School effectiveness research and Educational Policy

by

Harvey Goldstein[1] and Goeffrey Woodhouse

Institute of Education

Introduction

The field of activity known as 'school effectiveness (SE)' is little more than 20 years old, although some would trace its origins back to the influential work of Coleman (1966) and Jencks et al. (1972). These studies and the influential work of Rutter and colleagues (1979) argued that the effect of schools per se upon pupil performance had been neglected. They attempted to show that, even when social and other factors were taken into account, there remained differences among schools which could be ascribed to the quality of schooling itself. Despite the methodological weakness of some of this early work (Goldstein, 1997), school effectiveness research flourished during the 1980s and 1990s, becoming more sophisticated both in the kinds of data used and the statistical modelling techniques applied. It also appealed to the Conservative government from the late 1980s and the subsequent Labour government, at least partly because it implied that changing schools could affect performance and hence that educational policy was relevant to educational 'standards'. Thus, in 1997, the new Labour administration set up a powerful 'Standards and Effectiveness Unit' within the department for Education and Employment, headed by someone who had been involved prominently in school effectiveness research.

This 'incorporation' of SE within government policy making raises issues to do with the role of research and also with the way in which policymakers view and seek to use research. In this paper we will use the history of SE to explore some of these issues. This will be done

---

[1] Email correspondence should be addressed to h.goldstein@ioe.ac.uk

principally by looking at some recent critiques of the area focussing on the methodology, the usefulness and the political consequences of SE research. The critiques will be judged in terms of their coherence and the evidential support for the cases that they seek to make. These critiques have emerged largely from UK academic institutions, reflecting no doubt the high profile that SE has enjoyed within the research community and within Government. The issues, however, are general and apply to SE research carried out in other educational systems.

The critiques can be grouped under the following broad headings:

1. Abuse by Government

2. Oversimplification of the complex 'causalities' associated with schooling and sidetracking into focussing on 'league tables'.

3. That 'theory' in SE work is little more than reification of empirical relationships

4. Too much SE research is simply poor quality

Abuse by Government

Pring (1995), Elliot(1996), Hamilton(1996), Gibson & Asthana(1998) and Willmott (1999) all touch upon this issue, and to the extent that there is a strong continuity between the policies of both the recent Conservative and Labour governments their arguments apply with equal force to both. These arguments can be summarised by saying that government has taken up school effectiveness because it emphasises the responsibility that schools have for 'standards' rather than government itself. The 1992 Conservative administration first set up a school effectiveness unit and the New Labour government commitment is illustrated by the size of the 'Standards and Effectiveness Unit' created at the D*f*EE.

Hamilton suggests that SE research 'pathologises' schools by accepting that economic and other problems of society can be ascribed to failings of education and of those who work in the system, especially teachers. He also points to the 'remedies' which have been suggested for curing ailing institutions, using lists of 'key factors' associated with being an effective school; leadership, collegiality etc. The particular focus of his criticism is the OFSTED commissioned review 'Key characteristics of effective schools' (Sammons et al., 1995) which he suggests promotes the idea that there is a small number of such factors acting as mechanisms of effectiveness and that the identification of these allows action to be taken to

initiate 'improvements'. He suggests that the authors' claims too readily reflect policy requirements for simple solutions within a framework of crude social engineering.

Pring (1996) also sees the strong government concerns with holding schools responsible for society's problems as driving SE research in unproductive directions and sees SE researchers working 'under the watchful eye of their political mentors'. He also notes how some SE researchers appear to have accepted the official view that schools can and should be held responsible for economic and social improvement.

Elliot (1996) sees a strong government influence, with SE findings being 'music in the ears of politicians' and points out that politicians are highly selective and will tend to listen only to those whose music they like. He also criticises Sammons et al as supporting 'a mechanistic methodology, (and) an instrumental view of educational processes'. He goes further and regards SE activities as 'ideological legitimisations of a socially coercive view of schooling'. Elliot seems to be well informed about the SE literature and he gives a good critical account of the early beliefs of SE researchers in 'key factors' and how these were undermined by research in the late 1980s and early 1990s. His analysis of the attempts by Reynolds (1995) to both critique early findings while continuing to accept them is clear and insightful and he regards Sammons et al., (1995) as similarly wishing to have their cake and eat it. Elliot subjects the factors in Sammons et al to a logical analysis and argues that these factors reveal pre-existing *assumptions about the nature of education*, as much as the outcome of empirical investigation. He highlights the explicit statements by several SE researchers that they need to make their research relevant and to have practical application, for example to school improvement. He suggests that such a concern endears SE research to government and can lead both to a superficiality of interpretation and a concentration on pragmatics rather than theory.

Gibson and Asthana (1998) are explicit that SE research has reinforced government policies which are concerned with identifying schools as the sole agents of 'success' or failure' and they illustrate another important issue, namely that not only will governments selectively use findings, they and their advisors also often simply fail to understand them. Gibson and Asthana quote the Literacy task Force Report (Barber, 1997) statement that schools with similar intakes of pupils achieve widely differing results'. In fact the 'intake' measure referred to was not an intake measure at all but the percentage of children in a school taking free school meals.

A dominant theme in all these critiques is the desire of SE research to get close to government, and the consequent distortions that this has produced. The critics make some important observations, for example the case against the reality of 'key factors' is a strong one and it does seem clear (see also Goldstein, 1997) that there really is little evidential support for such lists. Yet some of the critiques are spoilt, either through an unfamiliarity with the range of the SE literature, or through simple misunderstandings. Thus, while Hamilton has pointed to some key problems in SE research his criticisms are in fact too all-inclusive. There is much SE research which recognises the real complexity of education, which is concerned with subtle processes rather than simple exam and test scores, and a great deal which is overtly critical of government attempts to subvert the research for its own ends (see Goldstein and Myers, 1996 for a discussion). Similarly, Gibson and Asthana, having made out a clear case for the abuse of SE research seek, unfairly, to generalise to all SE researchers.

Likewise, Pring reiterates the distinction made by Brown and Riddell, (1995) between school effectiveness and school improvement (SI) research. These latter authors claim that SE studies 'focus on outcomes that are readily measurable … and on generalisable findings', whereas school improvement is concerned with 'the individual institution with its distinctive processes and multiple goals'. In other words they identify SE with quantitative research and SI with qualitative case studies, which is a view by no means shared by those working in these areas.

Elliot also quotes with approval the views of Brown and Riddell, (1995) who believe that multilevel modelling necessitates a view of schooling 'which is essentially one of top down management'. Again we see the qualitative/quantitative divide emerging and this is especially unfortunate because, while there are serious problems with some of the ways statistical modelling has in fact been used and interpreted (see Mortimore and Goldstein, 1996), these models certainly have no built in requirement for any particular managerial structure.

Willmott sees inherent 'right wing' assumptions in SE research. Yet he is vague about what 'right wing' means and does concede that caution is required when using such a term. He goes further than most other critics in being very explicit about the political values associated with SE research, yet hardly goes beyond what is little more than an assignment of 'guilt by naming'.

The enthusiastic espousal of SE research, or at least selected elements of it, which has characterised educational policy, particularly that of the New Labour government, has been an

important factor in arousing criticism of the field. Even though some aspects of these critiques have been misplaced, the basic concerns expressed are important. SE researchers would do well to re-examine closely their relationship with government and policy making in general, and we shall return to this point later.

Oversimplifications

We have already touched on some of the criticisms made of the way that some SE research has been content to present simple sounding solutions to complex problems such as exemplified in lists of 'key factors'.

One of the enduring features of SE research is its conceptualisation of schools as discrete non-interacting entities which have characteristics derived from their student body, their staff as well as more permanent structures. These characteristics can, in principle, be measured and a study of the relationships among them, for example between intake and output achievements, is assumed to be capable of yielding insights into the processes of schooling, and possibly to indicate paths for positive change. In practice, of course, it is often extremely difficult to obtain good measures of the things which one might suppose matter, especially subtle process factors, motivations, classroom organisation, etc. and this is the basis for a particular set of critiques (see e.g. Brown et al., 1995). There is some evidence that SE researchers are taking note of such criticisms (Goldstein, 1997), but there is a fundamental problem with the basic assumption mentioned at the start of this paragraph, namely that schools are assumed to be *non-interacting* entities.

A cursory reflection on the way schools function, especially in England, reveals that the actions and characteristics of any one school are linked to those of other schools. Schools which, in one way or another select students on the basis of capabilities, influence the capacities of surrounding schools to do so. Indeed, the 'market' model of schooling introduced by the Conservative government in 1987 makes this explicit. Yet, there is very little attempt within SE research to take this into account. This failure underlines and gives credence to many of the critiques, yet interestingly enough none of the critics mentioned above describes it clearly. Mortimore and Whitty (1997) make a similar point about the impossibility of separating the functioning of schools from the constraints of the wider society.

Perhaps the most important consequence is in its impact on SI research and practice. Attempts to change the practices of each school considered as an independently functioning unit not only seem ill-conceived, but also likely to meet with limited success, however well-meaning the intention may be. Schools function within a social and political system which has its own structures and processes, whether these be ones of inter-school competition or those determined by externally imposed constraints of curriculum or resources. Despite the recognition of this by many within SE and SI (for example Mortimore and Whitty, 1997), there seems little in the research itself that seriously attempts to address this problem. While SE research has gone some way towards modelling within-school complexities, it has made almost no attempt to contextualise schools within the wider environment. To do so, of course, would involve political as well as social and cultural considerations and it is doubtful whether this would be welcome to government. Yet until SE and SI do move towards such a wider field of study the usefulness of their conclusions will remain limited.

Coe and FitzGibbon (1998) are particularly critical of the way in which much SE research identifies the 'unexplained' variation between schools, after adjusting for intake, as measuring 'effectiveness'. They point out that there may be other factors, outside the control of the school, which could explain such variation. In other words, the term 'effectiveness' needs to be treated tentatively when applied on the basis of models which may have omitted important factors. These authors also point out the need to develop models which take account of within-school structures.

SE 'theory'

A number of authors have attacked SE research on the grounds of its inadequate theoretical structure. This is perhaps best articulated by Lauder et al (1998) who point out that the early SE literature leant heavily on notions of 'ethos' and 'climate' in order to lay claim to a 'theoretical' basis for the field. In fact the definition of such terms is almost totally pragmatic, and measures of these concepts tend to be derived from a consideration of how well their component parts predict such things as student achievement, and the highly influential early work of Rutter et al (1979) was criticised for just this circularity of argument (see e.g. Goldstein, 1980). A similar criticism applies to the other notions such as 'leadership', which SE research has used to provide a theoretical-sounding support for its activities. Lauder et al. acknowledge that SE research has moved on and accumulated some valuable insights into the determinants of student achievement, and while they do not raise it explicitly, they are in fact

questioning whether SE, without any strong core theory, has a right to claim that it is a genuine field of study. We will return to this question later.

Lauder et al. go on to query whether the outcomes of SE research can be used to effect changes in schools. They point out that schools are complex organisations and that the culture and environment act as important constraints upon what can be achieved. They suggest that because SE research has not been able properly to define and incorporate such factors into its analysis, the application of SE findings will be unlikely to have a substantial effect. They also make the point, mentioned above, that studying each school in isolation is unrealistic. In their discussion they refer a great deal to the application of management theory ideas to schools, and develop what they refer to as a 'heretical' model from which they argue that attempts to change schools will be ineffective unless these attempts recognise the complexity of schools. In effect they are pointing to a problem with the attempt to link outcomes of SE research to SI which attempts to effect changes. This is another issue for the future of SE to which we will return. Finally, Lauder et al. describe what they call a 'contextual model' for SE. They advocate research which seeks to incorporate the wider political and social constraints into models of student progress, to include large scale quantitative work as well as case studies of 'good practice'. These ideas are strongly emphasised also by Riddell et al. (1998) who argue, on the basis of their own case study work, that the interactions between social background, parental behaviour, teachers and schools need to be taken into account. These suggestions echo those of Goldstein (1997) who sets out the requirements for valid inference from statistical modelling within SE research. Lauder et al. also suggest, at least initially, that a strong theoretical foundation may be unnecessary, so that SE research is perhaps best viewed as an attempt to continue what exists, jettisoning dubious pretensions to a basis in theory, while attempting to make it a more realistic description of reality.

Research Quality

All research fields contain work demonstrating a wide range of 'quality'. One measure of the health of a field of study is the extent to which it progresses by eliminating the poor quality work, through a shared recognition of what counts as 'good'. The standard way to judge this is to review the journals in the area over time. SE has been in existence long enough for such a judgement, perhaps provisionally, to be made. Since there appears to be no systematic attempt to do this, any comments will necessarily reflect personal preferences. Nevertheless,

it seems useful to attempt to identify the quality of contributions in order to assist the development of a shared recognition.

We have already alluded to criticism of the work of Reynolds, and some of his recent attempts in the area of international comparative studies do seem to exemplify weak methodology. In a report for OFSTED (Reynolds and Farrell, 1996) the authors draw causal conclusions about factors responsible for country differences in the international educational achievement studies carried out by the IEA, while in another part of the report they caution strongly against drawing any causal conclusions from such cross-sectional studies. This seems to be another example of SE research wanting to satisfy government demands for policy action while also attempting to portray itself as academically respectable. The 'International School Effectiveness Study' (Creemers et al., 1996) provides a further example of an inadequately conceived piece of research. It attempts to carry out comparative cross-systems analyses. It uses non randomly selected samples of schools within each country yet implicitly assumes representativeness and comparability. The number of schools involved in each country is between 5 and 12 which provides an extremely poor basis for coming to any definitive conclusions, Furthermore, there is no proper adjustment for intake achievement, so that this study cannot really claim to satisfy the basic requirements for valid SE research. Coe and FitzGibbon (1998) give other examples of 'causal' conclusions derived solely from cross-sectional studies.

We have already referred to the deficiencies in the report of the literacy task force. Another example of poor research (from the same principal author) which nevertheless achieved considerable publicity, is the report on homework provision (Barber et al., 1997) which concluded that increased 'homework' was associated with effectiveness. Yet it studied just 14 schools, used OFSTED inspectors judgements as to 'effectiveness' and was purely cross sectional. It is relevant to note that this same author has been principally responsible for developing the standards and Effectiveness Unit within the DfEE since 1997.

All these authors have, at least until recently, been prominent within the SE community, and this does lend some substance to the criticisms of the quality of SE work. It would seem that the quality of SE research in general will only improve if the SE community can itself eliminate much of such poor quality work from its discourses.

Quantitative methodology

One feature of the criticism of SE work is that it has emanated largely from those researchers not engaged in large scale quantitative work. The criticisms themselves also tend to focus on conclusions derived from quantitative analyses. Some critics claim that the process of quantitative, statistical, modelling itself necessarily limits SE research. For example Scott (1997) claims that multilevel modelling (MLM) techniques rely upon 'matched pairs of students', yet this is simply untrue. Willmott claims that statistical models are incapable of measuring the 'gap' between 'competence' and 'performance', yet accepts that they can measure the latter. He concludes in effect that 'competence' cannot be measured on the grounds that it is 'variable' – yet that is generally no bar to measurement of any quantity and is explicitly what statistical models are designed to do.

Statistical modelling, of course, is only as good as the data which it attempts to model. It is also often the case that such models oversimplify reality to the point of distorting it and producing misleading inferences. To point to specific inadequacies, however, or to list inappropriate uses of such models does not invalidate them per se. Statistical modelling of a complex kind can provide insights, and unexpected relationships. Furthermore, the use of sophisticated models, such as multilevel ones, has produced important counter arguments and demonstrations to crude attempts by governments and others to judge school performances on the basis of league tables. Also, much of what we know about the effects of factors such as social class is derived from careful statistical modelling.

Perhaps the most extreme criticism of the use of such models comes from Willmott who mounts a general attack on 'positivism'. The difficulty with this critique is that it is very difficult to recognise any real researchers as practising positivists – at least according to Willmott's definition. For example, he asserts that positivism is unable to take account of subtle variations of social relationships within institutions over time, yet there are many studies which do model changing relationships of various kinds over time – the main problem being the availability of data rather than with the technical procedures required to do the modelling.

Responses from the SE community and future directions

The concentration of critical assessments of SE research during the mid to late 1990s is a healthy sign and there is much here that the field would be well advised to take note of. This

needs to be done with care, however, since there is also much in these criticisms that is unjustified and some arises out of ignorance, as we have pointed out. Of particular concern is the failure of some of the 'qualitative' critics properly to understand the nature of quantitative techniques. This is quite serious since these critics do have valid points whose force is often dissipated through lack of proper understandings of the nature and substance of quantitative models.

Unfortunately, the response to criticisms from the SE community has been disappointing. Too often it has been over-defensive, and while this may be understandable, it does little to advance a constructive debate. For example, in their response to Hamilton, Sammons et al. (1996) claim that one of their purposes was to simplify the issues so that information would be 'accessible to non-researchers'. Yet they do not directly respond to his concern that the listing of 'key factors' encourages a mechanistic approach to policy making, although they do recognise the dangers of oversimplification. Interestingly Sammons et al. are clear that they believe researchers have a responsibility to disseminate their findings to policy makers and the general public, and in so doing appear to concur with the view of Hillage et al., (1998), in their review of research for the DfEE, that research should inform policy. In so doing Sammons et al. appear to support the view that SE researchers should become very closely involved with policy makers. The response of Sammons and Reynolds (1997) to Elliot's criticisms further underlines this. They are at pains to emphasise how SE research has focussed on critiques of league tables and how it needs to respond to policy requirements.

Likewise, Mortimore and Sammons (1997) are too sanguine about the ability of techniques such as MLM to 'make valid and reliable judgements'; the real problem is to obtain valid and reliable data and a failure to do that is the point of much of the criticism. Unfortunately Mortimore and Sammons have little to offer in terms of how practically we can obtain large scale long term longitudinal data that captures the relevant variables and allows exploration of subtle relationships across time. MLM is a powerful set of techniques for extracting interesting patterns from data but has nothing per se to do with the quality of the data presented to it. As OFSTED has demonstrated (Mortimore and Goldstein, 1996) it can also be used to obscure a weak research design.

Similarly, Reynolds and Teddlie (1999) praise the use of multilevel models with little appreciation of their limitations. They argue in favour of more inter-country comparative studies, without any indication of the severe practical and theoretical difficulties of so doing. They also appear to believe that the increased variability introduced by incorporating

between-country variation would allow more interesting comparisons - forgetting that they would then have to find ways of adjusting for all the other country level factors likely to be associated with such variation. In a similarly naïve vein, these authors also accept implicitly the belief, discussed above, that schools can be treated as independently acting institutions. Their discussion of 'effective' and 'sick' schools shows a serious misunderstanding in that they appear to regard the identification of such schools as relatively unproblematical, whereas in fact it is precisely the difficulty of defining 'effectiveness' that so disturbs many of the critics.

We have already indicated our belief that the future health of SE research depends on a collective recognition by the community that all is not well; that there is a clear need to come to agreement on separating the good from the bad and that there is a need to move onto a sounder methodological basis which, among other things, will require the collection of different kinds of data. The SE community would also do well to be more receptive and less defensive about its critics, and to avoid the kinds of superficial responses we have discussed above. Likewise, there is a need for some critics to demonstrate a deeper understanding of SE research, and also to recognise some of its positive achievements such as its exposure of the limitations and dangers associated with the publication of school performance league tables.

In terms of future directions, we have some sympathy with Lauder et al who question the need to develop 'theory'. It is not incumbent upon every research endeavour to provide a strong theoretical basis of the kind that allows interesting predictions and shapes our interpretation of the world being studied. There is, for example, often an important period during which empirical evidence needs to be accumulated before coherent theories can be developed. Naturally, such empirical accumulation makes particular theoretical assumptions about the phenomena it studies, but these can be *general* ones already in use rather than specific ones that attach exclusively to the particular field of enquiry. Thus, a particular view of institutional structures within a quasi-market based capitalist system might be a useful basis for designing SE projects, but other frameworks, for example with a psychological emphasis, might also provide evidence that can provide insights into schooling.

The choice of framework will itself determine the nature of any inferences which are drawn, and different frameworks can lead to real or apparent conflicts. It is important to appreciate this, since there is a notable lack in the current school effectiveness literature of serious attempts to expose the underlying assumptions that the research is making. Prior to developing useful theories, existing assumptions need to be exposed and questioned and this is where many of the critiques are particularly useful. Thus, Elliot (1996) is concerned with

the underlying assumptions made about educational processes and Gibson and Asthana (1998) point out how research has uncritically accepted government assumptions about the role of schools.

In our view there are several key things which need to happen if SE research is to thrive.

First, there has to be a clear separation of SE research from government influence. Of course, many SE findings will have relevance to government agendas and one way of viewing the role of SE researchers is as that of 'critical friends', who can contribute to general public policy debates while maintaining their integrity, especially when ostracised or threatened by government. This may be difficult when government itself tends to regard the research community as consisting of either committed friends or subversive enemies. Furthermore, an explicit attempt to distance a field of enquiry from short term practical application runs counter to many contemporary official views of the role of social, and particularly educational research (Hillage et al. 1998), but that is ultimately a problem for government rather than for research..

The existence of something called the 'Standards and effectiveness unit' at the DfEE exemplifies the way in which government, with the support of certain key individuals, drawn originally from the research community, has attempted to appropriate SE. For this reason, and also because SE is no longer a proper description of the research area which is concerned with institutions other than schools, and with components of schools such as classes and other groupings, we suggest that the area renames itself as that of Educational Effectiveness (EE), with the relevant journals and Centres altering their titles accordingly. Apart from anything else this would symbolically distance it from politicians.

While distancing itself from politicians, EE should consider carefully its relationship with School Improvement (SI) which also perhaps should consider calling itself something different, like Institutional Change (IC). There are certainly many opportunities for mutual learning across these two areas of study, but neither exists merely to serve the other. Thus, while insights from EE may inform both research and practice in IC, there is no formal requirement that EE *needs* to do this. Educational effectiveness should attempt to become a field of study which, while it may result in practical policies, has as its primary justification an increase in knowledge and understanding. It follows that any movement to bring EE and IC into a unitary framework (see for example, Reynolds, 1995), with EE research linked to the practical concerns of IC, may well be misplaced.

The requirements for better data, for a concern with knowledge not driven by particular policy agendas and for attention to the possibility of whether and how to develop well grounded theory, pose a considerable challenge to EE researchers. Unless this challenge is met successfully EE will not establish itself as a reputable field of study. In particular, if it continues to exist as SE, if many of its proponents remain superficially defensive and it ignores or fails to understand the warnings of its critics, we have very little optimism that it will survive its present state of adolescent turmoil to emerge into a full maturity.

References

Barber, M., Denning, T., Graham, J. and Johnson, M. (1997). *School performance and extra-curricular provision*. London, Department for Education and Employment.

Brown, S. and Riddell, S. (1995). School effectiveness research: the policy makers' tool for school improvement? *European Educational Research Journal*(March 1995): 6-15.

Coe, R. and FitzGibbon, C. T. (1998). School effectiveness research: criticisms and recommendations. *Oxford review of education* **24**: 421-438.

Coleman, J., Campbell, E., Hobson, C., MacPartland, J., et al. (1966). *Equality of Educational Opportunity*. Washington, D.C., US Government Printing Office.

Creemers, B. P. M., Reynolds, D., Stringfield, S. and Teddlie, C. (1996). *World class schools: some further findings.* Paper presented at American Educational Research Association.

Elliot, J. (1996). School effectiveness research and its critics: alternative visions of schooling. *Cambridge Journal of Education* **26**: 199-223.

Gibson, A. and Asthana, S. (1998). Schools, pupils and examination results: contextualising school performance. *British Educational Research Journal* **24**: 269-82.

Goldstein, H. (1980). The Statistical procedures in Fifteen Thousand Hours. *Fifteen Thousand Hours: a discussion.* B. Tizard. London, Institute of Education.

Goldstein, H. (1997). Methods in school effectiveness research. *School effectiveness and school improvement.* **8**: 369-95.

Goldstein, H. and Myers, K. (1996). Freedom of information: towards a code of ethics for performance indicators. *Research Intelligence*(57): 12-16.

Hamilton, D. (1996). Peddling feel good fictions. *Forum* **38**: 54-56.

Hillage, J., Pearson, R., Anderson, A. and Tamkin, P. (1998). *Excellence in research on schools.* London, DfEE.

Jencks, C. S., Smith, M., Ackland, H., Bane, M. J., et al. (1972). *Inequality: A reassessment of the effect of family and schooling in America..* New York, Basic Books:

Lauder, H., Jamieson, I. and Wikeley, F. (1998). Models of effective schools: limits and capabilities. *School effectiveness for whom?* R. Slee, G. Weiner and S. Tomlinson. London,

Falmer.

Mortimore, P. and Goldstein, H. (1996). *The teaching of reading in 45 Inner london Primary schools: a critical examination of OFSTED research.* London, Institute of Education.

Mortimore, P. and Whitty, G. (1997). *can school improvement overcome the effects of disadvantage?* london, Institute of Education.

Mortimore, P. and Sammons, P. (1997). Endpiece: a welcome and a riposte to critics. *Perspectives on school effectiveness and school imporvement.* J. White and M. Barber. London, Institute of Education.

Pring, R. (1995). Educating persons: putting education back into educational research (The 1995 SERA lecture). *Scottish Educational Review* **27**: 101-112.

Reynolds, D. (1995). The effective school: an inaugural lecture. *Evaluation and research in education.* **9**: 57-73.

Reynolds, D. and Farrell, S. (1996). *Worlds apart?* London, OFSTED.

Reynolds, D., Creemers, B. P. M., Stringfield, S. and Teddlie, C. (1998). Climbing and educational mountain: conducting an international school effectiveness research project. *Doing research about education.* G. Walford. Lewes, Falmer Press.

Reynolds, D. and Teddlie, C. (1999). The future agenda of studies into the effectiveness of schools. *Enhancing educational excellence, equity and efficiency.* R. J. Bosker, B. P. M. Creemers and S. Stringfield. Dordrecht, Kluwer.

Riddell, S., Brown, S. and Duffield, J. (1998). The utility of qualitative research for influencing policy and practice on school effectiveness. *School effectiveness for whom?* R. Slee, G. Weiner and S. Tomlinson. London, Falmer

Rutter, M., Maughan, B., Mortimore, P. and Ouston, J. (1979). *Fifteen Thousand Hours.* London, Open Books:

Sammons, P., Mortimore, P. and Hillman, J. (1996). Key characteristics of effective schools: a response to "Peddling feel-good fictions. *Forum* **38**: 88-90.

Sammons, P. and Reynolds, D. (1997). A partisan evaluation - John Elliot on School Effectiveness. *Cambridge Journal of Education* **27**: 123-135.

Sammons, P., Hillman, J. and Mortimore, P. (1995). *Key characteristics of effective schools.*

London, Institute of Education/OFSTED.

Scott, D. (1997). The missing hermeneutical dimension in mathematical modelling of school effectiveness. *Perspectives on school effectiveness and school improvement.* J. White and M. Barber. London, Institute of Education.

Willmott, R. (1999). School effectiveness research: an ideological commitment? *Journal of philosophy of education,* 33, 253-268.

Abstract

This paper discusses a series of recent critiques of school effectiveness (SE) research from within the academic community and the responses to them by SE researchers. It uses these as a basis to explore the nature of current SE research and its relationship with government policy. It is argued that much SE research has been to closely concerned with specific government policies as well as having weak theoretical and empirical support. The general response of the SE community to these criticisms is judged to be inadequate and recommendations for future directions are made.

Acknowledgements

Note on Authors

Harvey Goldstein is professor of statistical methods at the Institute of Education. His principal research interest is in the development of techniques and software for multilevel modelling. He has a particular interest in the application of these techniques to the study of educational processes and is concerned with how these techniques can be used constructively for school change (school improvement) purposes.

Geoffrey Woodhouse was a lecturer in statistics at the Institute of Education from 1989 to 2000. His principal research interests have been concerned with the influence of measurement errors on inferences from multilevel models and the uses of performance tables for school comparisons.