiveness as having reached a new stage in which the earlier common sense identification of variables used in statistical modelling has now given way to a more theoretically informed approach to the variables used in school effectiveness models. Bearing this in mind, the key research questions will include the following: which other variables that have been consistently omitted by SER are likely to add to the explanatory power of the ML Models? What theoretical framework can be used to develop and test models including those variables? What proportion of the variation in the school outcomes depends on students', schools' and state's characteristics? Does the effect of these models on students' outcomes vary for different schools' context characteristics? (School composition: family economic, social and cultural capital, and school modality).

Abstract*: Based on Lakato's notion of research programmes, the paper analyses the structure of the School Effectiveness Research (SER) programme and reviews the main criticisms that have arisen, stressing those regarding its objectivity and theoretical limitations. Then, some proposals are made to address these criticisms, namely: to adopt a critical realist approach to the study of SE and an Abductive Theory of Scientific Method that lead to the development of sound theory in the field. Based on this analysis the paper concludes that, in terms of Lakatos, a movement towards a new research programme is needed in order to ensure that the main objectives originally set for SER can be eventually reached.*

Introduction

School effectiveness research (SER) has reached a point where it is necessary to review and evaluate the critiques it has received in the past in order to guarantee the accomplishment of the goals originally set for it. For this paper I am going to adopt the concept of SER given by Luyten *et al.*: "…a line [programme] of research that investigates performance differences between and within schools, as well as the malleable factors that enhance school performance, usually using student achievement scores to measure the latter" (Luyten, Visscher and Witziers, 2005, p 249). So, if progress is going to be made in the understanding of these differences and in the explanation of the relations between these factors and the schools' effectiveness, a new realism as to schools limits and capabilities is needed as well as a re-examination of how research can inform these explanations (Lauder, Jamieson and Wikeley, 1998).

The objective of this paper is to review the main criticisms posed against SER and to make some proposals to address them. Criticisms made of SER can be classified into objectivity issues, theory limitations and methodological limitations (Teddlie and Reynolds, 2001). Only the first two are discussed in this paper.

This work is structured following Lakatos' (1970) notion of research programmes. Thus a review of its main concepts is presented first, followed by an analysis of the structure of the SER programme proposed by Lauder and colleagues (1998), then the main criticisms to the SER's objectivity and its theory limitations are reviewed and, based on the analysis of the previous points, an evaluation of the mainstream tradition of SER is made. Finally some recommendations to address the critiques and their consequences are presented.

Lakatos' notion of research programmes

According to Lakatos (1970), a research programme consists of three main parts: a hard core, a protective belt and positive and negative heuristics.

The hard core consists of very general hypotheses that give the research program its essential characteristics and bedrock assumptions.

The protective belt is a set of auxiliary theories underlying and protecting the hard core from falsification, although part of the research program can be changed or adjusted as a result of anomalies or empirical challenges without abandoning the program itself.

The heuristics (positive and negative) are a set of rules or hints about how to treat the hard core and the protective belt in order to aid discovery or invention. The negative heuristics state what the scientists are advised not to do, for example that the hard core of the program should not be abandoned or modified. On the other hand, the positive heuristics specify what scientists should do within the programme – what issues they should address, in what order – for changing or adjusting the protective belt.

Finally, according to Lakatos, the major indication of merit of a research programme is to what extent it leads to novel predictions or explanations that are confirmed. In this way, a progressive research programme is one which remains coherent between its hard core and protective belt, and eventually leads to novel predictions or explanations (theoretical progressiveness) that are subsequently corroborated (empirical progressiveness); whereas a degenerating research programme is one that loses its coherence and/or is not capable of leading to novel predictions or explanations and/or is not capable to corroborate them.

The hard core of SER

Lakatos (1970) coined the term 'rational reconstruction' which refers to what the researchers have logically and historically been committed to, independent of their own thinking, that is to say that it is not necessarily related to what a researcher believes or their personal commitments. Drawing on this notion Lauder et al.(op cit, pp 52-53) suggest that a rational reconstruction of the mainstream tradition of SER would be as follows:

1. Schools as organizations do have an effect on student outcomes in terms of exam success, in contrast say to the view that it is teachers rather than schools who have effects on student outcomes, or to the view that when contextual factors are fully taken into consideration there is no significant difference between schools.

2. These school effects are not caused by chance, and effects that improve school performance, in relation to exam success, can be engineered on the basis of SER.

3. Schools are like any other organization where staff and children will respond to the systems, sanctions and rewards of a school so that successful re-engineering is possible.

4. Schools as organizations are structured as nested organizations. For example, with central / local government as the broadest organizational structure and within that, the school, departments, classrooms, etc. It is further assumed that effectiveness can be identified by analysing the performance of each level and by inference the effectiveness of the relationship between each level.

5. There is a degree of relative autonomy between levels and between the education system and society. Therefore schools can generate effects independently of many of the factors external to the school which may impinge on exam outcomes.

Some critiques of SER

There has been a long debate within SER and a vast debate among critics and defenders of this programme. However, there are some outstanding compilations of criticisms and counter-criticisms. Among the criticisms, the books of Slee *et al.* (1998) and Thrupp (1999) stand out. As for the counter-criticisms, Teddlie and Reynolds (2001) wrote an exact standoff in defence of SER. Luyten *et al.* (2005) is another document mentioned throughout this paper as it is a recent work that confronts critiques and recommendations related to this issue. This paper, as has been previously stated, will be focused exclusively on those aspects of the debate related to the objectivity of SER and to its alleged lack of theory.

The objectivity of the SER

In this section an overview of some of the main critiques raised about the objectivity of SER is provided. These kinds of critiques have been mainly made on two different grounds: methodological objectivity and ideological objectivity.

i) Methodological objectivity
This point refers to the SER assumption that applying rigorous quantitative methods will necessarily generate objective knowledge (Luyten *et al*., 2005, p 251). It certainly can be said that the use of a rigorous quantitative methodology can be one of several requisites to obtain more or less reliable results, but it is only one of them. In this sense, critics of SER point out that objectivity must be sought before that, when deciding what kind of information is going to be collected, what kind of methodology is going to be used to do it and how it will be done; and after, once the results have been obtained, when interpreting them (Grace, 1998; Wrigley, 2004). Here is where ideological objectivity comes into play.

ii) Ideological objectivity
The critiques in this sense are related to the close ties between researchers and policy makers which might give the perception that the SER agenda reflects governmental concerns more than scientific considerations; or that SER is not as much a scientific endeavour as it is an ideological force (Luyten *et al*., 2005, p 251). According to some critics the processes of formulating research questions, collecting information, and interpreting findings unavoidably involve ideological and political choices, which may serve particular interests and are more evident when the research is dominated and funded by government or governmental related agencies (Luyten *et al*., 2005). Other authors consider that it is not that SE researchers readily align themselves to certain social philosophy, but that the biases in SER have to do with its epistemological commitments. Willmott, for example, considers that the underlying thread of the critique of SER's objectivity is its commitment to positivist epistemology (and its concomitant ontology) that necessarily has congruity with neo-liberalism, since in positivism 'society' is held to be constituted solely by constantly conjoined events, somehow brought about by externally related individuals and their doings. As a result, non egalitarian social structures are reduced to individuals and it is this fallacy of

composition that provide the philosophical backdrop to the government's 'marketisation' of the education system (Willmott, 1999a, p 255).

Moreover, some critiques point out that the tendency of SER to focus on the relationship between student performance and school factors, while attributing the popularity of schools to their performance rather than their intakes and other context variables, leads to a culture of blame (Hargreaves, 1994; Thrupp, 1998, 1999), that is to say that governments have often used SER arguments to show school failure to be the responsibility of schools alone without any reference to the broader socio-political context, such as the impact of poverty. According to this, the claim that student achievement is significantly affected by socio-economic status and other context characteristics is excluded as a plausible excuse for poor school performance (Thrupp, 1999).

Theory limitations

It is in the critiques regarding the theoretical grounding of the SER where more agreement has been reached: both promoters (eg, Teddlie and Reynolds, 2001) and detractors (eg, Slee *et al*., 1998) of SER concur that there is a lack of theory in the programme.

The critiques in this sense have to do with the lack of theoretical basis for the selection, operationalisation and explanation of the relations between variables; and with the lack of theories to support the fundamental assumptions of SER regarding the nature of schools, students and teachers.

The procedures challenged by the first group of critiques provide the basis (eg, variables) to develop the hypotheses challenged by the second group, which represent the protective belt of the SER programme. According to several critics these theory limitations can be said to be related with its positivist affiliation (Thrupp, 1999; Willmott, 1999b; Wrigley, 2004).

i) Selection, operationalisation and explanation of the relationship between variables
Authors like Coe and Fitz-Gibbon (1998) and Lauder *et al.* (1998) claim that normally the inclusion of variables considered for SER are justified more on statistical than theoretical criteria, and that the operationalisation of these variables tends to rely on no more than common sense. An example of this is the technique commonly used by SER to generate concepts like 'school climate' or 'leadership', which consists of operationalising these concepts with a battery of items (selected with no more basis than common sense) which when combined 'add up to 'school climate' or 'leadership'. If they then turn out to be statistically significant they are regarded as important to school effectiveness (Lauder *et al*., 1998, p 54). A further problem of these common sense operationalisations is that they vary considerably across studies.

Furthermore, Coe and Fitz-Gibbon (op cit) accuse SER of 'fishing for correlations' between particular factors associated with school effectiveness and particular characteristics of schools, without specifying why or how it is expected that these particular characteristics be related with school effectiveness. In addition, these authors point out that most of SER rely on a simple linear logic, eg, more of this is associated with more of that. In this fashion, such conclusions are neither able to indicate causal relations between variables nor explain the mechanisms behind these relationships.

Another critique related to the selection of variables considered by SER says that to date, the programme has focused mainly on investigating direct effects of school level variables on school performance, when some of the variation in school performance may be explained by the relationship between school context and processes occurring within schools (Luyten *et al*., 2005). That is to say, the indirect effect of the school context on the school performance through the school processes. There is also a set of variables that have been often omitted

by the SER. Thrupp (2001b) for example, refers to the background characteristics of students, the composition of student population within schools and the curricula used by schools; Lauder *et al.* (1998) mention the school culture or politics and the external policy environment; and Luyten *et al.* (2005) point out the question of which school and classroom structures and strategies are more profitable for different types of student groups.

Regarding the variables used to measure the school outcomes, authors like Knuver and Brandsma agree that SER has traditionally focused on basic cognitive pupil outcomes such as language and mathematics achievement (Knuver and Brandsma, 1993, p 189), without considering a broader set of objectives (eg, the degree to which the student feels at home in the school environment, the academic self-concept, and the social integration in the class) that can be or are imputable to formal education. On this point it is important to say that even when there are other studies that use non-cognitive dependent variables (eg, Cervini, 2003b; Kyriakides, 2005; Van Landeghem *et al*., 2002), they are not common in SER.

ii) Fundamental assumptions about the nature of schools, students and teachers
According to the analysis of Lauder and his colleagues (1998), the SER programme appears to be coherent within its hard core assumptions and between the hard core and its protective belt of hypothesis. For example the hard core assumption that no difference exists between schools and other kinds of organisations can be supported with the protective belt of 'quasi managerialist theories' (Lauder *et al*., 1998), which use the notions of leadership, goal definition, etc. to develop hypotheses about the organizations / schools (Bosker and Visscher, 1999; Cervini, 2003a). Nevertheless fundamental assumptions concerning the nature of schools, teachers and students are still subject to debate.

On this point, the criticisms can be summarized in two questions. Are schools, teachers and especially students, as Bowles and Gintis (1976) considered, essentially passive individuals regarding the structure and demands of the educational system? Are the schools actually like any other organisation or, as considered by Ball (1996), they have particular characteristics which make them substantially different?

According to Ball, "schools are complex, contradictory, sometimes incoherent organisations like many others. They are assembled over time to form a bricolage of memories, commitments, routines, bright ideas and policy effects. They are changes, influenced and interfered with regularly and increasingly. They drift, decay and re-generate". (Ball, 1996, p 1)

The hard core assumptions of SER also include a belief that schools as organisations do have an effect on student outcomes. In other words, once all the contextual variables are fully taken into consideration there is a significant difference between the schools outcomes of different schools. Moreover, that the effects that improve school performance (once identified) can be engineered on the basis of SER to be replicated in non-effective schools.

To protect those statements SER has developed a vast group of hypotheses using the contextual factors that seem to have some influence over school outcomes. Among the most recurrent in the literature are socio-economic status, cultural capital and social capital, students' gender, students' and their parents' expectations regarding the academic future of the former, and so on. Although nowadays there is a consensus that these factors play an important role in determining schools outcomes, the relative importance of each is still a topic of debate within the literature. That is to a great extent because of a problem that was mentioned before: the operationalisation of these variables varies considerably across studies. At the same time little has been said about the mechanisms that these relationships depend upon (Coe and Fitz-Gibbon, 1998). Even though there is acknowledgement of some studies that make use of social theories – like the ones developed by Bourdieu (1977), Bernstein (1971) and Boudon (1974), among others – when justifying the inclusion of context variables in their analysis (eg, Cervini, 2003b; Fernández, 2003a, 2003b; Taylor, Muller, and

Vinjevold, 2003), they are rather scant and do not use these theories to explain in depth the relationships between the independent variables and school outcomes.

Besides, there are severe criticisms to the claim that SER's findings, regarding the characteristics that make a school or classroom effective, can be used as a basis for school improvement interventions; namely, to try to improve low performing schools by encouraging them to adopt the characteristics detected in effective schools. This is because, according to the critics, schools differ so much in relevant aspects, such as the causes underlying their specific performance, capacity for change, contextual characteristics, etc. These differences are stressed when considering the practice of importing school effectiveness models from one country to another. That is to say that one-size-fits-all solutions cannot be used to improve school performance, instead school improvement efforts should carefully consider the 'power of site or place' (McLaughlin, 1998; Miles, 1998).

Finally, there are some voices arguing that SER must lift the level of abstraction from mere empiricism to a more conceptual level (Wyatt, 1996). An example of this is another of the concepts developed by the protective belt of the SER: school climate. This concept plays an important role in the internal structure of the SER, since once a set of factors have been identified as producers of school effectiveness it is necessary to organize them around a concept capable of putting the constituent parts of the recipe together; school climate seems to do the trick (Lauder *et al*., 1998). Even when several studies show that the school climate can vary across schools and can influence the teachers and students performance, there is no theory that explains how it is created (Lauder *et al.,* 1998) or how it interacts with the teachers and students performance. According to the critics, to address these questions it is necessary first, to move the analysis to a conceptual level, and then deal once again with the problem mentioned above; this concept – like many of those developed by SER – has been usually defined by an arbitrarily selected set of items in a questionnaire that varies greatly among different studies (Miller and Fredericks, 1990).

The protective belt of SER

This section of the paper is mainly based on the article written by Teddlie and Reynolds (2001) to address the criticisms of SER contained in two books: Slee and Weiner with Tomlinson (1998) and Thrupp (1999). The format of the paper is in the 'Point: Counterpoint' style and presents a series of fourteen criticisms of SER followed by the counterpoints, which range from simple statements of agreement to vigorous defences. The importance given to the work of Teddlie and Reynolds relates to the title of principal gatekeepers of School Effectiveness conferred to them by authors like Wrigley (2004) due to their key role as editors of The International Handbook of School Effectiveness Research (Teddlie and Reynolds, 2000).

Teddlie and Reynolds (T&R) classify the criticisms against SER in terms of ideological objectivity, theory limitations and methodological issues. These points are addressed below:

i) Ideological objectivity

SER is tacitly presented as politically neutral as it has been criticized politically on contradictory grounds. T&R claim that, paradoxically, besides the critiques of 'progressive' educators (eg, Elliott, 1996; Slee *et al*., 1998; Thrupp, 1999) which link SER to a right-wing social ideology; SER has also received critiques from 'conservative' scholars (eg, Ralph and Fennessey, 1983; Rowan, 1984), labelling SE researchers as liberal reformers more interested in improving the circumstances of poor children than in conducting good science (Teddlie and Reynolds, 2001, p 42). From this, T&R conclude that as long as SE researchers are accused of conducting research that supports both liberal and conservative causes, these criticisms can be accepted as simply a part of the territory in which they work. Furthermore, they claim that SER will always receive a certain amount of political criticism

regardless of the nature of the results and their interpretation, and that SE researchers will be only partially successful in refuting it, since it comes from all directions (idem: 43).

ii)Theory limitations
With respect to the critiques related to the lack of theory, T&R merely state their agreement by saying that these criticisms still have validity. However they point out that researchers and theoreticians are attempting to solve this 'perennial' problem. In fact, there is some evidence of works using different theoretical developments to explain the relationships between some school and context variables and school outcomes. In this sense the works of Fernandez (2003a, 2003b), Cervini (2003b), Murillo (2004b) and Annevelink (2004) can be cited. In the last, the author built her work on a hypothetic model that shows how reduction in class size affects students achievements in three different lines; Fernandez and Cervini attempt to use some of the classic theories of sociology of education to give a frame to their works; and Murillo analyses different theoretical traditions in education (curriculum, organization and behavioural theories) that could provide elements to be validated empirically by SER. As for T&R, they add the work of Scheerens and Bosker (1997) to shed some light on possible theoretical explanations between various school and classroom factors by interrogating existing studies with the predictions of 'meta' or 'grand' theories, such as contingency theory or public choice theory (Teddlie and Reynolds, 2001, p 44).

iii) Methodological issues
T and R present the methodological criticisms as internal debates of the field and it is said that many of them have been muted by recent methodological developments. I agree with this statement and some examples are presented in the next section of this paper.

Finally, regarding the alleged positivism of SER, even though T&R admit that some SE researchers work under the (post) positivist paradigm, they disassociate themselves from this tradition by claiming they have assumed a pragmatic position, and explicitly state their research has the following orientations:

- Inclusion of both quantitative and qualitative methods for gathering and analysing data;
- Use of both deductive and inductive logic depending on the phase of the research project;
- Use of both objective and subjective epistemological approaches depending on the data being worked with;
- An axiology in which values play a large role in interpreting results;
- An ontology that accepts external reality, yet denies that truth can be determined once and for all.

Yet, as Wrigley (2004, p 283) points out, here pragmatism appears to refer simply to a greater methodological variety rather than to a epistemological affiliation. What is more, a rough analysis of these points shows that none of them conflict with a positivist posture.

The SER programme: Progressive or degenerating?

The main criteria proposed by Lakatos (1970) for evaluating research programmes is their capability to lead to novel explanations through theoretical and methodological innovation. Based on that, research programmes can be classified as progressive or degenerating.

There are clear examples suggesting that SER can be considered as a progressive research programme, at least on methodological grounds. One of the most remarkable is the use of Multilevel Modelling (MLM) analysis techniques. MLM gains its direction from the hard core assumption of 'nestedness', which has made it possible to expand the range of both 'within'

and 'between' school analysis in terms of the treatment that schools give to specific groups of students (Lauder *et al*., 1998). It has also allowed decomposition of the schools outcomes variance between the different levels considered in the analysis, giving more precision to the results and making it possible to deal with the technical problems that arise when working with data of multilevel nature.

The development of so-called 'value added' studies is another example. These studies have made possible a fairer comparison between schools since the progress of all pupils counts in evaluating school performance instead of raw test examinations results, thus schools are held accountable only for those things they can influence and not for all the existing differences between their intakes (cf. Sammons, 2001).

As mentioned before, the evidence of theoretical developments, even though they can be considered as incipient, is another example that can give credit to considering the SER programme as a progressive one.

Now, even though the theoretical and methodological developments are doubtlessly plausible and may support the claim that SER is a progressive programme, there are still criticisms and questions about its fundamental assumptions with no answers. So, what steps have to be taken to address them? Next, this paper tries to answer this question in terms of the positive heuristics of the research programme.

The heuristics of SER: Towards a new research programme

Although the importance of the issues concerning the objectivity of SER is undeniable, they cannot be addressed as part of the heuristics since they do not interfere in a direct way with the core assumptions (ie, hardcore) of the programme. Nevertheless, I wish to present my position on this debate. Even though many of the quantitative methods used in SER involve nothing more than using computer algorithms to process data and that data processing itself can be considered as an objective procedure, the decisions on what and how information in collected necessarily imply personal choices and the output of the analyses always requires some interpretation. Thus, SE researchers should provide enough information about data sets, analysis techniques and outcomes for the readers to be able to detect possible biases. Research results derived from public datasets that are open for secondary analysis also help in this sense (Luyten *et al.,* 2005). Moreover, given the intimate relationship between SER and its governmental sponsors, it is not acceptable that SE researchers wash their hands of any responsibility for the application of their research (Teddlie and Reynolds, 2001); in turn, they need to recognise the existing dilemmas and speak out against the misuse of their research findings (Wrigley, 2004).

Acknowledging the importance of theory for the development of SER, the positive heuristics of the programme would point to the development of a scientific methodology that supplies tools for the construction of explanatory theories. This suggests a quite different methodology to the positivism presupposed by SER. The detection of robust empirical regularities–phenomena, in Haig's terms–would be the starting point (Haig, 2005a), then researchers would have to move to the construction of possible [theory driven] explanations, which can subsequently be tested out (Bhaskar, 1975) with a return to the formulation of the problem, this time in more precise terms and with better understanding of the reality (Haig, 2005a).

The critical realist paradigm provides the elements to frame such a methodological account, as its ontology assumes many of the key theories that might be used in this context to refer to the dynamics of power within school and between school and society. But power is hard to observe, except in the most obvious cases (Lukes, 1974) thereby suggesting that critical realism which admits unobservable entities into its ontology and epistemology provides a starting point. However given that theories may be competing in seeking explanations for

observed phenomena, the knowledge gained from such a methodology will be from the best theories available at a given time.

Here, the Abductive Theory of Scientific Method (ATOM) developed by Haig (2005a) is especially relevant, since it shares the ontological and epistemological commitments of the critical realist approach. According to its author, ATOM is meant to provide a guide-line to systematically describe how the existence of phenomena can be established and then how theories can be constructed to explain those facts. ATOM proceeds in a facts-before-theory sequence, claiming that the search for understanding empirical phenomena is what gives explanatory construction its point. In this sense, empirical phenomena exists to be explained rather than to be used as the objects of prediction in theory testing (Haig, 2005a).

In this context, qualitative methods may clearly have a place. For example, narrative explanations (Nash, 2002) can help to identify the possible connections between school effectiveness and particular school or context and in determining whether the causal mechanisms revealed by such a method are merely accidental or more widespread in explaining elements of school effectiveness.

The following points explain these ideas in more detail, and in doing so, some elements are provided to better understand the research questions set for the work of which this literature review forms part.

1. As it was said, the first stage would be the detection of phenomena. A wide range of methodological strategies can be used in order to establish the reliability of data for detecting phenomena. In SER context, an example of data could be a set of findings regarding the statistical correlation between student's achievement in public primary education in Mexico City, and variables like family income and if the students' family owns certain household goods. From several studies carried out in different contexts (in different levels of education, in different cities, in private education, using quantitative and qualitative analysis techniques, using information from different academic tests, etc. and showing a constant pattern (eg, a positive correlation between economic variables and academic performance), it could be inferred that there is a phenomena existing in the relation between the economic capital of a family and the school attainment of its members.

ATOM proposes a model of data analysis which assists in the detection of phenomena by attending in turn to i) data quality, which can be verified through exploratory data analysis; ii) Identification of patterns, that refers to the effect of school and context related variables on student achievement and are normally identified by statistical techniques such as MLM; and iii) Generalization of patterns, that is, verifying the stability of these emergent data patterns through the use of confirmatory data analysis procedures. Coincident research results derived from public datasets that are open for secondary analysis can be a way to confirm the existence of patterns. When different researchers, using the same data sets but different samples, analysis methods and approaches, get to identify the same factors as enhancers of school performance it can be said that those patterns are confirmed.

2. Once phenomena have been detected, the next stage would be the construction of explanatory theories. Theory construction refers to the generation of explanations of the detected phenomena by abductively inferring the existence of underlying causal mechanisms. Statistical techniques such as exploratory factor analysis (Haig, 2005b) and multilevel structural equation models (Goldstein and McDonald, 1988) are especially relevant here as they are designed to facilitate the postulation of latent variables that are thought to underlie patterns of correlations in new domains of manifest variables.

An alternative approach could be the use of qualitative research to develop theoretical notions of school processes and to test them through quantitative large-scale research

(Luyten *et al*., 2005; Scheerens and Bosker, 1997; Thrupp, 2001a). The opposite direction can also be followed, in other words, quantitative methods can be used as a diagnostic tool to identify relationships between variables (data patterns), and then develop the theoretical notions based on qualitative research.

Additionally, the generation of theories to explain phenomena can be supported by making more use of theoretical developments from other disciplines (Luyten *et al*., 2005; Thrupp, 2001b). In this sense, it is necessary to keep on evaluating and analysing the links that have been tried so far and, at the same time, look for other theories that could contribute with elements to frame and support SER. Without trying to be exhaustive, the next table shows some links between groups of variables and theoretical lines that can be further developed.

Table 1. Some theoretical developments from other disciplines that can be used in SER

Level of the variables	Groups of variables	Theoretical developments
States / Geographical areas	Level of economic development	Heyneman y Loxley (1976, 1982, 1983)
School	Socio-economic context	organisational theories[1]: Environment approach (DiMaggio and Powell, 1983)
	Socio-cultural context	
	School management	
	Pedagogic practice	organisational theories[1]: Structure approach (Buckley, 1967)
	School climate	
	Socio-economic-and cultural indicators of teachers and head-teachers	organisational theories[1]: Organisational link approach (Seidl, 2003)
Student	Economic capital	Reproduction theory (Bourdieu, 1983; Bourdieu and Passeron, 1977)
	Cultural capital	
	Social capital	Bourdieu (1977, 1983), Coleman (1988) and Putnam (1995) social capital theories
	Academic expectations	Boudon (1974) and Goldthorpe (1996) rational choice theory
	Family economic structure	Bourdieu (1977, 1983), Coleman (1988), Putnam (1995), Boudon (1974) and Goldthorpe (1996) theories
	Non-educational activities	

In this fashion, an example of the links between these theories and a realist account of SER can be given by acknowledging that some explanations about school processes and outcomes are based on unobservable mechanisms. This can be understood in two senses: effects in school may be observed but their causes may lie outside schools (eg, cultural / social capital); and contradictions between schools' policy/management and the broader

educational policy produce tensions (unobservable mechanisms) that have an actual impact on the school outcomes. In the first case what is unobservable are elements in a web of relationships, which in principle may be observed but for practical reasons cannot; while in the latter case the tensions are in principle unobservable but may nevertheless have real effects.

Summarizing, the realist account of SER pushes for the consolidation of a theoretical framework in the study of school effectiveness. This framework must be based on theoretically driven disciplinary explanations for school processes and outcomes that lead to the development of theory in the field. The approach should move from a positivist paradigm to a critical realist one in an ontological and epistemological sense, which means that the knowledge gained from it will be from the best theories available at a given time; and in the epistemological sense refers to the idea that theories can postulate unobservable mechanisms that explain observable phenomena (Lauder and Brown, 2007). The theoretical developments then, can be done by adding up different parts of different theories (special attention should be paid to their ontological and epistemological compatibility and to the processes or mechanisms used to evaluate and judge between competing theories) and by developing SER's own theoretical notions of the school processes. The premise is that only through explanations furnished by the best theories available will we be able to understand how schools might improve in any given context (Lauder and Brown, 2007).

Finally, even when the methodological progressiveness of the SER programme could be claimed on the basis of its methodological developments, it is clear that the proposals made to address the criticisms analysed in this paper suggest that the advice given by the negative heuristics should not be followed (ie, to not change or modify the assumptions contained in the hard core). In this sense the SER programme cannot be considered as a progressive one. Therefore, a movement towards a new research programme is needed in order to assure that the main objectives originally set for SER can be eventually reached.

References

Annevelink, E. (2004) *Class size. Linking teaching and learning. Enschede*. The Netherlands: Print Partners Ipskamp.

Ball, S. (1996) *Good school / Bad school.* Lancaster: BERA.

Bernstein, B. (1971) *Class, Codes and Control* (Vol. 1). London: Routledge and Kegan Paul.

Bhaskar, R. (1975) *A realist theory of science*. Leeds: Leeds Books.

Bosker, R. and Visscher, A. (1999) Linking school management theory to school effectiveness research, in A. Visscher (Ed.), *Managing schools towards high performance* (pp 291-322). Lisse, The Netherlands: Swets and Zeitlinger.

Boudon, R. (1974) *Education, opportunity and social inequality.* New York: John Willey.

Bourdieu, P. (1977) *Outline of a Theory of Practice.* Cambridge: Cambridge University Press.

Bourdieu, P (1983) Forms of capital, in J. C. Richards (ed.) *Handbook of Theory and Research for the Sociology of Education*. New York: Greenwood Press.

Bourdieu, P. and Passeron, J.-C. (1977) *Reproduction in education, society and culture.* Trans. 2nd edn. R. Nice. London: SAGE.

Bowles, S. and Gintis, H. (1976) *Schooling in Capitalist America.* London: Routledge.

Buckley, W. (1967) *Sociology and modern systems theory.* Englewood Cliffs, NJ: Prentice Hall.

Cervini, R. (2003a) *Factores asociados al aprendizaje del lenguaje y las matemáticas en 13 Estados de México (3° y 4° en educación básica)*. México: Instituto Nacional para la Evaluación de la Educación.

Cervini, R. (2003b) Diferencias de resultados cognitivos y no-cognitivos entre estudiantes de escuelas públicas y privadas en la educación secundaria de Argentina: Un análisis multinivel. *Education Policy Analysis Archives,* 11(6).

Coe, R. and Fitz-Gibbon, C. T. (1998) School Effectiveness Research: Criticisms and Recommendations. *Oxford Review of Education*, 24, 4, 421-438.

Coleman, J. S. (1988) Social capital in the creation of human capital. *The American Journal of Sociology*, 94 (Suplement), S95-S120.

DiMaggio, P J. and Powell, W. W. (1983) The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review,* 48, 2, 147-160.

Elliott, J. (1996) School Effectiveness Research and its Critics: alternative visions of schooling. *Cambridge Journal of Education,* 26, 2, 199-224.

Fernández, T. (2003a) *Tres estudios sobre determinantes sociales del rendimiento escolar.* México: Instituto Nacional para la evaluación de la Educación.

Fernández, T. (2003b) *Contextualización sociocultural de las escuelas de la muestra de Estándares Nacionales (1998-2002).* México: Instituto Nacional para la Evaluación de la Educación.

Goldstein, H. and McDonald, R. P (1988) A general model for the analysis of multilevel data. *Psychometrika,* 53, 455-467.

Goldthorpe, J. H. (1996) The quantitative analysis of large-scale data-sets and rational action theory: for a socilogical alliance. *European Sociological Review,* 12, 109-126.

Grace, G. (1998) Realizing the mission: Catholic approaches to School Effectiveness, in R. Slee, G. Weiner and S. Tomlinson (eds) *School effectiveness for whom? Challenges to the school effectivenes and school improvement movements* (pp 117-127). London: Falmer Press.

Haig B. D. (2005a) An Abductive Theory of Scientific Method. *Psychological Methods,* 10, 4, 371-388.

Haig B. D. (2005b). Explanatory Factor Analysis, Theory Generation, and Scientific Method. *Multivariate Behavorial Research,* 40, 3, 303-329.

Hargreaves, D. (1994) *Changing teachers, changing times.* London: Cassell.

Heyneman, S. P. and Loxley, W. A. (1976) Influences on Academic Achievement: A Comparison of Results from Uganda and More Industrialized Societies. *Sociology of Education,* 49, 3, 200-211.

Knuver, A. W. M. and Brandsma, H. P. (1993) Cognitive and Affective Outcomes in School Effectiveness Research. *School Effectiveness and School Improvement,* 4, 3, 189-204.

Kyriakides, L. (2005) Extending the Comprehensive Model of Educational Effectiveness by an Empirical Investigation. S*chool Effectiveness and School Improvement*, 16, 2, 103-152.

Lakatos, I. (1970) Falsification and the methodology of scientific research programmes, in I. Lakatos and A. Musgrave (eds) *Criticism and the Growth of Knowledge* (pp 91-195). Cambridge: Cambridge University Press.

Lauder, H. and Brown, C. (2007) *An appraisal competing research programmes on school effectivenes and school improvement.* Bath: University of Bath.

Lauder, H., Jamieson, I. M. and Wikeley, F. (1998) Models of Effective Schools: Limits and Capabilities, in R. Slee, G. Weiner and S. Tomlinson (eds) *School Effectiveness for Whom? Challenges to the School Effectiveness and School Improvement Movements.* London: Falmer Press.

Lukes, S. (1974) *Power: A Radical View.* London: Macmillan.

Luyten, H., Visscher, A. and Witziers, B. (2005) School Effectiveness Research: From a review of the criticism to recommendations for further development. *School Effectiveness and School Improvement,* 16, 3, 249-279.

McLaughlin, M. W. (1998) Listening and learning from the field: Tales of policy implementation and situated practice, in A. Hargreaves, A. Liberman, M. Fullan and D. Hopkins (eds) I*nternational handbook of educational change* (Vol. 5, pp 70-84). Dordrescht / Boston / London: Kluwer Academic Publishers.

Miles, M. B. (1998) Finding keys to school change: a 40-year odyssey, in A. Hargreaves, A. Liberman, M. Fullan and D. Hopkins (eds) *International handbook of school change* (Vol. 5, pp 37-39). Dordrescht / Boston / London: Kluwer Academic Publishers.

Miller, S. and Fredericks, J. (1990) The false ontology of school climate effects. *Educational Theory,* 40, 3, 333-342.

Murillo, F. J. (2004b) Un marco comprensivo de Mejora de la Eficacia Escolar. Revista *Mexicana de Investigación Educativa,* 9, 21, 319-359.

Nash, R. (2002) Numbers and Narratives: further reflections in the sociology of education. *British Journal of Sociology of Education,* 23, 3, 397-412.

Putnam, R. (1995) Bowling alone: America's declining social capital. *Journal of Democracy,* 6, 1, 65-78.

Ralph, J. H. and Fennessey, J. (1983) Science or Reform: Some Questions about the Effective Schools Model. *Phi Delta Kappan,* 64, 10, 689-694.

Reynolds, D. and Teddlie, C. (2001) Reflections on the Critics, and Beyond Them. *School Effectiveness and School Improvement,* 12, 99-114.

Rowan, B. (1984) Shamanistic rituals in effective schools. *Issues in Education*, 2, 76-87.

Sammons, P (2001, 2-4 April) *Fairer comparisons in schools. The role of school effectiveness research in promoting improvement*. Paper presented at the The meaning of Quality in Education Conference, Karlstad.

Scheerens, J. and Bosker, R. (1997) *The Foundations of Educational Effectiveness*. Oxford: Pergamon.

Seidl, D. (2003) Organizational identity in Luhmann's theory of social systems, in T. Hernes and T. Bakken (eds) *Autopoietic Organization Theory. Drawing on Niklas Luhmann's Social Systems Perspective* (pp 123-150). Copenhagen: Copenhagen Business School Press.

Slee, R., Weiner, G. and Tomlinson, S. (1998) *School Effectiveness for whom?Challenges to the school effectiveness and school improvement movements*. London: Falmer Press.

Taylor, N., Muller, J. and Vinjevold, P (2003) *Getting schools working: Research and systemic school reform in South Africa.* Cape Town: Pearson Education.

Teddlie, C. and Reynolds, D. (2000) *The International Handbook of School Effectiveness Research.* London and New York: Falmer Press.

Teddlie, C. and Reynolds, D. (2001) Countering the Critics: Responses to Recent Criticisms of School Effectiveness Research. *School Effectiveness and School Improvement,* 12, 1, 41-82.

Thrupp, M. (1998) Exploring the Politics of Blame: School Inspection and Its Contestation in New Zealand and England. *Comparative Education,* 34, 2, 195-209.

Thrupp, M. (1999) *Schools Making a Difference: Let's Be Realistic!: School Mix, School Effectiveness, and the Social Limits of Reform*. London: Open University Press.

Thrupp, M. (2001a) Recent School Effectiveness Counter-critiques: problems and possibilities. *British Educational Research Journal,* 27, 4, 443-457.

Thrupp, M. (2001b) Sociological and political concerns about school effectiveness research: Time for a new research agenda. *School Effectiveness and School Improvement,* 12, 7-40.

Van Landeghem, G., Van Damme, J., Opdenakker, M. C., De Frairie, D. F. and Onghena, P. (2002) The Effect of Schools and Classes on Noncognitive Outcomes. *School Effectiveness and School Improvement,* 13, 429-451.

Willmott, R. (1999a) School Effectiveness Research: an ideological commitment? *Journal of Filosophy of Education,* 33, 2, 253-268.

Willmott, R. (1999b) Structure, Agency and School Effectiveness: researching a 'failing' school. *Educational Studies,* 25, 1, 5-18.

Wrigley, T. (2004) 'School effectiveness': the problem of reductionism. *British Educational Research Journal,* 30, 2, 227-244.

Wyatt, T. (1996) School effectiveness research: Dead end, damp squib or smouldering fuse?. *Issues in Educational Research,* 6, 1, 79-112.