

Presented at British Educational Research Association meeting,  
York, September 11-14, 1997

**School effectiveness research: a bandwagon, a hijack or a journey towards enlightenment?**

By

Harvey Goldstein and Kate Myers<sup>1</sup>

Institute of Education

University of London

**1. School effectiveness as political legitimation**

Under the New Labour Government school effectiveness seems set to continue to occupy a high profile, associated with policy imperatives to 'raise standards' and 'achieve targets'. In the 20 or so years during which school effectiveness research has existed, it has commanded a remarkably high level of political interest, and has been viewed by many politicians as a legitimating device for a wide range of new policies. In this paper, we start by looking a little more closely at this phenomenon, and then move on to discuss a more positive potential for school effectiveness (SE) research.

Several academics (Pring (1996), Hamilton (1996), Elliot (1996), White and Barber (1997)) have recently been critical of SE research, and have drawn responses from SE researchers (Sammons et al, 1996; Mortimore and Sammons, 1997; Mortimore and Whitty, 1997). The criticisms have centred around 3 main issues. The first is that SE research has claimed too much for its 'findings', and this is a view with which we have some sympathy and to which we shall return. The second accusation is that SE *inevitably* concentrates on restricted 'cognitive' outcomes of schooling and ignores the many other aspects which are important. With this view we have little sympathy. We believe that it misunderstands the nature of current research and we will elaborate below. The third charge against SE research is that it has assisted in the process of governmental centralisation and control of education and educational professionals. We both agree and disagree with this!

We disagree because we do not accept that SE researchers as a group have consciously supported such government moves, although would be prepared to concede that some individuals involved in SE may be culpable. Nevertheless, we do agree that the government and its quangos have 'cherry picked' what they wish to use in order to help to legitimate their policies. There is no particular shortage of examples. Many research discussions have been quoted out of context e.g. OFSTED's work on reading (OFSTED, 1996) has sought to justify some dubious 'research' by appealing to aspects of the SE literature (Mortimore & Goldstein, 1996) and the literacy task force report produced for the Labour party sought to justify its comparisons among primary schools by questionable references to 'intake adjustments' (Goldstein, 1997).

Worst of all, the school effectiveness research base is used to justify blaming schools for 'failing' on the assumption that because some schools succeed in difficult circumstances the reason others do not must be their own fault. In this scenario complexity and context are ignored. Furthermore

---

<sup>1</sup> Kate Myers is now Professor of Education at Keele University.

some politicians and policymakers have found it possible to deny their role in 'failure' by shifting all the blame onto individual schools (Myers and Goldstein 1997)

In evaluating SE research we need to recognise that it is a very young activity, struggling to find its identity and, most recently, becoming aware of the need to distance itself from the short term demands of policy makers. Indeed, we believe that the future health of the area depends crucially on establishing SE activity as an autonomous area of study rather than by always responding to immediate external requests. Whilst it should certainly be addressing policy concerns it should not be dependent on a too-cosy relationship with politicians.

## **2. Towards enlightenment**

The last 20 years have seen a development of understanding of school effectiveness from the early studies such as those of Edmonds (1979) and Rutter et al (1979), through the writings of Fullan (1982) to research of Mortimore et al (1988) and to the more recent work by Gray et al (1996) and Hill and Rowe (1997). There are already a number of lessons we can learn to help us think about the future.

The first lesson, and perhaps the most important, is that the term 'school effectiveness' is a misnomer. Effectiveness, if it means anything, is multidimensional. Schools differ in their effectiveness by curriculum subject and are differentially effective for different groups of pupils: their effectiveness also changes over time (Gray et al, 1996; Thomas et al., 1997a; Thomas et al., 1997b). The absence of a single continuum of 'effectiveness' raises issues about 'effective schools', and leads to questions about the importance of 'leadership' and the relevance of other school level factors such as 'common clear goals'. Such notions require not only a more careful definition but a detailed explication of how schools, teachers and other factors interact to produce the complexity of relationships that empirical studies are beginning to unearth. This is a role here for careful theory and as we see it one of the major challenges for those working in this area.

The second lesson is that we need to tailor our study designs and the analysis of data to match the real complexity of schooling processes. It is now almost universally agreed among researchers that a necessary condition for valid inferences about 'effectiveness' is the existence of longitudinal data on pupils following them through their school careers. Indeed, recent evidence (Goldstein and Sammons, 1997) suggests that this should encompass more than one stage of schooling, since, for example, *both* the pupils' primary as well as their secondary school affect achievement at the end of secondary schooling. While such very long term research is desirable, there is also a place for more intensively sampled longitudinal data, extending over only a few years but with frequent measurements to chart short term changes and influences. An example of this is work looking at individual teacher effects (for example Hill and Goldstein, 1997).

In addition to following students through schooling, it is important to realise that schools are complex interactive structures and this again requires study at a very detailed level. It seems to us that we need both more intensive case study approaches to illuminate such detail as well as larger scale quantitative studies to allow a more detailed modelling of the processes and their effects. Such work is not easy, and it also is not cheap. Yet, unless SE research moves into this area it is not going to provide new insights: nor is it going to become the intellectually challenging area of study that it could.

Finally, if it is accepted that we have to match our analytical tools as closely as possible to the complexity of the real world of schooling, then we need to spend some effort in developing those tools. At the basic level of measurement we have relatively primitive instruments. Certainly, the use of exam results and crudely graded test scores at a few stages during a pupil's educational

career are not adequate. Affective and behavioural measures such as those developed by MacBeath et al (1996) are needed as well as more accurate and reliable cognitive measures, and it will be important to study the interrelationships among these over time. Detailed measurements on pupils, including their out-of-school activities and home backgrounds are important. We need comprehensive information on school and teacher behaviours and the ways in which they change over time, and we need to take into account the general environment, social and political, within which schools operate. All of this should inform both small scale and large scale investigations.

### **3. Technicalities of School Effectiveness research**

A somewhat unfortunate feature of existing SE research is the general separation of what has been labelled 'qualitative' from 'quantitative' work. The former has tended to adopt a detailed descriptive role of studying school processes while the latter has tended, with one or two exceptions, to make use of whatever data happen to be available to draw conclusions. This is both unnecessary and inefficient. If the field of SE research is to achieve coherence it needs to agree on common approaches to what constitutes adequate measurement and on what constitutes an appropriate set of analytical procedures for summarising, modelling and drawing conclusions. We believe that one of the more fruitful developments of recent years has been the application of multilevel modelling techniques in this area. The main thrust of these statistical models is to attempt to set up mathematical descriptions which are sufficiently complex that they will be able to capture the real life complexity which exists. Thus, for example, Hill and Goldstein (1998) have shown how to cope with student mobility across classes within schools and movement between schools. Techniques now exist for handling more generally multidimensional structures which change over time. Interestingly, one of the successes of applying these models to current debates about school accountability has been to show how, when a description of sufficient complexity is adopted, simple-minded approaches to comparing schools through either raw or 'value-added' league tables are inadequate and potentially misleading.

One of the problems with the adoption of sophisticated statistical techniques is that they may become difficult to explain. This has led some commentators to reject their use in favour of 'simpler' techniques. Unfortunately this can easily lead to *oversimplification* as when, for example, straight line, simple regression relationships are used in 'value added' analyses (SCAA, 1997), leading to incorrect inferences (O'Donoghue et al., 1997). While we accept the difficulties associated with the explanation of complex analyses this should not become an excuse for using inferior or misleading methodologies. Rather, ways need to be found to enable this to take place.

The criticisms of some commentators seem to reflect a misunderstanding of the technicalities. Thus, for example, Winch (1997) is unaware that there are solutions to the problem of school mobility when fitting multilevel models (Hill and Goldstein, 1997). The concerns of Scott (1997) that mathematical modelling cannot cope with 'intentions, beliefs, (&) propositional attitudes can be addressed, and there is nothing in the technicalities of modelling which precludes the study of attitudes, beliefs etc, although we do accept that this is far from straightforward. Scott's worries about comparisons over time have also been studied in technical literature (Beaton and Zwick, 1990, Goldstein and Wood, 1989).

Finally, some commentators appear to be arguing against school effectiveness research on the grounds that many of its findings are logically trivial. Thus White (1997) suggests that many findings could have been acquired on the basis of purely *logical* reasoning without any empirical observation. Such views seriously misunderstand the nature of empirical investigation and make it very difficult to engage in *any* debate about modelling in school effectiveness studies.

#### **4. Slowing down the bandwagon**

School effectiveness research has been something of a bandwagon. Its enthusiastic espousal by politicians, while suggesting success, also raises warnings. Thus, the bracketing of 'school effectiveness' with 'standards' in the title of a new unit within the DfEE seems to indicate that the politicians view SE as a key element in 'raising standards'. This is unfortunate: SE research has no necessary direct connection with 'standards' - and it is by no means clear even how we are suppose to define the term 'standards' . We certainly welcome any desire by politicians to be informed by research findings, and it is quite proper that SE findings, if appropriate, be used to affect policy on, say, student achievement, just as the findings of a survey into the nutritional content of school lunches might do so. Nevertheless, SE exists in its own right and to link it too closely with a particular political programme can only be to its detriment.

Clearly there is a lot in a name: if we are to strive towards enlightenment then a change of name might be rather a good place to start. Perhaps the DfEE could be persuaded to change the name of its unit to something more descriptive of what it actually does? Our own suggestion, which at least has a memorable acronym, is the unit for "Teacher Evaluation, Student Testing and Institutional Targets". If that fails then we may have to re-label ourselves. Something which is politician-proof would be good, but in the meantime, and in all seriousness, the term 'educational effectiveness' is a more accurate description of what we are really about.

## References

- Beaton, A. E. and Zwick, R. (1990). *Disentangling the NAEP 1985-1986 reading anomaly*. Princeton, Educational Testing Service:
- Edmonds, R. R. (1979) Effective schools for the urban poor, *Educational Leadership*, 37 (1), 15-27.
- Elliott, J. (1996) School effectiveness research and its critics: alternative visions of schooling, *Cambridge Journal of Education*, 26 (2), 199-223.
- Fullan M G (1982) *The Meaning of Educational Change* Toronto: OISE Press
- Goldstein, H., (1997). *The Labour Party Literacy Task Group report - a Critique*. Education-Line document (<http://www.leeds.ac.uk/educol/>).
- Goldstein, H. and Wood, R. (1989). Five decades of item response modelling. *British Journal of mathematical and statistical psychology* 42: 139-167.
- Goldstein, H., and Sammons, P. (1997) . The influence of secondary and junior schools on sixteen year examination performance: a cross-classified multilevel analysis. *School effectiveness and school improvement*. 8: 219-230.
- Gray, J., Goldstein, H. and Jesson, D. (1996). Changes and improvements in schools' effectiveness: trends over five years. *Research papers in education* 11: 35-51.
- Hamilton, D. (1996) Fordism by fiat, *Forum*, 38 (2), 54-56.
- Hill, P. and Rowe, K. J. (1997). *Modelling student progress in studies of educational effectiveness*. Melbourne, University of Melbourne.
- Hill, P. W. and Goldstein, H. (1997). Multilevel modelling of educational data with cross classification and missing identification of units. *Journal of Educational and behavioural Statistics*. To appear.
- MacBeath, J., Boyd, B., Rand, J., Bell, S., *Schools Speak For Themselves. Towards A Framework For Self-Evaluation* NUT/QIE, University of Strathclyde
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D. and Ecob, R. (1988)  
*School Matters: The Junior Years*. Wells, Open Books.
- Mortimore, P. and Sammons, P. (1997). In: White, J. and Barber, M. (Eds.) (1997). *Perspectives on School Effectiveness and School improvement*. London, Institute of Education, Bedford way Papers.
- Mortimore, P. and Whitty, G. (1997). *Can School improvement overcome the effects of disadvantage?* London, Institute of Education, Occasional Paper.

Myers, K. and Goldstein, H. (1997 forthcoming) *Failing schools or a failing system?* in Stoll, L., and Myers, K. No Quick Fixes: Perspectives on Schools in Difficulty Falmer Press

O'Donoghue, C., Thomas, S., Goldstein, H. and Knight, T. (1997). *DFEE study of value added for 16-18 year olds in England*. London, Department for Education and Employment & London, Institute of Education.

OFSTED (1996). *The teaching of reading in 45 Inner London primary schools*. London, Office for Standards in Education.

Pring, R. (1996) Educating persons: putting education back into educational research, Scottish Educational Review, 27 (2), 101-112.

Rutter, M., Maughan, B., Mortimore, P. and Ouston, J. (1979)

Fifteen Thousand Hours: Secondary Schools and Their Effects on Children. 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.

SCAA (1997). *The value added National Report*. London, Schools Curriculum and Assessment Authority.

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.

Thomas, S., Sammons, P., Mortimore, P. and Smees, R. (1997a). Stability and consistency in secondary schools' effects on students GCSE outcomes over three years. *School effectiveness and school improvement*. 8: 169-197.

White, J. and Barber, M. (Eds.) (1997). *Perspectives on School Effectiveness and School improvement*. London, Institute of Education, Bedford way Papers

Winch, C. (1997). Accountability, controversy and school effectiveness research. *Perspectives on school effectiveness and school improvement*. J. White and M. Barber. London, Institute of Education.

---