

**A SCIENCE IN THE SERVICE OF AN ART?  
THE USE OF 'VALUE ADDED' ANALYSES OF SCHOOL  
PERFORMANCE TO AID SCHOOL IMPROVEMENT**

**LESLEY SAUNDERS**

**Thesis submitted to the University of Oxford for the degree of D.Phil.**

**Kellogg College**

**Trinity Term 2001**

## CONTENTS

**Thesis title:** A science in the service of an art? The use of 'value added' analyses of school performance to aid school improvement

**Abstract**

**Acknowledgements**

**Chapter 1** Introduction: aims and scope of the project

***PART ONE: REVIEW OF THE RESEARCH SO FAR***

**Chapter 2** **Defining and Measuring Educational 'Value Added': Some Functional Ambiguities**

- 2.1 Introduction: 'choice and diversity' in the definition of value added?
- 2.2 Value added in an educational context
- 2.3 Why 'value added'? Pressures, precursors and presumptions
- 2.4 Conclusions

**Chapter 3** **Angels on a Pinhead? Review of the Academic Literature on Educational Value Added**

- 3.1 Introduction: clarifying the line of inquiry
- 3.2 The provenance of value added principles in educational research relating to the compulsory sector
- 3.3 Changes over time in the relationship between 'value added' and school effectiveness discourse
- 3.4 A note on analytical models and methodological strategies used in the measurement of value added
- 3.5 What have value added analyses revealed about 'effectiveness' and educational quality? A short summary
- 3.6 What further questions and queries do value added analyses raise?

<b>Chapter 4</b>	<b>A Sharper Stick to Beat Them With? Review of the Policy Literature on Educational Value Added</b>
4.1	Introduction
4.2	Changes in official policy on value added
4.3	The value added national project
4.4	Reconciling academic criteria with policy requirements
<b>Chapter 5</b>	<b>Getting Beyond the Numbers Game: the Evidence on How Schools Use Value Added Data</b>
5.1	Introduction
5.2	Schools' use of value added data: the relationship between 'school effectiveness' and 'school improvement'?
5.3	The evidence on schools' use of value added data
5.4	Issues for further research
<b><i>PART TWO:</i></b>	<b><i>THE FIELDWORK AND ITS IMPLICATIONS</i></b>
<b>Chapter 6</b>	<b>Framework, Methodology and Protocol for Conducting the Field-Based Research</b>
6.1	Introduction
6.2	Fieldwork aims and objectives
6.3	Research issues
6.4	Design of the study
6.5	Setting up the study
6.6	Research protocol
6.7	Main methodological problems encountered
6.8	Subsidiary evidence
<b>Chapter 7</b>	<b>Analysis of Case Studies I: Descriptive Framework for Organising Case Study Evidence and Preliminary Findings</b>
7.1	Introduction
7.2	The case studies
7.3	Overview of preliminary findings and issues raised

## **Chapter 8**

### **Analysis of Case Studies II: Further Exploration of QUASE Database Information at School Level**

- 8.1 Introduction: issues for further exploration
- 8.2 Construction and analysis of school records
- 8.3 Exploration of schools' 'relative effectiveness'
- 8.4 Exploration of departmental effectiveness
- 8.5 Summary and Conclusions

## **Chapter 9**

### **Analysis of Case Studies III: Discussion of the Uses of QUASE Data with Reference to Different Participants' 'Narratives'**

- 9.1 Introduction: the task
- 9.2 Taking the analysis further: some issues and false starts
- 9.3 Headteachers: narratives of leadership and school improvement
- 9.4 Other senior managers: narratives of counterpoint?
- 9.5 Middle managers: narratives of leadership and teaching
- 9.6 Teachers' narratives
- 9.7 Governors' narratives: a watching brief?
- 9.8 Summary and conclusions

## ***PART THREE: CONCLUSIONS***

## **Chapter 10**

### **Value Added and School Development: Some Examples of QUASE Analyses in Use**

- 10.1 Introduction: using QUASE analyses developmentally
- 10.2 A note on the presentation of data
- 10.3 Example 1: Using the data to ask questions about teaching and learning
- 10.4 Example 2: Developing a dialogue around QUASE data
- 10.5 Good practice in using value added for school improvement

10.6 Conclusions: improving the use of value added data

## **Chapter 11**

### **Summary and Conclusions: Some Ways Forward for Value Added and School Improvement**

- 11.1 Introduction: concluding the study
- 11.2 Summary of key findings from the study
- 11.3 Disquisition on the findings
- 11.4 Reprise of the main problems encountered in the study: how to 'complete the description'?
- 11.5 Conclusions: ways forward for value added and school improvement?

## **REFERENCES**

## **APPENDICES:**

- A. QUASE Sample Profile (tables and diagrams)**
- B. QUASE Sample Profile (text)**
- C. Invitational letter to QUASE schools**
- D. Holding letter to respondents**
- E. Research Information Sheet**
- F. Confirmatory letter**
- G. Non-participant letter**
- H. Semi-structured interview schedule**
- I. Sample Information for Governors**

Approximate length of thesis = 72,850 words

**A SCIENCE IN THE SERVICE OF AN ART? THE USE OF ‘VALUE  
ADDED’ ANALYSES OF SCHOOL PERFORMANCE TO AID  
SCHOOL IMPROVEMENT**

**Lesley Saunders, Kellogg College, Trinity Term 2001  
Thesis submitted for the degree of D.Phil.: ABSTRACT**

The thesis is concerned to explore whether and how ‘value added’ data analyses can contribute to school improvement, and to identify some of the conditions under which this might be so. In the course of conducting the study, the author experienced a tension between the ‘research’ and the ‘development’ dimensions of her work, and this is used to inform the outcomes of the thesis.

The thesis is underpinned by three related aims: first, to provide a historical and evaluative overview of how the idea of ‘value added’ came to enter and influence the debate on educational quality in England. The study of the literature demonstrates that the main principles of ‘value added’ were already well developed before the term was in common use; it also reveals that the ambiguities in the term are not merely reflective of disagreements about how best to calculate value added but actually central to how the idea has been made to function within a particular political agenda for education having to do with ‘raising standards’. Because of the relentlessness of that agenda, ‘value added’ measures of attainment have undergone considerable methodological development over the past ten years, to the point where sophisticated statistical data on pupils’ and schools’ performance is being generated and used by government agencies, LEAs and schools themselves to an extent virtually unknown elsewhere. However, this thesis indicates that the technical and conceptual issues involved in putting such data to practical use in schools are likely to test the interpretative and organisational skills of users. The literature search confirms that not much investigation has been done into how data is actually used, but that what there is suggests some important lines of inquiry. The second aim of the thesis is accordingly to explore, through a small-scale empirical study, the use by secondary school staff of value-added data as exemplified by the NFER’s value added service QUASE. The study was conducted in nine schools, with staff at senior and middle management levels, and focused on mathematics, English and science departments. The third aim of the thesis is to assess how far the case-study findings might shed further light on the issues entailed in using such data for school improvement. The evidence suggested that value added data are seen as complex and often ‘high stakes’ and that – at the time of this study – the uses of value added data were rather more limited than expected; furthermore, the meanings of value added would seem to be socially constructed by the political and institutional environment, and to be closely related to individual teachers’ values and attitudes. This in turn suggests that better insights into, and management of, ‘the psychology and sociology’ of how value added data are perceived and used are necessary.

Nonetheless, the study concludes that there is potential for value added analyses to contribute to school improvement under certain conditions; crucially, the study indicated that these included a culture which emphasised self-evaluation – rather than external accountability – within the school or subject department, combined with input from a ‘champion’ or facilitator who understood the technicalities and significance of the data. Value added analyses seemed to be used most actively by staff who were able to use them ‘heuristically’, that is, to pose informed questions about teaching and learning (rather than as literal truths or statistical fictions). The study argues that ‘value added’ measures of performance are accordingly better understood as a technology than a science – that is, a practical application of knowledge which interacts dynamically with its social and cultural environment.

## ACKNOWLEDGEMENTS

The fact that this thesis exists at all owes a great deal to two people: Geoffrey Walford, Reader in Education Policy in the Department of Educational Studies, Oxford University, who, in the tactfulness of his tutorial guidance and the acuteness of the questions he posed, gave the best possible encouragement to undertake and complete this study; and Malcolm Rigg (Managing Director of the Survey Division at the British Market Research Bureau) who, through his constancy of love and practical support as well as his insights into aspects of the debate on ‘value added’, has been a true companion throughout my pre-occupation with the subject.

Next, I should like to thank the headteachers and staff in the schools who participated in the study, by giving generously of their time to contribute to the evidence, views and insights on which this thesis is based. Their willingness to spend time being interviewed is doubly valued in the light of all the other statutory and non-statutory demands made of them – especially in the area of data production and analysis.

I owe a large debt of gratitude to the National Foundation for Educational Research, where I worked, which published (through the Membership Programme) the reviews that summarised the first section of this thesis and which also funded the fieldwork study which makes up the second section. Moreover, I was particularly fortunate that the NFER gave me professional colleagues whose contributions and support I deeply appreciate. My colleague Peter Rudd carried out fieldwork in five of the schools – ably assisted by Deborah Davies – and made his excellent fieldnotes available to me; he was then kind enough to check that what I had written using his material accorded with his view of what was actually happening in the schools.

The bibliographical dimension of the work would not have been possible without Pauline Benefield, NFER Research Librarian, who not only managed to find everything I asked for but always seemed to know what it was I wanted before I asked for it! Christine Taylor, also a member of the Library staff, gave generously of her own time to correct, update and merge the references.

The knowledge and expertise unstintingly contributed by Dougal Hutchison, Jo Morrison, Lesley Kendall and Ian Schagen, my colleagues in the Statistics Research and Analysis Group at NFER, were central to this work, not only in providing additional analyses of schools' performance data and generating tables and diagrams, but also in encouraging me to draw the right inferences from the data and from the wider work of their team in the area of 'value added'. Keren Beddow, in Field Research Services, helped me to follow up different trains of thought by digging deep into the schools database. Mark Rickinson gave me a very helpful and detailed commentary on the penultimate draft of the thesis, and thereby a very welcome sense of collegial support. Julia Rose used her considerable lay-out skills to make the final draft presentable.

I am conscious of the great benefit I have derived from discussions and exchanges with eminent scholars whose insights have enriched my understanding of specific issues in the value added field, in particular Professor Harvey Goldstein (of the Institute of Education, London University), Professor John Gray (Homerton College, Cambridge University), Dr Alma Harris (Centre for Teacher and School Development, University of Nottingham), Dr Pam Sammons and Dr Sally Thomas (both of the Institute of Education, London University); and, in relation to the conceptual and ethical tensions between 'exploratory' and 'evaluative' research, Professor Martyn Hammersley (The Open University). I was also very glad to have the ideas, support and constructive criticism of Dr Sheila Stoney, Assistant Director at the NFER, Ralph Tabberer, now Chief Executive at the Teacher Training Agency, and Dr Philip Davies, Lecturer in the Department of Continuing Education at Oxford University. Any flaws which remain in the text are entirely due to my oversight or obstinacy.

The Economic and Social Research Council awarded me a studentship for the period of my study, for which I was very grateful.

In the course of conducting the doctoral research, I have adapted and submitted some parts of this thesis for publication; these have appeared in the places listed below, and I am grateful to the publishers/editors:

SAUNDERS, L. (1998). *'Value Added' Measurement of School Effectiveness: an Overview*. Slough: NFER.

SAUNDERS, L. (1999). *Value Added Measurement of School Effectiveness: a Critical Review*. Slough: NFER.

SAUNDERS, L. (1999). 'A brief history of educational "value added": how did we get to where we are?' *School Effectiveness and School Improvement*, **10**, 2, 233–56.

SAUNDERS, L. (2000). 'Understanding schools' use of "value added" data: the psychology and sociology of numbers', *Research Papers in Education*, **15**, 3, 241-58.

SAUNDERS, L. and RUDD, P. (1999). 'Schools' use of "value added" data: a science in the service of an art?' Paper presented at the British Educational Research Association Annual Conference, University of Sussex, Sussex, 3 September. Available on-line.

## ***CHAPTER ONE***

### **INTRODUCTION: AIMS AND SCOPE OF THE PROJECT**

This introductory chapter provides a brief background to the subject of the thesis, outlining the scope and scale of the research together with its aims and objectives; it shows briefly how each chapter contributes to the whole.

The background against which the work is set is the policy agenda for rendering schools more accountable in terms of student performance and for raising the levels of student attainment. The thesis was able to take as its starting point the idea that ‘value added’ data on performance is now a major feature in the educational landscape – although the thesis was begun before the introduction of the national value added system in autumn 1998. As I argue in Chapter Eleven, this decision could be seen as the culmination of a decade of sustained and public argument about how to measure the performance of pupils in the nation’s schools in a way which sheds light on progress as well as standards. At the very least, the academic debate on school effectiveness and how to measure it is now integrally linked with the national political agenda for educational quality.

This in turn means that the range of parties legitimately interested in value added extends from politicians to school senior managers, and from academic researchers to lay governors. There is consequently a need for continuing discussion not only of the technical questions – how most sensitively and/or informatively to measure added value – but also of the different expectations and requirements of different ‘stakeholders’ and users: the psychology and sociology of the numbers, as it might be said.

Before the introduction of the national value added system, various methods had been – and continue to be – tested and developed for ways of producing information which can be used, on the one hand, to assess the relative effectiveness of different schools or local education authorities (LEAs) and, on the other, to provide diagnostic assistance for school/LEA managers and staff. So there is no shortage of information, particularly of a quantitative kind, about the performance of schools in England and Wales: school managers are now required both to generate and to be on the receiving end of an unprecedented amount of data, covering everything from test/examination results and other performance indicators to information on individual students’

ethnic group, special educational needs and family circumstances, and to relate all this information to financial planning and budgetary data.

On the other hand, the focus of most of the work in the school effectiveness and value added fields seems to have been on methodological accuracy, from conceptualising the statistical model appropriately to collecting the right data in the right form. I suspected that the question of what headteachers and staff might do with the analyses if and when they arrived in schools was something either taken for granted or not addressed.

However, headteachers and staff must continue to address the requirements of a changing national curriculum, and to contribute measurably to the achievement of the national targets for education and training. Underpinning this research project, therefore, was the personal belief that the only defensible rationale for introducing so called value added measurements of performance into an already crowded educational agenda is in order to assist with school improvement and the raising of students' levels of attainment – not simply to add to the burden of information.

The aims of the research project were three-fold:

- first, to undertake some further **conceptual clarification and elaboration** in what is meant and understood by 'value added' as it is applied in educational assessments, so as to be able to agree some credible criteria for measuring value added;
- secondly, to set up and evaluate a fairly contained **fieldwork exercise** with which to explore how school staff use value added analyses of their students' performance;
- thirdly, to see how far these case-study findings might be related to, and shed further light on, the substantive issues and problems entailed in **using school effectiveness evidence, in the form of performance data, for school improvement.**

The thesis is accordingly divided into three parts, which discuss, respectively: the research and policy literature in which the various issues and problems broadly related to 'value added' analyses of student performance can be discerned; the findings from a specific fieldwork exercise based on case study schools selected to be at different stages of development in terms of student performance; and finally the implications of these findings for understanding

whether, how, and under what conditions, performance data of a sophisticated statistical kind can be used in the service of school improvement.

To set out the structure of the thesis in rather more detail: Part One (Chapters Two – Five) is essentially a review and discussion of the research and policy literature, made necessary by the contested and changing significance of the idea of ‘value added’ as it pertains to education. I attempt to trace the historical development in relation to key research discoveries and to influential policy events or initiatives: for the measurement of the amount of ‘value added’ by schools to the performance of their students appears to have a very interesting history, bound up with the study of school effectiveness, developments in the statistical measurement of relative performance and changes in educational ideology respectively (see, for example, Saunders, 1997).

But what does the availability of value added measurements contribute in any way to school improvement and raising students’ attainment, and if so how? This would seem a crucial question to answer, given the recent proliferation of value added approaches provided by individual LEAs, by university departments and by commercial agencies, and, of course, the national value added system. The literature search confirms that not much investigation has been done in this field and suggests that there are interesting issues to be pursued about, for example, the different agendas within schools. It also suggests that different kinds of schools will probably need different kinds of strategy – schools are complex social and ‘political’ organisations with different backgrounds and starting points. A key hypothesis of the research was therefore that the notion of a simple formula or checklist which can be applied regardless of circumstance is not tenable.

Part Two (Chapters Six – Nine) moves on to describe and discuss the fieldwork project, whose objectives were to:

- use existing value added performance data analyses, and some additional background data on the schools, to set up a systematic study;
- relate the foregoing argument about the issues involved in value added measurement of performance to what was discovered during fieldwork;
- explore the introduction and use of value added analyses as an integral part of school improvement strategies. The key questions for this aspect of the study were concerned with understanding the context for the introduction and use of value added analyses, and in particular:

- who manages the data, who uses it for what purposes and how is it regarded? is there an identifiable whole-school management ethos which is more compatible with the developmental usage of performance data by staff?
- what are the strategies for embedding evidence-based school improvement approaches at the departmental and classroom levels? how *do* teachers use and respond to value added data? what are the constraints and barriers?
- how are the new expectations and new uncertainties that these approaches bring with them managed?
- to what extent has the use of value added data contributed to strategies for raising students' attainment?
- (if at all possible, given the size and nature of the sample) which strategies for evidence-based school improvement work best in different kinds of school?

Since the NFER was engaged in providing, through its QUASE (Quantitative Analysis for Self-Evaluation) service, value added data on a confidential basis to secondary schools, the fieldwork used nine of those schools to set up an empirical investigation lasting nine months. Chapter Six describes the methodological considerations which informed the design of the project, and gives an account of the research instruments.

The expectation that difference would be as important an analytical principle as similarity was reflected in the selection of case-study schools (discussed in Chapter Six), which takes account of some 'headline' differences between schools; and the findings are discussed from explicitly contrasting perspectives. Chapters Seven, Eight and Nine report the outcomes and attempt different kinds of analysis of the findings. In Chapter Seven, a descriptive analysis is given, in the individual school context, of how value added data was perceived and used; Chapter Eight adduces some additional background information to help understand the apparent anomalies and complexities that emerged from this analysis of findings; and Chapter Nine attempts a more speculative approach by devising a framework based on the notion of senior and middle management 'narratives'.

In Part Three (Chapters Ten and Eleven), an attempt is made to arrive at some conclusions about whether, how and under what conditions value added analyses can help with school improvement. In Chapter Ten, the thesis

examines some detailed examples of how QUASE analyses have been successfully used in practice. This is done, first, in order to indicate the sorts of conditions under which such data can be used with worthwhile outcomes. Secondly, some provisional conclusions can be drawn about how QUASE and, by extension, other systems and services concerned with the use of performance data analyses, can be further developed and improved.

Chapter Eleven summarises and concludes the exploration undertaken in the thesis, which has tried to:

- test the hypothesis that school effectiveness evidence can contribute in non-trivial ways to school improvement;
- identify some of the conditions under which this might be so;
- explore the potential benefits and costs to schools, as social organisations, of using value added data;
- identify what the possible reservations and qualifications are, in terms of remaining conceptual and practical problems.

## *CHAPTER TWO*

### **DEFINING AND MEASURING EDUCATIONAL ‘VALUE ADDED’: SOME FUNCTIONAL AMBIGUITIES**

#### **2.1 Introduction: ‘Choice and Diversity’ in the Definition of Value Added?**

The context and rationale for calculating so-called value-added measurements of performance has changed dramatically since the 1980s, when ‘value added’ was regarded as a quasi-technical idea which had strayed into education from economics and therefore little more than an esoteric preoccupation of certain civil servants. By the late 1990s, the term was being used, apparently quite comfortably by journalists, Government politicians, local policy-makers, school managers and teachers, as well as academic writers, to mean something like ‘a fairer or more accurate measurement of students’ performance and therefore of the quality of their education’. ‘Something like’, indeed: it is part of the purpose of this chapter to tease out the diverse and sometimes conflicting, unclear or inconsequential meanings contained within the range of common usages of the term. Sometimes value added seems to mean whatever the writer/speaker chooses it to mean.

As will become clear, this is not surprising. This chapter gives a historical account of how and – as importantly – why the term value added has come to be used in an educational context, including a brief discursus on the predominant political pressures on education. It outlines the ways in which the term value added was developed, understood and defined in the earlier stages of its educational usage, leading to what are called ‘functional ambiguities’ in the term; that is to say, ambiguities which cannot be altogether eliminated because they are necessary to how the term is made to function. Nonetheless, some criteria emerge in the course of the chapter for making evaluative judgements of what should and should not count as meaningful definitions of value added. The chapter therefore undertakes some important conceptual clarification on which the rest of the thesis is premised.

Because this chapter is more an attempt to trace a historical lineage than to comment critically on the current literature, the discussion centres on items published before the proliferation in the mid-1990s of value added work as we now know it. Furthermore, the literature reviewed for this and the next two chapters has focused on items in bibliographical searches whose titles contain

the words ‘value added’.<sup>1</sup> This is because, although there are several pre-existing and very extensive research fields whose connection with value added is crucial for a full understanding of the term’s import – these include the economics of education, assessment and measurement of performance, the development of performance indicators (including creation of ‘league tables’), the definition and measurement of school effectiveness, the identification of factors affecting performance, and ultimately the question of how performance data can be used for raising students’ levels of attainment – it is only sensible to impose some boundaries on this initial discussion. Relevant literature in these other fields is reviewed and/or used in the discussion in later chapters.

## **2.2 Value Added in an Educational Context**

This section introduces the educational discourse in which the term value added has come to be used in a way which goes beyond its origin as an economic concept. It also, however, notes that the ambiguities associated with the term were inherited from the economic context.

Value added in an educational context, moreover, is one of those terms which comes with an agenda already attached, the agenda in this case being the political preoccupation in the UK (as in other nations) with standards and quality in education, or rather the lack of quality and/or of deterioration in standards therein and the justifiable wish of politicians to get better value for public expenditure. For some time, this agenda has been explicitly attached to the issue of global competitiveness, and the consensus belief that education and training are important levers for economic competitive advantage. At the same time, many practitioners have been concerned to demonstrate that standards (particularly in their own institutions) have not been slipping, that the profession is giving value for money. To put it another way, from the vantage point of the late 1990s value added seems in many ways to have been an idea waiting for its time.

This preoccupation with standards and quality both fed and was sustained by what has been called ‘*an industry*’ (Woodhead, 1997, p.3) of academic research into school effectiveness. From the late 1980s onwards the debate about measures of performance which reveal the ‘*increment on performance of*

---

<sup>1</sup> One indication of the rapid proliferation of interest in value added is the fact that there were twice as many bibliographical references on the topic of value-added in 1995 as in 1990.

*each individual child that goes to [the school]*' – in Smith and Tomlinson's (1989, p.301) phrase – became vociferous and many-faceted. It led on the one hand to highly technical expositions of particular methodological approaches and in this the achievements of the research community were extensive and important. The decade of the 1990s saw an impressive development in identifying, and solving, problems of statistical methodology so as to enable both individual student and school academic achievements to be more rigorously analysed and more acutely interpreted. This work is discussed, in broad terms, in the following chapter.

On the other hand, the complexity of the issues involved in assessing schools' effectiveness, both relatively and absolutely, became inescapable. It is worth remembering that when value-added measurements began to be talked about outside the restricted circle of academic researchers and statisticians, the then Secretary of State for Education (John Patten) was scathing about what he was convinced were ever more sophisticated ways of 'cooking' schools' results: he wanted to stick with 'raw' ones that were simple to compile and understand. By 1995, however, the Secretary of State for Education and Employment (Gillian Shephard: the Departments had been merged in the meantime, thereby confirming the assumed link between education and the economy) who had replaced him had been won round to the extent that in 1995 a research contract was awarded to the Centre for Educational Management (CEM, then based at the University of Newcastle) by the Schools Curriculum and Assessment Authority (SCAA) to design and pilot national systems for value added measures (Fitz-Gibbon, 1997). The policy concerns surrounding value added are addressed in Chapter Four.

Before exploring the measurement and range of possible applications of value added, however, it is necessary to spend a little time unravelling some of its meanings since, as Spours and Hodgson (1996, p.7) point out, '*value-added is used as a short-hand or as an umbrella term to describe a diverse field of practices and projects*'. In acknowledging that the term originates in economics, they attempt an economic definition (as many educational commentators do not), as follows:

The term 'value-added' originates in economics when referring to the difference between inputs, such as raw materials, and energy and final outputs represented by the value of sales... In recent years, the economic concept of 'value-added' has taken on an additional meaning when associated with the need for

advanced Western economies to focus on high quality manufacturing as a way of competing successfully with rapidly growing economies... The term 'value added' in this context is applied to goods whose quality and value is increased by high levels of technology and skill in the manufacturing process. (p.5.)<sup>2</sup>

Implicit in this definition is clearly some principle of transformational activity on the part of human agency, and this should be a signal that there is perhaps more to the economic meanings of the term than meets the eye.

Clearly it is not feasible to undertake a comprehensive review of economic theory for this thesis; but some minor annotation of what seem to be the difficulties in arriving at an economic definition of value added is relevant. One might wish to explore further, for example, the apparently naive question, 'value added to what, by whom and for whose benefit?' This modest task is not without problems of understanding and interpretation, however, because a seemingly inherent characteristic of the term value added is its ambiguity. Is it, for instance, an absolute or a relative concept? And with descriptive or evaluative force? Thus, does 'adding value' simply describe what all companies do in transforming inputs into outputs; or does it refer to what efficient companies do in transforming inputs into outputs with a lower cost-benefit ratio than their competitors?

The answer seems to be, it depends. Mainly it depends, on the one hand, on a prior definition of 'value', which itself depends on the purposes for which it is being defined and, usually in an economic context, (ac)counted. The other critical factor is the level of abstraction within the economic system at which the definition is made. So, whilst it may be well-known that the term 'value-added' has been borrowed from economics, what is not so well-known (at

---

<sup>2</sup> Intriguingly, this discussion by Spours and Hodgson leads one to speculate whether there has not also been some feedback in the opposite direction, that is into the economic context from the educational one. The discourse within which what one might call the evolved definition of value added is situated concerns, once again, the relationship between the (national) economy and education/training provision. It also leads one to wonder to what extent education and training can be demonstrated, or are only presumed, to be integral to national economic health. Whilst this is not the place to discuss that discourse and its own long history, it is worth saying that some doubt has recently been cast on whether the quality of education/training should be taken as the key variable in the success of an economy (see, e.g., Hutton, 1996, quoting Bluestone, 1994, Freeman and Katz, 1994, and Nickell and Bell, 1995).

least amongst educationists) is that its economic definitions and usages are as supple and multiple as the educational ones.

To take an everyday usage, ‘value added tax (VAT)’ is the tax paid on the difference between the buying price and the selling price<sup>3</sup>. This meaning of value added is consonant with the definition given by Spours and Hodgson (quoted above), which enables them to go on to say that:

using the same input/output notion as in economics, value-added has been used to describe the difference between the state of knowledge or qualifications of a student on course entry and her/his state on exit. (p.5.)

Common-parlance usages are actually premised, however, on complex theoretical investigations of economic value, which in turn have been shaped by substantive ideological assumptions. The key question of agency – of who adds value to what and for whose benefit – which is often obscured in classical economic theory, was comprehensively treated, of course, by Marx in his published work from the Paris Manuscripts of 1832 to *Das Kapital* (first published in 1867). In Marxist theory, value added is construed as part of the expropriation of labour-power by capitalists, as in this extract from Mandel (1962, p.305):

Labour-power, as we know, has the dual function of conserving the value of constant capital (the stock of machines, raw materials, buildings) by transferring part of this value to currently produced commodities, and of producing all the new value available to society. The first-mentioned property of labour-power makes it possible to conserve the accumulated stock of social wealth and instruments of labour, which determines the average level of the productivity of labour and the material civilisation of the given society. The second makes it possible to create an income – a ‘value added’ – which in capitalist society is divided between income of labour (wages) and income of capital (surplus-value).

The object of quoting this excerpt is not to enter into a debate on Marxist versus classical economic theory, but to underline the point that, when educationists say that value added is a term borrowed from economics, this

---

<sup>3</sup> It should be noted, of course, that in the case of VAT, the ‘value’ or revenue accrues to central government; more usually, it is shareholders in private companies who are identified as the key beneficiaries from value added.

does not serve well as in any clarification of its meaning. The conclusion to be drawn is that there is no definitive definition, no *ur*-meaning to value added: it is in essence a ‘conflicted locution’. Paradoxically, perhaps, this realisation will be of help as the educational applications are explored. The issue of advantage or benefit is a particularly important one to pick up on: value added, even in an economic context, has an evaluative significance even if it is not actively present in all uses of the term. In an educational context, I would argue, the evaluative *potential* located within the term has been converted into an evaluative *purpose* in all its applications.

It is not easy to find the definitively first use of the term value added in an educational context in the UK. Certainly the early 1990s were, to judge from anecdotal reports and personal experience, a time when many people in educational policy circles were talking with varying degrees of confidence about the need to understand and analyse the ‘value added’ by institutions to their students’ academic performance. By 1990 the term was in sufficiently common use to appear, without warranting elucidation, in the title of a key document published by the Council for National Academic Awards, *The Measurement of Value Added in Higher Education* (McGeevor *et al.*, 1990). An article discussing the work behind this document (Gallagher, 1991) then helped to map the value added territory more explicitly, though conceptually and methodologically the suggested approach was inchoate. Gallagher herself points out (Gallagher, 1991, p.20) that ‘*the term value added appeared in post-school education at least as long ago as 1978 [in Pratt et al.]*’, and this citation is followed up in detail in Section 2.3 below.

By 1992, the National Commission on Education considered the term well enough understood even by people outside the education profession to entitle one of its widely disseminated Briefing Papers *Measuring Added Value in Schools* (McPherson, 1992), which is one of the clearest and most accessible accounts available to a lay audience of the key technical and ethical issues.

It will now be helpful to consider the provenance of the term value added in education and the route by which it came to be used as a term connected with the measurement of quality.

### 2.3 Why ‘Value Added’? Pressures, Precursors and Presumptions

This section explores some of the history of the term value added as applied to the economics of education and traces two distinct strands in its development.

The previous section noted specific dates by which the term value added was evidently in common usage in educational circles. But this is not to imply that the ideas which lay behind the introduction of the notion of value added to education were not already very well established by then. In fact, the development of value added analyses of educational performance has two rather different albeit intertwined strands, which – in addition to the existing difficulties in the economic derivation outlined in the previous section – may well account for some of the continuing confusion in its definition and application.

As an explicit *term*, its first appearance seems to have been in the higher education (HE) sector in the context of the development of institutional performance indicators. As a critical political and educational *issue*, however, it originated from the school effectiveness research agenda which was particularly at home in compulsory secondary education. In both cases, there has been an associated methodological inquiry which has become a dimension of educational research in its own right.

What incontrovertibly gave both sectors a rationale – some would say an outright requirement – for being concerned with value added in education were the centrally imposed funding-and-accountability mechanisms introduced into the education system in the late 1980s. It is interesting to note that a key document in the justification of how these mechanisms should and could work was compiled by a firm of accountants (Coopers and Lybrand, 1988), an event which was surely a message about the way things were to be from now on<sup>4</sup>. The subsequent legislation included the Education Reform Act of 1988 affecting primary and secondary schools in England and Wales, and the various statutory changes affecting the funding, control and status of institutions of further and higher education (F/HE) in those countries.

---

<sup>4</sup> On a parochial note, it is evident that over the last decade sponsored evaluation projects in the field of educational management which used invariably to be granted to university departments or independent research organisations have often been awarded instead to accountancy and management consultant firms, who are now major players in educational research of one kind.

A search of the literature of that time makes it possible to discern how exponents in the different sectors developed rather different kinds of argument about value added measurement. Essentially, the HE sector seems to have been particularly concerned with assessing efficiency and effectiveness *within* educational institutions in order to justify funding allocations; whilst the school sector was more pre-occupied with ways of making fairer comparisons *between* educational institutions, partly in order to explain the apparent disparity between dissimilar results for similar funding levels in different institutions or local education authorities. And again this distinction of emphasis should be no surprise, given the wider educational policy context. At the same time, one should not make too much of the difference, since what the two conceptualisations have in common is the underpinning question: ‘How can pupil/student progress be measured in such a way as to throw light on the performance of institutions?’ This question is probably the key to understanding the methodological principles of value added, and will be returned to in the following chapter.

However, it suits the purpose of this chapter to pursue the historical perspective a little further and temporarily to treat the two strands as separate developments. So far as the HE sector is concerned, the response from the Council for National Academic Awards (CNAA) and the Polytechnics and Colleges Funding Council (PCFC) (McGeevor *et al.*, 1990) to the 1988 Government directive on efficiency in higher education, gives some strong clues. This document defined value added as:

a measure of student achievement which takes into account the effects of differential student inputs (i.e. entry qualifications). Student achievements thus conceived can be used to contribute to the evaluation of institutional quality in the teaching and learning functions. (McGeevor *et al.*, p.1.)

Work in the further, as distinct from higher, education sector on value added proceeded along similar lines (see, for example, FEU 1993 for a definition of value added in FE very similar to that for HE); and most of Spours’ and colleagues’ influential body of work has been in the 14–19 sector (see, for example, Spours and Young, 1994; Post-16 Education Centre, 1996). This essentially economic notion of inputs (and outputs) can be traced back to the argument developed over a decade earlier in Pratt *et al.* (1978). The authors of this thought-provoking book undertook some ground-clearing discussion in Chapter 3, ‘Economic issues: the economics of education’, when they asked:

*'Input and output – can there be efficiency in education?'* (p.154.) Their view was that:

The concepts of input and output are central to economics. They underpin notions of efficiency and productivity and calculations of *value added* and of costs and benefits. Yet they seem to have caused economists little but trouble in their application to education. There are indeed some economists who, when faced with an education institution or service, cannot tell the difference between inputs and outputs. (p.154.) [My emphasis.]

In a critique of Government cost studies, the authors expand on the question of what should count as inputs and outputs in an educational context. In budgetary terms, of course, a major input measure must be that of financial cost; in a complex service such as education, the main issue might well be seen – and was so seen by the then Department of Education and Science (DES) – as one of how to define costs. But the authors argue (p.156) that:

It is not a matter of a satisfactory definition of costs. It is a matter of deciding what costs to calculate, and the decision will be different for different purposes. *The question of what to count as input is not a general or definitional question*: it is a question that has meaning only when related to both the output and the process involved. It will differ, for the same institution or service, in different circumstances. [My emphasis.]

The sentence which has been highlighted is arguably a crucial piece of conceptual clarification which serves to underline a similar point made in the previous section. It is an argument to which it will be necessary to return.

The question of how educational 'outputs' were to be defined was apparently even more tricky (Pratt *et al.*, pp.156-8). It is salutary to be reminded that in 1978 it apparently could not be assumed that the then DES, still less the University Grants Committee (UGC), would think of outputs in terms of student performance or, more globally, educational standards. The authors characterise official thinking on outputs as '*unsatisfactory*' and '*muddle[d]*' (p.157), in that factors such as '*student/teacher ratios*' and '*quality of teaching*' as well as '*graduates*' themselves, are instanced as examples of output or product. They go on to argue firmly for a '*simple but serious*' application of economic concepts to education; and therefore for a view of

inputs which would comprise all resources, including staff, and of outputs which would comprise ‘*educated students*’. But, they say:

This may seem a simple and obvious conclusion; in fact it overturns all current practice in costing, budgeting and accounting; both administrative and academic... In education we ... need to be able to discuss what mix of inputs is most apt for producing desired outputs... (p.159.)

These extracts should serve to indicate how the discussion of efficiency and quality in HE was closely tied into a concern with funding, and therefore how ‘educational value added’ started out as a truly economic pre-occupation, at least in this sector. The chapter goes on to propose an explicit definition of value added:

The question of quality will then relate to the difference which the educational institution or service has made to the entrants... The course’s contribution, the value added by it, is precisely the difference it makes to the student. (pp.160-3.)

What follows in the remainder of their chapter is a discussion of the problems such a definition raises, such as the inappropriateness both of economic measures like graduate salaries and of existing examination and assessment systems to measure the difference courses make, combined with the tendency of educators to make unsupportable claims about the contribution of their courses to students’ maturity and so forth. Pratt *et al.* conclude:

We have discussed various economic ways of measuring the [difference which the college makes to the uneducated student] and have concluded that these are not apt... It is for the educators to create measures which, apt in principle, are also convincing in practice.’ [p.166.]

It might be argued that this is the challenge which the value added ‘project’ – in so far as the various individual academic studies, research and development reports and government policy statements add up to such a thing – has in principle accepted; except for the fact that the arguments and conclusions presented in this book seem to have sunk almost without trace. What has happened in practice is that educators have defaulted to using proxy variables which already exist and are measurable, for understandably pragmatic reasons – but this is to pre-empt the discussion in the following chapter of this thesis.

So far as it is possible to check, Pratt *et al.*'s work was last cited, in the context of discussing educational value added, in Gallagher (1991, p.20), although '*the advice and assistance*' of one of the authors was acknowledged in the CNA/PCFC document (McGeevor *et al.*, 1990, p.51) cited above. It is not proposed to undertake a critique here of the computational bases for value added presented in the latter document, important though they undoubtedly have been for the work of university and college managers. What is more at issue for this thesis is the way in which Pratt *et al.*'s clear though challenging definition of value added was rendered imprecise by a series of assumptions and elisions made in subsequent documents such as the CNA/PCFC report, which thus allowed a basic confusion to be perpetrated and, as it has turned out, perpetuated.

First of all, as the excerpt quoted earlier in this section reveals, the authors of the CNA/PCFC report seem to have pre-empted the discussion of what should constitute 'input' and 'output', since they implicitly equate 'output' with '*a measure of student achievement*' and explicitly define 'input' as, and thereby restrict its meaning to, '*entry qualifications*'. The imputed relationship of input to output is rather vaguely glossed as a matter of '*taking into account*' (p.1). What the measurement of value added therefore comes down to is something '*based on a comparison of entry and exit qualifications*' (p.1), which is not at all the same thing as a computation of how efficiently inputs in the normally understood meaning of the word are deployed to achieve desired outputs.

Further on in the same document, the following note is appended to a table purporting to show '*value added scores for male/female sociology graduates*':

all [methods of calculating gender differences in value added]  
indicate that women *obtain greater value added* than men;  
(p.23.) [My emphasis.]

A little syntactical deconstruction is called for. It will be remarked that this phrasing contains an ambiguity about the agency to whom the capacity to add value is being attributed, in that 'obtain' could have an active or a passive meaning (being equivalent either to 'women achieve greater value added than men' or to 'the results for women show a higher level of value added than do those for men' respectively). Now this vagueness of verbal expression might be regarded as a trivial matter – pointing it out at all might be dismissed as

sophistic nit-picking; but its implications are decidedly non-trivial. For the ambiguity amounts to a conceptual equivocation (not necessarily intended by the authors, it must be said) about what sort of activity, in what category, ‘adding value’ is. To make this as plain as possible, I am arguing that the ambiguity, when elided with the notion that measuring value added is something to do with ‘a comparison of entry and exit qualifications’, results in a blurring of boundaries between two logically and politically disparate concepts: namely, on the one hand, the progress made by an individual between, say, A-levels and first degree and, on the other, the amount – if any – of difference to that individual’s progress which can be attributed to what has been done by the institution s/he attended as distinct from any of the other factors at work.

As a rider to this discussion, it may be noted that the objectives of the CNA/PCFC project were to test different approaches to the measurement of value added:

‘... *relevant to the operational needs of the PCFC and higher education institutions*’ (p.1) [My emphasis.]

It may be speculated that those needs had more leverage on the project than a strict attention to what could be construed as a question of mere semantics.

The distinction made above, between individuals and institutions, is material to the application of, and the methodology associated with assessing, value added in education. It is an argument of the present thesis that the former concept (progress made by individuals), whilst integral to both the discussion and the measurement of value added, does not constitute value added in itself. Adding value, certainly according to Pratt *et al.*, is what establishments (colleges, they would say, as well as companies) do, by virtue of being economic micro-systems with resources at their disposal and decision-making structures for disposing in a variety of alternative ways of those resources; adding value is, on a broad economic definition, something logically and practically different from, something over and above, what individuals can do for themselves by way of making progress.

What is striking is that this conceptual ambiguity has continued right into the most recent literature, where it is the case not only that different authors mean different things by the same term but also that the same authors use the term to

mean different things at different times (see, for example, Spours and Hodgson, 1996, *passim*; Fitz-Gibbon, 1997, Glossary, under *Regression* and *Reliability* as well as *Value Added*, pp.122, 124 and 127 respectively). As was said earlier, it is quite hard sometimes to spot such instances or to get to grips with the issue at all, however, precisely because it is concerned at least as much with things which are *not* said or clarified as with things which are.

Turning now to the Gallagher article (1991), which, as was earlier noted, is a useful critical commentary on the CNA/PCFC document and the work which formed its basis, we can see that some of these problematic issues were adumbrated therein. Gallagher defines value added as [a performance indicator which] ‘is intended to indicate the contribution of an institution to a student’s achievement, to determine what an institution adds to a student as she/he passes through it.’ (p.19.) She glosses this as follows: ‘although addressing largely qualitative issues, if value added is to be adopted as a performance indicator it requires measurement in quantitative terms.’ Each of these points, one could say, constitutes a key principle to bear in mind when considering value added, namely:

- the need for *standardisation* of measurements;
- the notion that institutions ‘*make a difference*’ to student performance which can be assessed;
- the acknowledgement that what is being looked at is a matter of qualitative judgement – or, as one might prefer to spell out, involves *making value judgements* (whether or not these have been made explicit and subject to scrutiny).

But a major finding to be highlighted at this point is that educational value added in HE is usually coterminous with ‘progress made’ or ‘distance travelled’ (two commonly-occurring phrases in the literature); and whilst the idea of ‘distance travelled’ clearly entails a starting and an end point, it does not require that there must be a known correlation between the two; still less that beginnings and ends must be defined in precisely the same way in different institutions. Spours and Hodgson’s (1996, p.5) general definition – repeated here for ease of reference – succinctly sums up this view:

Using the same input/output notion as in economics, value-added has been used to describe the difference between the state of knowledge or qualifications of a student on course entry and her/his state on exit. By taking the starting point of the learner into account, it is possible to consider the extent to which both a course and an institution have been responsible for any progress made.

Such a definition, whilst not actually incorrect, conceals several conceptual and technical difficulties. Gallagher (1991, p.22.) was moving towards a more rigorous view of measurement when she said that *'some [additional] notion of relativity is required if judgements are to be made about the institution and comparisons made between institutions.'* In other words, what Gallagher calls *comparative* value added needs to be adduced, which will create 'a level playing field' and one which is based, not on an arbitrary score, but on *'an empirically derived expected value'* (Gallagher, 1991, p.22, quoting McGeevor *et al.*, 1990). As Gallagher points out (p.22): *'using this approach, it is no longer up to institutions to decide the value of entry and exit qualifications.'*

The work of Fitz-Gibbon and Tymms, then at the Curriculum, Evaluation and Management Centre of the University of Newcastle upon Tyne, in the 16–19 sector is not mentioned by Gallagher, presumably because there were felt to be too many differences between further and higher educational qualifications and institutions for the principles and methods to be applicable. Yet Fitz-Gibbon's development, with Tymms, of ALIS (A-level Information System) was a pioneering effort in using regression analyses to measure and report on relative progress, in which students' O-level (to be superseded by GCSE) results in participating schools and colleges were taken as the starting point for assessing their A-level performance. (See, for example, Fitz-Gibbon, 1991a and b, 1992; Tymms, 1990.) Less progress has been made in establishing national systems for measuring value added in vocational programmes (see, for example, FEDA, 1995; Hewitt and Hutchison, 1995).

Fitz-Gibbon and Tymms applied their ALIS model to the development of systems for measuring value added in the secondary and primary sectors (with YELLIS and PIPS respectively), and also won the SCAA contract for exploring the feasibility of a national value added system (discussed in Chapter Four). This is therefore the point at which the discussion of value added in post-compulsory education begins to exhibit similarities to the applications developed in the compulsory sector. So let me now turn to the

second strand in the development of educational value added: how value added has come to be understood in secondary as distinct from tertiary education and in particular what it owes to the school effectiveness debate.

If, without further argument for the moment, one accepts that there *is* a different meaning of value added in the compulsory sector and therefore takes the value added question to be pretty much as Sweetman (1997, p.77) states it: ‘*at its simplest, how much better do students in a good school perform than they might have been expected to do from their performances on their first day?*’, then it is possible to track this back through the accumulating research into school effectiveness, which already by the early 1990s had a long and strong pedigree. This research is summarised and discussed, so far as it is relevant to the present thesis, in the chapter which follows, but it is perhaps worth risking some over-simplification here to say that at the core of the research was the demonstration that some schools are measurably more effective than others (for example, Rutter *et al.*, 1979; Smith and Tomlinson, 1989). Rutter *et al.*’s research arguably contributed to the shift of focus of concern away from national monitoring of aggregate statistics towards a scrutiny of the performance of individual schools. The conclusions reached by the end of the 1980s in this area could be represented thus:

Research and analysis showed that schooling is not effective as a means of reducing individual inequality. It did not show that schooling has little effect on whether or not children can read, write and do arithmetic... The result of going to an effective school can be seen as an increment on the performance of each individual child that goes to it. This increment may be large enough to be very important for its effect on what each individual is actually capable of, yet small in comparison with the differences between individuals.’ (Smith and Tomlinson, 1989, p.301).

It was inevitable that this new emphasis on performance (as distinct, say, from other legitimate concerns like curricular process or learning experience or educational equity) – which, ironically from some points of view, sat rather comfortably with the New Right’s determination to identify and ‘weed out’ poorly performing schools and teachers – should have concentrated attention on how best to *measure* performance or ‘*achievement outcomes*’ in Mortimore and Stone’s phrase (1990, p.73), in ways which would show how far schools had actually contributed to the achievements of their students. (Hardly surprisingly, then, the introduction of value-added analyses has in turn – as is

argued later in the thesis – had an impact on the direction and scope of school effectiveness research.)

The evidence from, and methods of, this body of research were able to be pressed into the service of making fairer comparisons between institutions, which has continued to be a huge pre-occupation amongst head teachers and others concerned with the effect on schools of ‘league tables’, the public reporting of schools’ academic performance in rank order of results obtained in national curriculum tests and public examinations, introduced by Government in 1992. (Schools were already enjoined by the 1980 Education Act to publish public examinations results in their school prospectuses.) Secondary schools and their managers were most visibly and immediately affected by league tables, because of the presumed role of such tables in informing parental choice of school (extending parental choice had been another avowed intention of the Education Reform Act), and also because GCSE examination results for students at the end of Key Stage 4 of the National Curriculum – unlike National Curriculum tests for students at the end of Key Stages 1, 2 or 3 at that point in time – were already a standard means of assessment and accreditation which could readily be deployed for drawing up school-based lists in the public domain. One could go so far as to say that making comparisons between schools was – up to and including the government’s announcement in 1997 of the development of a national value-added system – the major purpose of measuring, and *a fortiori* of defining, value added in the compulsory sector.

Technical and conceptual complexities of the kind alluded to earlier in this section were therefore squarely in the frame from early on. We may note, for example, the publication of ‘The search for a fairer way of comparing schools’ examination results’ (Gray *et al.*, 1986) in the mid-1980s. In posing the question: ‘*under what circumstances might one be justified in claiming that [school A] has done ‘better’ than [school B]?’* (p.91), the authors could be said to have encapsulated the aim of (comparative) value added analysis *avant la lettre*. The findings from this study – for example, that ‘*differences between schools did not appear to be a very important source of differences in the results achieved by individual students of differing ability/attainment levels attending different schools*’ (p.118); and that there was ‘*competing evidence on the question of whether individual schools were differentially effective with students whose abilities/attainments differed at intake*’ (p.118 again) – also

pre-figured a very great deal of the evidence on relative measurable effectiveness which has since been accumulated.

As the authors argue, methodologically speaking two things follow from the intention to compare schools' performance: first, some concept of 'contextualisation' must be introduced, to allow for the fact that schools are dealing with student populations which differ considerably in terms of prior achievement and other key variables (the issue of 'input'). Secondly, since schools are '*part of an educational system which has natural hierarchies or levels of "nesting" [e.g. students within classes within year groups within schools]*' (p.117), the statistical model adopted to analyse performance data should reflect this structure. These arguments are taken up in the following chapter.

## **2.4 Conclusions**

The evidence presented so far indicates that, although by the late 1990s it seems fair to say that the combination of rapidly developing school effectiveness studies together with a government-led focus on the value-for-money of further/higher education has fundamentally changed the attitude of most educational managers towards sophisticated quantitative data and its relevance to the everyday life of schools and colleges, the acceptance of the term value added in an educational context should not be taken to mean that everyone understands the same thing by it (nor, indeed, understands it *tout court*).

Moreover, the evidence suggests that no final definition could ever be arrived at, since value added is a term which contains within itself an evaluative function: its meaning and therefore its definition (as well as its application and associated methods of measurement) will depend on the purpose for which it is to be deployed in any given context.

It is no accident that education has borrowed the term from economics, since there is a consensus in most countries of the world that education makes a direct contribution to national economic performance and is thus a sub-system of the economic system; however, no convincing model of how that contribution can truly be measured has yet been developed, and so educators

and economists alike, and economists of education, have continued to use proxy variables such as qualifications, often without question.

Even so, it seems to be possible to say that the ideas contained within the usages of 'value added' in an educational context contain the following items (variously combined and/or emphasised) of assumption and evidence:

- educational performance – as a measure of standards and quality – matters not only to the individual but to the health of the national economy;
- measures of standards and quality, and of the contribution of education to the economy, are necessary; it is not enough to rely on the assertions of educationists about the efficacy of what they provide;
- value added has come to be a meaningful term in an educational context and can be deployed in the assessment of standards and quality;
- institutions make a difference to students' educational performance;
- it is possible, and from many points of view desirable, to make comparative assessments of institutional performance; in consequence, some institutions can be shown to be measurably more effective than others in terms of specified student outcomes;
- it is possible to identify and account for key factors beyond an institution's immediate control which are implicated in its students' performance;
- it is possible to calculate 'predicted' or expected levels of performance based on these factors, and then to assess individual institutions' results against these (the model presupposes that positive school/college-level 'residuals' amount to their value added – this will be further explored in the following chapter).

It is now appropriate to scrutinise the academic literature more thoroughly, for a further exploration of the ambiguities identified in this chapter, for a discussion of the methodological developments in measuring value added and for an understanding of how far value added – as a term which has been adequately explained and understood – has a useful contribution to make to the understanding of educational quality.

### **CHAPTER THREE**

## **ANGELS ON A PINHEAD? REVIEW OF THE ACADEMIC LITERATURE ON EDUCATIONAL VALUE ADDED**

### **3.1 Introduction: Clarifying the Line of Inquiry**

This chapter pursues the task of conceptual clarification by reviewing key texts in the academic literature on value added in the educational sphere since the mid-1980s, with the aim, first, of exploring further the argument developed in the previous chapter about the principles of, and ways of doing, value added analysis; and secondly of identifying the strengths and weaknesses of the approach in the context of assessing 'school effectiveness'. In doing so, it reveals how the later usages of value added need to be 'read back' into previous research pre-occupations. The chapter demonstrates how value added has shifted in function from being a contested methodology concerned with making better (fairer and more valid) comparisons between schools and 'owned' by researchers on behalf, so to say, of educational stakeholders generally, to being a management tool concerned with raising pupils' attainment and internal school improvement, owned by national government, LEA advisers and school senior managers on behalf of an education-management culture. (That this latter development has been extant for some years in at least some states in the USA is demonstrated in Webster *et al.*, 1994.)

The line of inquiry it seems desirable to pursue can be posed as a series of questions, thus:

- what meaning(s) of value added seem to have been pressed into the service of educational effectiveness discourse?
- what kinds of analytical models and methodological strategies have been developed as a consequence?
- in turn, what impact have these models and strategies had on the way effectiveness has been understood in practice?
- what have value added analyses revealed about 'effectiveness' and educational quality? Are these findings valid? Are they interesting and important? To whom?
- what further questions and queries have value added analyses raised? In particular, what do value added analyses leave out of account and what serious weaknesses, if any, do their models contain?

The chapter discusses these questions in some detail. Section 3.2 investigates the provenance of value added principles in the context of research in the compulsory education sector; Section 3.3 comments, in a historically organised narrative, on the relationship between school effectiveness literature and the notion of value added. Sections 3.4 and 3.5 each offer a brief discursus on value added methodology and on the contribution of value added to school effectiveness evidence respectively. Section 3.6 concludes the chapter by recording the limitations of value added approaches and then revisiting the proposition made at the beginning of the chapter, that value added has shifted its functional meaning over time.

### **3.2 The Provenance of Value Added Principles in Educational Research Relating to the Compulsory Sector**

It was claimed in the previous chapter that the calculation of ‘value added’ measures in the school effectiveness context was initially and primarily concerned with making fairer and more valid comparisons between schools in terms of their pupils’ academic performance, the impetus for which came from the Government’s decision to publish comparative tables of school performance. If the relevant literature is examined more closely, ample evidence can be found that this is indeed so, although – to understand the issues properly – the distinct issues integral to that general theme must be identified. These may be thought of as follows:

- the measurement of educational quality and the development of educational performance indicators (see Mortimore and Stone, 1990);
- the problems associated with attempting to compare schools’ effectiveness on the basis of published examination/test results (see Cuttance and Goldstein, 1988);
- the interpretation of schools’ test/examination results, and especially of the differences between schools in this respect (see Gray *et al.*, 1990);
- ‘disentangling the variety of influences on a pupil’s examination performance’ (Education, 1992).

It is plausible to construe the Education Reform Act of 1988, and in particular the recommendations of the Task Group on Assessment and Testing (TGAT), as providing both a focus and a momentum – ‘the moment’ – for research which up to that point had comprised innovative but isolated projects on what was characterised as ‘the school effect’. Much of this work had been overtly concerned with an equal opportunities agenda, with specific reference to selective versus non-selective schools (for example, Gray *et al.*, 1984; Maughan and Rutter, 1987) or to the

achievements of minority ethnic group pupils and/or pupils in inner city schools (for example, Maughan and Rutter, 1986; Smith and Tomlinson, 1989). Some work was also already being done on developing systems for comparative performance analysis (see Gray, 1979, 1981a and b; Gray *et al.*, 1986).

It was due to the strength of the conceptual and methodological frameworks thus established – supported by hierarchical modelling techniques (see Raudenbusch and Bryk, 1986; Goldstein, 1987) – that the implications of the Government’s radical proposals in 1988 for a national system of assessment and testing, and its intention to publish the results school-by-school to assist the extension of parental choice of school, could be challenged so quickly and cogently by academics.

Thus before the end of that same year a short paper by Cuttance and Goldstein (1988) was published in which the authors explicitly forewarned of, and tried to forestall, what they called the ‘handicapping system’ (p.200) which would in all likelihood do an injustice both to schools as would-be providers and to parents as would-be selectors of educational quality. The authors argue that if the proposed system were to make comparisons between schools on the basis of pupils’ attainment in circumstances where there was variation, either social and/or academic, in pupil intakes to schools – and it could be assumed that in the real world there *would* normally be such variation – this would mask the true extent of the progress made by pupils in different schools, as distinct from the standards reached by them.

The basis for their argument was the ‘*vast research literature*’ demonstrating that ‘*children from socially disadvantaged areas tend to have lower exam scores and test results than those from more socially advantaged areas*’ (p.197) and, even more importantly, that ‘*the attainment of... children when they first enter the school is the single most important determinant of subsequent achievement*’ (also p.197). If these factors were not taken into account, the performance of schools could be neither fairly assessed nor validly compared, as TGAT themselves intimated. But, according to Cuttance and Goldstein, TGAT could be accused of wanting it both ways: on the one hand, failing actually to mention the school effectiveness research so far published whilst on the other – by recommending that local socio-economic (SE) factors should be published ‘alongside’ schools’ results – passing the buck to parents and LEAs for making the necessary adjustments to schools’ results in order to get a truer picture of their effectiveness. The authors conclude (p.201) that:

the system of school evaluation proposed in the Report [of TGAT: GB.DES/WO, 1988] would be quite misleading, and school boards, parents and teachers would be right to protest loudly at the use of such poorly constructed performance indicators.

Smith and Tomlinson (1989, p.303) who must have been writing their conclusions in the same year, protest in a similar vein as follows:

[Our findings] show that a comparison of the raw test results between schools would be highly misleading... If the results of the tests are to be made publicly available, it is essential that they should be analysed by methods akin to those used in the present study [i.e. using multivariate analysis]...

They too go on to say:

There is no support in the findings of this study for the idea that increasing parental choice will improve standards... it is quite clear that parents cannot identify the schools that are doing well in terms of pupil progress. This is hardly surprising, since it takes a complex analysis to identify those schools.

There are other key points made by these two pairs of authors which must be noted in an historico-narrative account such as this. Presciently, Cuttance and Goldstein introduce their paper by warning that 'publication [of comparative school results] could well become the most visible and perhaps most important aspect of the whole system, irrespective of the actual content of the assessment or indeed of the National Curriculum itself.' (p.197.)

On more substantive matters, Cuttance and Goldstein point out that there are flaws in the system at a basic level, in that a comparison of school averages:

tells us nothing about the relative achievements of different types of pupils within the schools... Consequently schools which perform well relative to other schools for the average pupil in the population may perform less well for disadvantaged or advantaged pupils. (p.198) [My emphasis.]

This issue – of differential effectiveness – if found to be generally applicable would obviously have consequences for both the public reporting of school comparisons and for parents' choice of school for their individual children. Concern about the urgent need for '*Fair and comprehensible ways of presenting performance indicators in context, and to reveal differences between sub-groups of students*' fuelled another influential paper of around the same time (Nuttall *et al.*, 1989, p.769). Based on the findings from a multilevel analysis of a large dataset held about secondary schools in inner London, it concluded that:

...school effectiveness varies in terms of the relative performance of different sub-groups. To attempt to summarise school differences, even after adjusting for intake, sex and ethnic background of the students and fixed characteristics of the schools, in a single quantity is misleading... [T]he concept of overall effectiveness is not useful.' (pp.775-6.)

On a second crucial issue, Cuttance and Goldstein point out (p.198) that ranking schools in order of results – whether raw or adjusted for social background and/or prior attainment – is in any case highly problematic because the method used for arriving at estimates of school performance partly determines the ensuing rank order. We may assume that 'method' in this instance covers both the issue of what outcomes are used – particularly those which can differentiate between different groups of pupils (see above) – and the question of what statistical modelling techniques are employed. These matters are discussed in more detail in the section following.

Thirdly, these authors say that '*the more background factors which are taken into account, the more unstable and unreliable become the resulting comparisons*'. This is an important principle to keep in mind, because there have been calls – understandably from the point of view of those seeking to manage schools and improve standards – for a model which would include everything which might be thought of as a factor affecting schools' performance. But actually the underlying principle of most value added work is an eliminative rather than an accumulative one, relatively simple in conceptual terms though complex statistically. Its relationship with school effectiveness is therefore indirect: its main role has been

not so much to identify a range of effectiveness correlates but to discern different kinds of ‘noise’ or extraneous information in the analyses of effectiveness, and to get rid of it as far as possible.

Finally, a point made by both Smith and Tomlinson and Cuttance and Goldstein (and which can be easily inferred from Nuttall *et al.*) is one that follows inexorably from the ability to make fairer and more valid comparisons between schools on the basis of pupils’ progress, but which tends to be obscured by the term ‘value added’. Smith and Tomlinson put it like this:

A school having a low balance of intake could appear to be doing badly, when in fact it was doing well; while a school with a high intake balance could be flattered by the raw test results (p.303).

This relates to another principle which it is important to bear in mind: that, since value added analyses are norm-referenced in the way they are calculated, there will inevitably be ‘losers’ as well as ‘winners’. This implication is not always understood in the cut-and-thrust of getting schools’ public relations messages across: headteachers have been heard to say things like: ‘*we just want to be able to prove what we already know, that all our teachers are adding value*’.

Nuttall *et al.* mention in passing another ‘major’ issue for investigators of school effectiveness, that of the stability of school effects over time (p.770). This has since become a key area of research and analysis.

Two rather later articles are also worth citing in this context, Paterson (1991) and Schagen (1991). Paterson’s detailed paper (cited by McPherson, 1992, below), whose findings are based on a multilevel model of school effects and their association with both socio-economic status (SES) and prior attainment of pupils, concludes:

Any system of monitoring schools requires attention to be paid to SES if it is to inform the practice of education. Equally, though, the measurement of SES has to incorporate studies of the detailed ways in which the social and economic circumstances of school pupils influence their opportunities and incentives to learn. Simply knowing that SES is associated with attainment – even in the relatively subtle ways which we have used here – is merely a first step towards understanding why it is so, and towards developing practicable ways of overcoming its harmful effects. (Paterson, 1991, p.117.)

In other words, the fact that socio-economic status is empirically associated with attainment does not mean we are entitled to infer an *a priori* association, still less a causative one. Moreover, Paterson suggests that there is perhaps an even more fundamental conceptual issue at stake:

... one way of viewing this problem of membership [of neighbourhood, occupation and/or labour market] would be as an opportunity to test various theories about the ways in which individuals can be said to belong to groups... [But] it is likely that some of the multi-dimensionality has appeared because a cross-section of pupils captures parents who are at different points in different types of occupational or family careers. The multi-level nature of social processes has the consequence that time itself is multi-level. (p.117.)

This hint that the interplay of educational and social factors is really a complex narrative being played out through time at the individual, family, school, local and national levels simultaneously is a very useful and important reminder of the challenges posed by the social sciences to a reductionist ‘common sense’ view of issues like school performance and effectiveness.

Schagen’s (1991) article by contrast is brief and practical. The reason for including it here is that it aims, by giving clear worked examples, to ‘show that, given suitable data and the use of appropriate statistical techniques, it is possible to use numerical outcomes (e.g. test results) to obtain, valid and helpful results about schools and their effects on children.’ (p.216.) Premised on the claim that ‘league tables of schools based on a single output measure are statistically dubious’ (p.222), it discusses the operation, and the strengths and weaknesses, of both multilevel modelling and data envelopment analysis.

A key point to be made is that none of the five sources<sup>1</sup> discussed above once uses the term value added. And yet everything in the development of value added analyses since then confirms the truth of their painstakingly won insights. In unravelling the history of value added, one is therefore more or less impelled to ‘read back’ into their research concerns and strategies later understandings and usages of value added.

### **3.3 Changes Over Time in the Relationship Between ‘Value Added’ and School Effectiveness Discourse**

It seems to be in the post-ERA research literature from 1991 onwards that we need to search for evidence of how ‘value added’ has become part of the discourse on educational effectiveness, and on what terms. It was clearly necessary to set criteria for a search of such literature; of the many texts authored by researchers and others, only those which, first, specifically allude to ‘value added’ approaches, data and/or analyses in their titles or abstracts and, secondly, provide one or more of the following, have been selected for discussion in this sub-section:

- first-time published use, or clarification, of key concepts and assumptions;
- important differences from the consensus or conventional view;
- description of techniques for measuring value added;
- critical overview of value added approaches in the context of the broader research issues involved.

The key texts selected on these criteria are, in chronological order of publication<sup>2</sup>, Nuttall, 1991; Kennedy, 1991; Jesson, 1992; McPherson, 1992; Nuttall, 1992; Wiliam, 1992; Birnbaum, 1993; Fitz-Gibbon and Tymms, 1993; Cornall and Lofthouse, 1993; Gray, 1994a; Mortimore *et al.*, 1994; Schagen, 1994; Fitz-Gibbon, 1995; Goldstein and Thomas, 1995; Gray, 1996; Jesson, 1996; Thomas and Mortimore, 1996; Thomas *et al.*, 1997. Some other articles and papers which use the term ‘value added’ (and are often interesting enough in their own right) amount, for current purposes, to glosses on earlier or more detailed texts and so I have chosen not to discuss them at this point, although references are sometimes made in passing. Yet further articles are important later on in the

---

<sup>5</sup> I am very grateful to Professor John Gray for subsequently drawing his earlier papers (1979, 1981a and b) to my attention in this historical context.

<sup>6</sup> The order of publication may not always, of course, reflect the order of composition.

discussion of the application of value added findings to school improvement, and are dealt with in the appropriate sections and chapters.

Nuttall's (1991) article in the Times Educational Supplement is a plea, couched in lay-person's language but attempting to summarise the research intelligence so far, that educationists should quickly demonstrate the feasibility of establishing systems for analysing student performance in terms of their relative progress, as opposed to accepting the principle of raw league tables. He says:

For nearly 30 years, research on school effectiveness has used the *progress* made by students from their level of performance on entry to their level of performance at the time they leave, rather than just their 'raw' results at the time of leaving. This approach, well illuminated by the metaphor 'value added', makes intuitive sense and is readily comprehensible.

As well as this insistence on the need for measures of progress, the article makes several other points about the requirements of the value added 'principle' or 'model', viz:

- the need for a range of outcome measures (not merely five or more A-C grades at GCSE, for example);
- the need to allow for socio-economic variables;
- the need for analyses of differential effectiveness;
- the need to allow for the possibility that schools' results are not stable over time;
- the need to use multi-level modelling as the statistical technique.

One may take it that these principles are based on the analysis, and its implications, of the 1990 GCSE examination results undertaken by Nuttall *et al.* for the Association of Metropolitan Authorities, although not published until 1992. Kennedy's (1991) article for *Managing Schools Today* is less clear about the prerequisites for calculating value added; but this may not be altogether surprising given that he was writing from the perspective of the Audit Commission which at that time was engaged in exploring approaches to value added in the 16-19 sector – a sector that has remained to this day intractable in terms of establishing appropriate baselines and inputs for calculating value added (apart from those for GCE A-level examinations). As we saw in the previous chapter, the definition of and methodology for assessing value added in F/HE tended to be connected with a less problematic notion of 'individual progress made' than that which was emerging from the longer and stronger tradition of school effectiveness research. Thus

Kennedy, in taking issue with league tables of raw results, claims that *‘a sounder basis for comparing institutions is the progress made by students, which will reflect the “value added” by their education.’* It is evident from what he goes on to say that ‘progress made’ is equivalent to the ‘distance travelled’ by different students in different institutions between, say, their GCSE grades and their A-level grades. He notes, however, that *‘no generally accepted robust methods for measuring the value added have yet been established’* and that *‘a number of key technical questions’* would need to be resolved before the extent of the likely differences between the performance of individual schools could be gauged. He makes no mention of the well-known work on GCSE to A-level comparisons piloted and developed through the ALIS (A-level Information System) service since the early 1980s by Fitz-Gibbon, then at the Curriculum, Evaluation and Management Centre (CEM) at the University of Newcastle (see Section 2.3 of the previous chapter for selected references).

McPherson’s (1992) briefing paper for the National Commission on Education (chosen as the first in a long and generally impressive series of Briefings for a general audience, and mentioned in Chapter 2 above) is more sanguinely entitled *Measuring Added Value in Schools*; it does a good job of summarising existing research in order to make the case for a more balanced and complex view of school comparisons. Whilst noting that there is a place for raw outcome measures as a reflection of actual attainment, his main argument is that *‘a bad indicator system’*, such as that based solely on raw test or examination results, carries the hidden costs of *‘mistaken judgements, needless anxieties and fruitless “further investigations”... triggered by false signals’*. The measurement of added value – explicitly defined as the *‘calculation of the contribution schools make to pupils’ progress’* – is central to developing a good indicator system. Such measurement must be predicated on *‘an explicit theory of good standing’* which is open to scrutiny and refinement, and includes such features as:

- pupils' prior attainment;
- the longitudinal nature of progress;
- the multilevel nature of schools;
- the multivariate nature of the factors involved, especially 'non-school factors that boost or retard progress', such as pupils' socio-economic background;
- differential effectiveness for different groups of pupils.

This is a more careful and weighty operational definition than that given by Kennedy, and shows an understanding that some of the 'technical questions' alluded to by him were capable of resolution at least in principle, even if the practical collection of necessary data was still problematic. McPherson provides a diagram – based on Paterson's (1991) research – showing how schools' relative performance becomes more similar, and how their relative positions change, the more adjustments for non-school factors are made. Moreover, McPherson seems to be the first writer to acknowledge the role that value judgements will always play in arriving at a 'good' information system on schools' results; and that what is useful information to parents in choosing schools may be less useful for teachers in setting about raising attainment, and *vice versa*. But 'useful' information is not inevitably simple information, even – or especially – where parents are concerned:

...complexity is not in itself an argument against aiming for the best possible indicator system. Any attempt to improve schooling by means of informing choice presupposes that parents are capable of understanding at least the complexity of an adjusted outcome score. To reject that possibility is to reject the possibility of informing parents.

Three of the Digests published by the journal *Education* in the early 1990s dealt with value added for a non-specialist but professional audience, mainly LEA and school managers. The first of these (Jesson, 1992) is clearly presented and argued, and uses analyses of GCSE results in different ways to illustrate the importance of the following principles:

- the need to base calculations on individual (not aggregate) data;
- the need to take account of prior attainment;
- the need to take account of gender;
- the need to use multilevel modelling;
- the need to use the model ‘not simply as a means of generating a more appropriate league table of “effectiveness” or “value added” but as a resource to prompt genuine questions of comparison between practices and procedures in contrastingly effective schools.’ (p.ii.)

It is worthy of comment, given that he argues in favour of the inclusion of gender as a background variable, that Jesson seems to discount the use of a social deprivation variable, on the grounds that if used ‘*as the only context within which examination results are evaluated, ... it can become an excuse for not expecting high achievement.*’ (pp.i-ii.) Jesson is not alone in expressing this concern (see, for example, Wiliam, 1992, cited below). But Jesson also supports the use of multilevel modelling as being ‘*essential to do justice both to the underlying relationships in the whole population of pupils as well as to the particular contribution made by each school*’. The point about multilevel modelling is precisely that it can include several background variables simultaneously and assess their relative importance in accounting for outcome variables. In any case, we saw that at least one commentator (Paterson, 1991, cited above) appears to have recognised that empirical relationships such as that between attainment and SES may be construed, not as evidence of an unchangeable fact, but as support for the necessity to change educational and/or social policy.

The anonymous author of the second *Education Digest* (anon., 1992) was not so convinced as Nuttall, McPherson and Jesson of the transparency of value added approaches. The article starts by warning that the interpretation of value added results ‘*depends on a number of important assumptions and qualifications which are not particularly easy to understand*’ and justified the emphasis on technicalities in the article by saying: ‘*However, if these techniques are going to be the main way we statistically analyse pupil achievement in the future, then it is vital that everyone in education understands what they can and cannot do*’ (p.i.). The article is based largely on a discussion of procedures used in one outer London borough on one year’s GCSE data, however, rather than on a critical overview of research evidence; and although the commentary makes some generally useful points about the importance of statistical validity and the role of the null hypothesis, as well as the varied influence of socio-economic factors, the text is (for this reader) too dense and laborious to fulfil its stated purpose. In concluding, the author takes up a

position on the use of the most up-to-date techniques (i.e. multilevel modelling) which contradicts that of McPherson and Jesson:

‘[The multilevel approach] is the most difficult to understand for the non-statistical reader... One should not underestimate the importance of this... Even the simpler approaches to value added can appear arid and remote... For this reason it may be that the use of multilevel analysis is not the best place to start.’ (p.iv.)

The third Digest on value added, authored by Birnbaum (1993), seems to make substantial use of, though without citing, this earlier article.

The fact that Nuttall’s, Kennedy’s, Jesson’s, McPherson’s and Birnbaum’s pieces – all using the term value added as if it were the *mot juste* that still required a little elucidation – were written in the early 1990s for a non-research audience is worth considering. As has been remarked, the relative lateness of the usage in compulsory education as compared with its established quasi-technical use in HE (discussed in the previous chapter) contrasts with the maturity of the research discipline which lay behind it. This suggests that the term ‘value-added’ was lifted from HE to describe a set of issues now of critical concern in the compulsory (and immediately post-compulsory) sector, but that it was initially reserved for generalist rather than specialist discussion because it seemed to be an idea whose common sense and common justice people could ‘intuitively’ (Nuttall’s word) grasp. (One may speculate that the *zeitgeist* had something to do with why this assumption – which the previous chapter showed to be premised on somewhat flimsy grounds – was so readily made by academics as well as by government policy-makers, who might be supposed not to mind so much about linguistic *laissez-faire*.)

So far as it is possible to tell, then, the term appears to have come into generalist use in work relating to the compulsory sector *before* its adoption in more scholarly discourse. This is hard to establish with complete certainty because of the time that often elapses between the writing and the publication of articles in academic journals. By 1992, Wiliam’s article for the *British Educational Research Journal* was addressing ‘*some of the technical difficulties in operationalising the notion of “value-added” in [the published results of aggregated school performance]*’ particularly as they related to National Curriculum assessments. (Wiliam gives a detailed discussion of the nature and structure of these assessments, which, whilst conceptually helpful about the limitations of the assessment framework, does not directly concern the present discussion.)

Wiliam uses quotation marks throughout to indicate the provisionality of the term 'value added', which he does not explicitly define himself but for which – since he cites McPherson (*op. cit.*) – he can be taken to have adopted McPherson's definition. He agrees, on certain conditions, with McPherson's statement that socio-economic factors as well as prior attainment may influence pupils' performance; but goes on to argue on theoretical and ethical grounds that SE data should not be included in a 'value added' analysis of performance. Although McPherson argues for '*an explicit theory of good standing*', Wiliam says, no such theory at present exists with regard to the type of SE indices that should be included. He argues that different studies have used a wide range of different variables, with differing results in terms of the correlations found. Furthermore, Wiliam considers that the act of computing socio-economic correlations may have social consequences of the 'self-fulfilling prophecy' type or, in Wiliam's words: '*it is extremely doubtful that the use of socio-economic factors in value-added models will do anything to combat under-achievement by students, and may serve to consolidate it.*' (p.338.) Given an adequate explanation of what value added analyses do *not* reveal as well as what they do, this should be a misplaced anxiety, albeit one shared by other commentators such as Jesson (1992). Wiliam also disagrees with the suggestion made by McPherson and Nuttall that differential effectiveness is important to report on, because: '*the use of multiple indices is likely to cloud, rather than clarify, the picture*' (p.339). Whilst one may in retrospect discount some of Wiliam's argument, his major contribution in this article was to raise the taxing practical issues – which the School Examination and Assessment Council (SEAC) had so far avoided – associated with using the new National Curriculum assessment framework for comparing pupils' progress in different schools. Most of the school effectiveness research at that time was by contrast based on public examinations and/or standardised test results.

The two Discussion Papers on value added issued by the Centre for the Study of Comprehensive Schools in 1993 (Fitz-Gibbon and Tymms 1993; Cornall and Lofthouse 1993) are chiefly interesting for the fact that, again, they are written for a non-specialist audience and with the weaknesses of raw league tables very much to the fore. It is perhaps unfortunate that the papers do not relate well to each other, and indeed on one or two questions – such as whether SE factors should be taken into account in calculating value added or if individual schools can 'do' value added analyses for themselves – offer conflicting answers. On the specific issue addressed by Wiliam (1992, see above), Fitz-Gibbon and Tymms are sarcastically dismissive of the possibility of using National Curriculum standard assessment task

(SAT) data for calculating value added (a position which surely needs to be set against the fact that they bid for and won the contract for using National Curriculum standard assessment task data in the national value added project – see Chapter Four); Cornall and Lofthouse write abstractly about the use of ‘pupils’ test scores’. The latter paper is really only a polemical statement of position and the authors do not attempt to tackle seriously the conceptual and technical problems raised by other commentators. The reason for inclusion here is the context, i.e. briefings for schools which might be thought to be particularly vulnerable to raw league table results.

In his report for the Scottish Office, Gray (1994a) states the purpose of the value added approach as ‘comparing like with like’ (p.2): *‘if there are two pupils who are identical in every respect how much difference does it make to their educational progress if they attend different schools?’* With the benefit of work piloted and developed over a number of years with different LEAs, Gray gives the basic requirements for building a value-added approach, whose similarities with criteria offered by Jesson (quoted above) are without doubt due to their collaboration over a number of years. Gray’s text, slightly abbreviated from the original, runs as follows:

- data on individual pupils rather than aggregate data, including:
  - a measure of outcome for each pupil which reflects all levels of pupil performance
  - a measure of each pupil’s attainment, preferably one which is finely differentiated, plus one or two other items of information about pupils’ background
  - in the absence of prior attainment, several items of information about each pupil’s background (ideally including a measure of social advantage as well as social disadvantage);
- analysis of data using multilevel modelling.

The report, though short, is the first publication (as far as I am aware) to cover issues and challenges in the implementation of value added approaches, as well as some discussion of basic technicalities. Amongst the sixteen questions he identifies for schools and LEAs to consider, for example, Gray points out that:

the most important issue to be addressed relates to the overall purposes of introducing value-added approaches. Is the main purpose to hold schools directly accountable for their performance? Is it to identify schools that are markedly under-performing and may even be ‘at risk’? Is it to help schools

justify their existing levels of performance? Or is it to help them to understand more about their own performance and how to improve it? If it is to be about several of these, are they all equally compatible? (p.9.)

The significance of this report lies not merely in the clarity and usefulness of the content but in its audience: the work was commissioned by the Audit Unit in the Scottish Office Education Department in order to inform the school development planning initiative in Scotland at a time when the English Department for Education was still acting as if afraid that anything other than raw results would be too complicated for public consumption and would encourage schools to make excuses for poor standards.

But value added was a persuasive idea and government agencies were slowly convinced of the need to have better measures of school performance than raw results could provide: the political watershed was marked in 1994 with the publication by the School Curriculum and Assessment Authority of *Value-Added Performance Indicators for Schools* (discussed in its own right in the following chapter). As the familiarity of the idea grew from 1994 onwards, the term, together with its functions and supposed advantages, became more or less taken for granted. Instead, as the literature shows, attention turned to other underlying and possibly intractable issues which needed further explication. Paradoxically – in view of the arguments initially advanced for value added by many of the commentators discussed above – one of those issues was the question of whether value added analyses could really be used for public accountability systems.

The immediate occasion for the 1994 paper by Mortimore *et al.* was the death of Desmond Nuttall: the paper was intended to be a tribute to his work and to that end the paper offers a summary of the main issues relating to school effectiveness and the measurement of value added, Nuttall's theoretical and practical contributions to which had become internationally recognised. This paper is explicitly situated in the school effectiveness tradition of academic research, and is important for the overview it gives of this research (albeit with a particular emphasis on Nuttall's work). It is also interesting, in this scholarly context, for its unremarked use of the term value added in the title and throughout. A definition of value added is not attempted (presumably because by now not felt to be necessary), but the term is used with a positive connotation, in the context of the provision of public information, thus:

Raw results tell parents the grades their children have obtained but they can say nothing about how well the school attended has performed. In contrast, value added results tell parents how effective their children's school is in promoting achievement. (p.319.)

The paper is exemplary for the way it foregrounds the areas where there are as yet unresolved issues and provides examples of conflicting evidence – in relation to stability of school effects over time, differential effects for different groups of pupils and the effects of whole-school contextual factors over and above individual pupil characteristics – to demonstrate ‘the varied and unpredictable impacts of schooling.’ (p.329.) It calls both for further research into these impacts and for a broader range of outcomes to be analysed beyond the narrowly academic.

A contrasting note is sounded by Saunders and Schagen (1994) who take issue with the way value added has become a new ‘buzz-word’ in education:

[value added] has become a rallying standard to signify opposition to crude comparisons between institutions without taking any account of the students with whom or the circumstances within which they are working. In the process, however, the meaning of the term has become understandably rather blurred (p.1.)

They say that ‘*much useful analysis is more properly classified under the general title of “school effectiveness”*’. The rest of the paper is a description and evaluation of how the ideas and techniques in Schagen's earlier work (1991) have developed operationally, through the piloting of the QUASE (Quantitative Analysis for Self-Evaluation) Service for secondary schools, sponsored by the National Foundation for Educational Research (NFER). This service was set up to operate on the same principles of ‘good practice’ as ALIS (mentioned above) which are that:

- it provides a confidential service to schools in which schools are given only their own results and no league tables are formed;
- results are fed back ‘*in a digestible form which can influence educational practice*’ (p.1).

Saunders and Schagen record the results of the QUASE feasibility study as follows (p.7):

- it was possible to combine results from disparate tests at age 11 or 12 to give a measure of pupils' prior attainment on entry to secondary school;
- multilevel models using such intake measures, plus school context data, can be used to provide schools with sensible performance information;
- the feedback from the QUASE pilot project was appreciated by schools and frequently tied in with information from other sources;
- some useful feedback can be provided to schools with no intake data, but this is necessarily more tentative.

They do not mention here that QUASE uses a series of separate departmental indicators as well as more a total of seven more generic ones like total and average GCSE scores, thus operationalising, at least in part, the notion of differential effectiveness. Saunders and Schagen conclude that:

The quantitative results should not be seen as ends in themselves but as a contribution to the process of school improvement. There is a need to work interactively with schools to help them integrate quantitative information of this type into their school development planning. (p.8)

Overall, the paper shows, in clear practical terms, how the criteria and principles agreed by people like Goldstein, Gray, Jesson and Nuttall may be operationalised. (See Appendices A and B for further information on QUASE.)

Fitz-Gibbon and Tymms (1995) is another milestone, in being the first of the many reports emanating from the so-called Value Added National Project. (The final report and accompanying technical reports are considered in the following chapter, as belonging more to the policy than the academic arena.) In it, this definition of value added is given: *'Measures of ... relative progress have come to be called "Value Added".'* (p.2.) This is elaborated on as: *'If some students make greater progress than other, similar students, the difference is referred to as the Value Added.'* (p.7.) As has been argued in the previous chapter, this is not the most rigorous nor indeed the most helpful definition, because it fails to make explicit the need to identify *schools'* contribution to pupils' relative progress; but it has come to inform most of government thinking since. To summarise a substantial piece of work, it may be noted here that the report considers issues of variables, coverage, analytical method, reporting and interpretation for each of the different approaches selected for comment, and where appropriate discusses the pros and cons involved. It concludes that: *'the statistical trialling undertaken to date has lent assurance to the view that simple, understandable and statistically acceptable Value Added indicators can be made available when suitable input and output measures are in*

*place over reasonable timespans.*' (p.29.) One of its findings – at odds with what was said in the earlier SCAA report (1994, p.84) – was that the National Curriculum assessments could provide such measures. Given the *caveats* made by the CEM team in the past about over-reliance on performance data and especially National Curriculum assessments for school evaluation purposes, it is noteworthy that – with the exception of school data used for the pilot studies about to commence, which was to be kept confidential – the purpose of a national value added system was at least in part to publish the results for '*parents and the public*' (p.19).

The view that value added data could and should be in the public domain was not shared by Goldstein and Thomas, 1995, who by the middle of the decade had arrived at the position that even the 'private' use of value added data by individual schools may be misleading. They point first to the inevitably historical nature of performance data by the time test or examination results are available to be used as outcome variables, such that schools may have completely changed their policy or practice in the meantime; and then to a much more serious problem which they state like this:

While we can study the factors associated with student performance and come to conclusions about which of them appear to be associated with 'success', yet it seems to be very difficult to identify precisely which schools are doing well or badly... In other words, research into school effectiveness is a useful activity in our attempts to obtain knowledge about the process of education, but a very poor tool for holding schools to account. (p.37.)

Gray's (1996) concern is similar to that of Goldstein and Thomas, and '*is premised on the assumption that ... comparisons [between schools] will continue and that the most helpful service can be rendered by reminding those who would make judgements of some of the problems to be encountered.*' (p.121.) In effect, the chapter is an elegant and concise summary of the technical and conceptual issues discussed elsewhere, and thus requires no further description. Its significance lies more in the fact that it is another example of the phenomenon that commentators who appeared to have much in common in the early 1990s are, by the latter half of the decade, taking up opposing positions.

Jesson's (1996) report performs the same sort of function for the DfEE as Fitz-Gibbon and Tymms's did for SCAA, that is, it reports on the feasibility of using current assessment data, in this case Key Stage 3 Assessments, '*for*

*evaluating differences in performances by pupils and their schools'* (p.12) at GCSE. Broadly, there was found to be sufficient correlation to provide '*a considerable opportunity for comparative evaluation of institutional performance using a common "starting point"*' (p.12.), although the author made it clear that he did not yet regard this as a reliable and consistent guide to value added assessments of school performance.

Thomas and Mortimore's 1996 study involved '*considering and evaluating a variety of different models for measuring school effectiveness using sophisticated statistical techniques (multilevel modelling).*' (p.6.) They particularly contrasted the results of controlling for prior attainment with those of controlling for SES factors. Amongst the findings, they report that: '*when prior attainment data are available no school context factors are significant and the fit of the model is substantially improved.*' Other studies, however, have shown some negative correlation between low SES and academic performance, even when prior attainment is controlled for. Paterson (1991) has a thorough discussion of statistical analyses purporting to show that '*SES is an important correlate of attainment over and above ability, and that the SES mix of the school has an effect on attainment over and above individual or family SES.*' (p.97) Sammons (1996) found that: '*by controlling only for prior attainment... a misleading picture of effectiveness was obtained for some individual institutions*' (p.119), and Schagen (e.g. 1996) has found consistently that although prior attainment data explained a great deal of the variance in schools' GCSE results, allowing for school context factors – including parental attendance at parents' evenings as well as overall proportions of FSM ([pupils entitled to] free school meals: see Section 3.4 below) – explained even more of the variance.

But, over and above the question of whether there is any statistically explanatory power in doing so, Thomas and Mortimore offer another version of the conceptual rationale for excluding context factors. They too argue that school socio-economic context may influence teacher expectations of pupils, with a resultant impact on pupil outcomes. This argument is addressed in Section 3.4 below.

Usefully, Thomas' and Mortimore's study also confirmed the likelihood in most schools of differential effectiveness, for different subjects and different groups of pupils. They conclude with a call for further research to explore '*the relationship between negative value added results and measures of school processes, in particular, the quality of teaching and learning.*' (p.28.)

Thomas' *et al.* (1997) report on research undertaken, using what they call 'a value added approach' (i.e. multilevel modelling to control for student intake measures) into the stability and consistency of schools' GCSE results. Their conclusion is that '*effectiveness is best seen as a feature which is outcome and time specific*' (p.194). They summarise the current position on school effectiveness as showing that any evaluation of performance needs '*to address three key questions: (1) Effective in promoting which outcomes? (2) Effective over what time period? (3) Effective for whom?*' (p.194).

By the late 1990s, then – the point at which the government had chosen to espouse the value added principle – there appears to be considerable agreement in principle about why the value added approach was found necessary in the first place. There is somewhat less consensus on what information of the kind that could loosely be called value added would most help teachers in improving, and parents in choosing, a school and what it consequently means in terms of methodology. The initial reason given by commentators for undertaking value added analyses seems primarily to add to, modify or otherwise elaborate on the published tables of performance based on raw results. The market principle that parents could and should be making informed choices about schools for their children thus seems to have been accepted by researchers including most of those discussed above. So we can say unequivocally that one of the intended functions of the value added approach was to render public information about schools' performance more accurate in a whole variety of ways for that purpose.

Another related, and for some head teachers an over-riding, function was to enable schools to mitigate the message given by their raw results, to counteract bad press. As Buck (1993) put it: '*This exercise [a method for contextualising GCSE results in one LEA]... provides a fast response to potential condemnation in the public domain.*' (p.92.) So far, then, there was an acceptance that the role of value added analyses was directly related to the reconstruction of published 'league tables'. Since the conversion of the then Secretary of State for Education under the Conservative administration, Gillian Shephard, to the principle of value added, this has continued to be the government's position up to and including the 1997 proposals to establish a national value added system.

But, as was suggested above, it is not a position held by everyone. Amongst others, Tymms (1990), Fitz-Gibbon (Fitz-Gibbon and Tymms 1993, Fitz-Gibbon 1994),

Saunders (1997), Saunders and Schagen (1994), Schagen (1991 and 1997) and particularly Goldstein (see, for example, Goldstein and Thomas, 1995; Thomas and Goldstein, 1995; and, most recently, Goldstein, 1997a, 1997c) have argued against this public accountability use of value added analyses. The dilemma can be stated like this, that on the one hand the *principle* of value added does seem to have been readily comprehensible, in Nuttall's (1991) words, an idea whose time had self-evidently come. On the other hand, the more rigorous the actual value added *analysis* becomes, in terms of underpinning theory and appropriate modelling techniques, the closer we get, in Thomas and Goldstein's (1995) words, to '*the complex, multi-faceted and often inconsistent nature of school effectiveness*'.

To put it at an extreme, the value added task began by appearing to promise better information for public consumption, but instead turned out to demonstrate that 'better information' and 'public consumption' were incompatible, if the latter depended on there being 'simple and straightforward' measures of progress (SCAA, 1994). As Thomas and Goldstein (*op. cit.*) remark, '*research emphatically demonstrates that the measurement of progress or value added... is neither simple nor straightforward*'. It is important also to record that an article for the TES (Kellner, 1997) drew attention to the obvious professional concerns of Fitz-Gibbon and Tymms about the national value added project and its applications. Their concerns are perhaps symptomatic of scholars' understandable caution when they see policy-makers and the public expecting too much from an approach they themselves have pioneered.

There are two ways of going on from here: one could make the case along with Thomas and Goldstein for further centrally-funded and '*well-designed, long-term research into the processes which shape the achievements of different kinds of pupils in different types and organisation of schools*'; or one could follow Tymms and Fitz-Gibbon (Tymms, 1990; Fitz-Gibbon and Tymms, 1993; Fitz-Gibbon and Tymms, 1995), at least in their earlier articles, Jesson (1996) and Saunders and Schagen (Saunders and Schagen, 1994; Saunders 1996, 1997) in arguing for the installation of 'feedback' systems designed to enable schools to make use of value added information for autonomous development and improvement. (This was the official position of the Scottish Office Education Department as far back as 1993: see foreword to Gray, 1994a.)

These alternatives are not mutually exclusive, of course; but at the moment stalemate seems to have been reached, in that – despite the weight of evidence

against it – the Labour government is if anything even more determined than the previous administration to pursue the notion of value added league tables, or at least to give prominence to numerical measures of progress and improvement within the public accountability system, without necessarily being committed to funding the more sophisticated research which would give such measures more validity and reliability.

The political climate has seemingly led to a situation in which practitioners of value added tend to fall into one of two groups: the pragmatists or the idealists. The pragmatists are those who are inclined to say that the idea of value added is here to stay and the sooner we develop a system which is workable and comprehensible the better (*sc.*: otherwise someone else will). Credibility with teachers and parents and above all user-friendliness for school managers – now charged with making year-on-year measurable improvements to their pupils' performance – are more important criteria than theoretical completeness or statistical purity. The government and its agencies are firmly in this category. Idealists see themselves as keepers of the flame of school effectiveness research, which long pre-dates current obsessions with public accountability and the workings of the educational market-place. They would argue that using procedures which are not as accurate and rigorous as state-of-the-art modelling techniques permit will bring the whole idea of value added into disrepute. The country's education system will end up (in their view) with a method for making judgements about educational quality on the basis of adjusted performance data which is even more unreliable and disinformative than raw data. For a good-humoured but incisive version of this position, see Gray, 1994a. Some, like Goldstein and Thomas (1995), would apparently go even further and say that value added, as an assessment of any individual school's performance, is inherently impossible, because of what they call '*a kind of uncertainty principle*' (p.37). So the wheel seems to have come full circle, with the government supporting value added approaches and some very respected academics arguing against them.

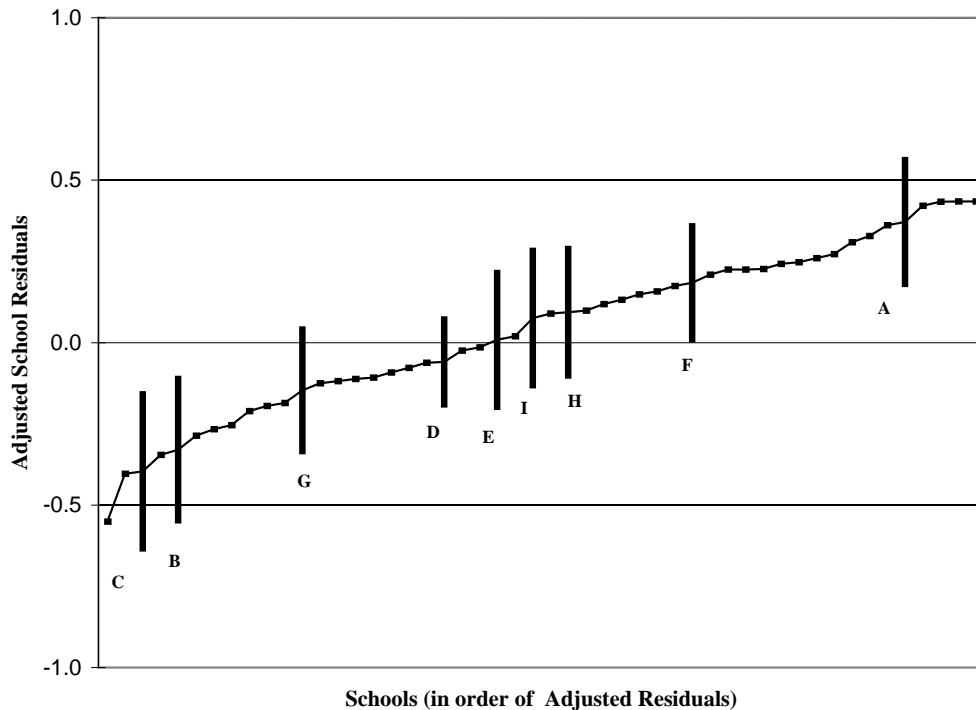
### **3.4 A Note on Analytical Models and Methodological Strategies Used in the Measurement of Value Added**

It may be helpful at this point to stand back a little from the historical narrative and focus briefly on what needs to be said in this thesis about the analytical models and methodological strategies associated with the measurement of value added. The thesis does not presume to be technically oriented, however, and what follows is no

more than a superficial gloss – mainly for clarificatory purposes – on an extensive body of academic work.

First of all (to reiterate a point made on page 29 above) it must be understood that, as a way of investigating school effectiveness, the application of what have come to be known as value added models has a peculiarity not readily obvious to the classroom teacher or statistical layperson. Unlike other forms of research, which rely on the steady accumulation of evidence, value added has an essentially *eliminative* basis. That is to say, as the author of an Education Digest (anon., 1992) pointed out, value added is an attempt ‘*to make some estimate of the total influence of these ... factors [i.e. factors outside the school’s direct control, such as pupils’ prior attainment] by a process of elimination.*’ Value added is the statistical discrepancy – the residual – between the actual score and the score calculated when factors known to have an effect on performance have been accounted for. Or rather, value *added* is represented by any residual above zero, i.e. with a positive value: since it is part of the modelling process to relate each school’s score to an average or expected value derived from the sample population as a whole, some schools or departments within the population are bound to end up with a negative residual. It is now normal good practice to state the degree of certainty with which the residual can be assigned, using some statistical index of significance, usually at the 95 per cent confidence limit. Very commonly, school residuals arrived at by the above method are not significantly different, statistically speaking, from the norm, that is, from what would be expected. Figure 3.1 below gives a diagrammatic representation (commonly known as a *Goldstein plot*, for reasons which must be self-evident) of this phenomenon. For an explanation of ‘residuals’, etc., see the Technical Glossary in Appendix A.

**Figure 3.1: Diagram representing nine secondary schools ranged in rank order of their mean average GCSE score, adjusted for pupil and school residuals and showing 95% confidence intervals**



Only three schools, A, B and C, have a pattern of mean average GCSE scores such that we can be sure that their performance is statistically significantly different from the average or what would be expected on the basis of pupil and school factors (this average is represented by the zero line). In the case of schools B and C, the scores are significantly lower than would be expected; in the case of School A, they are higher.

*Source: National Foundation for Educational Research, QUASE analysis 1999.*

Two or three things need to be said about this, from the point of view of interpreting data:

- Assigning value added to a school's results is equivalent to the theological notion of 'the God of the gaps'. The residuals demonstrate only that there is a positive (or negative) difference from what would be expected; they do not identify what is responsible for that difference.
- Usually, statistical residuals are quite small. Even so, 'in terms of differences between schools in students' outcomes they can be highly significant both educationally and statistically... [Such] differences... have significant implications for students' subsequent education and employment prospects' (Sammons, 1996, p.120).
- Since the residuals for most schools are not statistically significant, it is very hard to argue that in the majority of cases schools are performing much better or worse than would have been expected. That this is true does not help the cause

of league tables, whose *raison d'être* is to rank schools as if there were a real difference between them.

The second area which needs some further comment beyond what has already been said in this chapter (page 43 *sqq.*) has to do with whether or not socio-economic variables should be included in the analysis, provided that there is i) a reasonable proxy measure for SES which can be used and ii) that there are good statistical reasons to believe that SES is correlated with performance. To deal, rather cursorily, with the first proviso: eligibility for free school meals (FSM) has long been taken as a proxy for low SES, although with great reservations often expressed by those who do so. The major reason for using it is its relative ease of collection. More evidence is needed to assess whether having more detailed and accurate information on SES at the pupil level gives results significantly different from those arrived at by using the FSM indicator. Evidence in support of point ii) has already been presented above such that it would seem mistaken to conclude that there is no effect on pupils' and schools' results of poverty and other types of disadvantage.

The serious objections to including SES in the assessment of schools' performance therefore seem to relate more to questions of ethics and politics than of statistics. They can be summed up in the way the Minister for Schools responded when tackled about the continuing out-performance by schools in leafy suburbs of those in inner cities: 'poverty is no excuse'. Two distinct issues need to be separated out of this single soundbite, one concerned with the assessment of schools' performance and the other with managing the motivation of young people. On the first issue, since there continues to be an empirical association between pupils' academic performance and their SES, it would seem intellectually dishonest not to take account of such information when attempting to assess the relative effectiveness of schools serving widely contrasting catchment areas. Such assessment is, of its nature, historical: it has little to say about what might be the case in some differently constituted future (as has been said, the correlation is empirical not *a priori*) and even less about what ought to be the case (how far public educational resources should now be targeted on the poorest schools is a political question which the empirical data clearly poses). For a cogently argued position paper on this issue, the 1997 paper by Mortimore and Whitty would be hard to better (Mortimore and Whitty, 1997).

In fact, one could argue – as Saunders (1997) has done – that the *different* value added scores which a school may obtain depending on whether prior attainment

only or prior attainment plus school context data is included can be used to demonstrate to school staff whether or not their expectations have been influential in pupil outcomes. This suggests that the potential importance of using more than one model *in the same analysis* needs to be more fully explored. And of course this issue must be related back to the question posed in the previous chapter, about the multiple and conflicting purposes of undertaking value added analyses in the first place.

On the different question of managing pupils' motivation, it is easiest to argue from the statistical evidence itself. Although FSM is a significant variable correlating with performance at the school-level, most studies find that it has little or no explanatory, i.e. predictive, capacity at the individual pupil level. From a statistical point of view, what matters more than any other variable at this level is prior attainment. If teachers stick to what the data tells them, therefore, there should be no question of 'poverty being an excuse', in terms of withholding encouragement, support and challenge for every pupil regardless of background. The problem, it seems to me, is not so much that some teachers misuse statistical, including SES, data but that they go by subjective impressions. Sometimes they do so to the extent of (unconsciously) entertaining lower expectations of pupils from, for example, particular housing estates when the data – if used and understood – would tell a different story.

Jesson (1992, p.i.) goes even further in putting health warnings on the analysis of individual pupil performance. Even prior attainment is not to be trusted as a predictor: *'individual pupils are not pre-ordained to score at any particular level at GCSE by virtue of assessments of their "ability" made at some earlier stage in their career.'*

A key lesson to be learnt, then, is that statistical analyses can be very powerful when used to assess the past performance of aggregates; but weak and probably misleading if used to predict the future performance of individuals.

Finally, there is also a need to recap on what some might see as a fine distinction between 'value added analysis' and 'contextualisation'. TGAT conflated these two approaches, according to Buck (1993), and others have done so since. Contextualisation properly refers to the allowance made for SE factors in the computation of schools' results. Probably the best known example of this is to be found in the work of the National Consortium on Examination Results (NCER), which has for many years been providing GCSE results to LEAs (and thence to schools) on this basis. In this approach, the data is collected and analysed at school level, and regression analysis is the method used. Because of the demonstrable negative relationship between academic performance and SES, the outcomes of this analysis were perceived, and welcomed, by LEAs and schools as more realistic assessments of relative performance than reports based on raw results alone. But there are two main objections to equating this approach with an analysis of 'value added'. If we turn back to Jesson's and Gray's (1992/3) criteria, three of the five technical criteria remain unfulfilled, namely, the analysis of data at the individual pupil level, the inclusion of a prior attainment score and the use of multilevel modelling as the statistical technique. This last stipulation is made because, in trying to organise evidence on the performance of students for the purpose of measuring the relative effectiveness of schools, we need to deal with what is known as hierarchical data. Goldstein (1987) puts the argument like this:

... the existence of hierarchically organised data implies that we should take that hierarchy into account when we analyse data... the failure to account for hierarchies may lead us into trouble... If we were interested in the factors which influence student attributes, say their educational achievements, then among those factors we will generally wish to include the characteristics of the students themselves, their classes, and their schools. (p.3.)

There are now very few serious scholars in the field of school effectiveness who would argue against the use of multilevel modelling. Those few who do, like the

commentators reviewed in Section 3.3, tend to do so on pragmatic grounds, in that multilevel modelling is held to be less transparent than regression analysis and needs to be done by trained professional statisticians. But the point is surely that the method of analysis needs to be consonant with the complexity of the data and the issues it is attempting to elucidate rather than tailored to the limitations of one particular audience or end-user. It then follows that there is an interpretative as well as an analytical task to be done by the scholars (and others) providing value added analyses. This has ethical as well as practical implications, and is argued in greater depth in Chapter Five.

### **3.5 What Have Value Added Analyses Revealed About ‘Effectiveness’ and Educational Quality? A Short Summary**

Value added analyses, as I have explored them earlier in this chapter, have provided information on the following key issues, and – perhaps even more importantly – have made us able to think about those issues in evidential rather than just common-sense terms:

- the size of the school effect, relative to other effects;
- factors beyond schools’ control correlated with academic performance, in particular the difference made by the overall composition of a school, in academic and/or socio-economic terms;
- evidence of ‘differential effectiveness’ in the same school (e.g. positive results for some pupils but not others, related to their gender, ethnicity, prior attainment, SES);
- evidence of variable effectiveness (e.g. between departments in the same school);
- evidence of lack of stability in schools’ results over time.

Thomas (1998) gives a clear and comprehensive overview of what the technical research on value added measures of school effectiveness has achieved; this research is further updated (Thomas, 1999) in a comparison of optimal multilevel models of school effectiveness across regions.

Further questions which colleagues in the field of value added approaches are seeking to explore include:

- What can be said about social as distinct from academic effectiveness (as measured by attitudinal data and/or attendance)? Are the independent variables correlated with social effectiveness the same as for academic effectiveness?

What evidence is there of stability/consistency/improvement in terms of attitudinal factors in individual schools?

- How far is academic effectiveness correlated with social/motivational effectiveness within the same institution?
- How might the correlates of *changes* in school effectiveness (e.g. towards improvement) be identified?
- Given that research suggests that the greatest variance is within-school rather than between-school (see, e.g., Hill and Rowe, 1996), how might the relative effectiveness of teaching groups/classes, etc. be explored? What *caveats* would need to be made on such research from a statistical point of view?
- Is it possible to investigate ‘*how the variation [in effectiveness by... ability group, by social class, by ethnicity and by subject department] is maximised or minimised within different types of “effective” or “ineffective” schools?*’ (Reynolds, 1995.)
- Conversely, institutional ‘effectiveness’ might *prima facie* be thought to be correlated with the level and type of LEA or other external support. How could this be explored using multilevel approaches?
- Going on from here, what other important factors – notably i. the effect of previous institutional impact on pupils’ performance; ii. the relational or micropolitical aspects of schools; and iii. ‘partnership’ models of education which utilise agencies to work in or alongside schools – should now be brought into the analysis of school effectiveness? How could this be done validly and manageably?

The concept of value added, together with more sophisticated ways of treating data, seem here to stay, although the purposes and functions of value added continue to evolve in interesting and politically responsive ways.

What further or deeper questions are raised by value added analyses – in other words, what value added analyses as currently understood leave out of account as well as what they reveal not only about schools but about social and political processes in education – are worth summarising as a conclusion to this chapter.

### 3.6 What Further Questions and Queries Do Value Added Analyses Raise?

Whilst Fitz-Gibbon's 1996 definition of value added as '*the statistic of choice*' may not commend itself to everyone, it is a clear indication that the value added approach is now seen as both less problematic – in terms of how acceptable the principles are to, say, policy-makers and school senior managers – and less problematised, i.e. less likely to be questioned even in the most sceptical of staff rooms.

As we have seen in this and the previous chapter, value added started its educational life as a contested educational concept, in which questions of value judgement were explicit: was it right, and feasible, to measure schools' effectiveness by some other means than absolute academic standards? The rightness was accepted without hesitation by school senior managers and LEA personnel who were experiencing the fall-out from 'raw' league tables of performance. Questions about feasibility, and thus about statistical method, then became paramount and the mid-90s saw a proliferation of discussion papers and articles about how to do it. The more the methodological sophistication was developed, however, the more complex an undertaking value added was revealed to be, and *pari passu* the less appropriate for making simple summary judgements about 'better' and 'worse' schools.

But the public purposes of value-added have continued to be multiple, espoused by policy-makers as well as by school senior managers; broadly speaking, they can be grouped as follows:

- the development of better and fairer ways in which to assess school effectiveness, overall and by institution: 'effectiveness' has therefore come to mean academic progress made as well as standards reached;
- the deployment by schools and LEAs of 'finer grids' with which to examine local and departmental differences, and identify action points for school improvement.

That these purposes do not sit comfortably with one another seems to be recognised at present only by academics in the field, and not universally by them. In any case, there are limits to any kind of quantitative data and its application, *qua* value added, to the field of school effectiveness and improvement which need to be recognised. These can be summarised as:

- There is not yet an agreement that the right measures are being used (e.g. for measuring prior attainment and/or SES) nor that the most rigorous models are being universally applied to possible contextual effects. Two examples of the

latter are as follows: (i) Willms (1985) found that the higher the average attainment level at school leaving age of pupils in a school, the higher the achievement of its individual pupils; (ii) Sammons (1996) argues that the existence of a 'primary effect' has not been taken into account in most secondary school analyses and that these may therefore be '*seriously mis-specified*' (p.131). There is probably some way to go before scholars consider the conceptual models to be finalised.

- Moreover, the theoretical body of knowledge about the ways schools influence pupils' outcomes is not well developed: research is needed both into the generalisability of models of educational effectiveness and into the links between class/teacher and school effects (Sammons, 1996, p.142). As Goldstein (1997b) says:

... little of the true potential of school effectiveness research has yet been realised. Even the minimum requirements for valid inference are demanding ones. They imply considerable expenditure of thought and resources as well as the long term follow-up of large numbers of students, schools and classrooms. They need to be replicated across educational systems, phases and types of schooling. It follows that those engaged in school effectiveness research need to convey this modesty of knowledge when they make claims to understanding. (p.394).

- Quantitative data represents only one aspect of schools' performance. Not all desirable outcomes are easily measurable. (Nor is all that can be measured equally important.)
- Quantitative data is only one instrument of evaluation: it reveals correlations, not causes, and it deals in aggregates, not individuals. It needs to be used in dialogue with qualitative evidence and professional judgement.
- Because it is norm-referenced, value added analysis does not set new levels of goals/targets which may be needed.
- What is a good or effective or successful school? The existence of value added analyses has not done away with the need for establishing a public consensus on this question; and probably the need to redefine 'effectiveness' will constantly recur as changing labour and social markets continue to change our perceptions of what education is for.

At this point (i.e. around the time of the introduction of a national value added system) ‘effectiveness’, ‘improvement’ and ‘raising standards’ seem to be almost solely concerned in practice with test and examination results: value added has done nothing to undermine what one might call ‘the tyranny of the measurable’, and has even perhaps slowed down developments in more qualitative research on effectiveness (such as those developed by, for example, MacBeath *et al.*, 1996). Furthermore, the problems which most seem to concern policy-makers are those of methodological procedure. Even for the relatively restricted type of outcome measure which value added analyses currently provide, not many commentators have explicitly raised the issue of ethical considerations, including whether a protocol for the use of value added analyses is needed (but see Goldstein and Myers, 1996; Saunders, 1997). Whereas this may turn out to be the area where most effort is needed, especially with the introduction of ‘thresholds’ for awarding performance-related salary increases to individual teachers which are to be based partially on measurable pupil progress.

Notwithstanding these strictures, this chapter has provided evidence that value added has contributed enormously to understandings of the measurable aspects of school effectiveness; it has also served to show that, in its educational context at least, it ought to be seen as one of those ‘*metaphors and analogies, which [scientific thought] must both use and suspect.*’ (Byatt, 1997, p.251; my emphasis.) The potential must somehow be preserved for ‘*generat[ing] possibilities for things to happen that are closed off by the epistemologies of certainty.*’ (Stronach and MacLure, 1997, p.5.)

It is in the nature of governments to deal in epistemologies of certainty, however, and it is to the two major areas of current policy concern mentioned above – measuring relative progress and assisting with school improvement – that the next chapter is devoted.

## **CHAPTER FOUR**

### **A SHARPER STICK TO BEAT THEM WITH? REVIEW OF THE POLICY LITERATURE**

#### **4.1 Introduction**

This chapter briefly discusses the key policy-based literature on educational value added during the 1990s. This is done primarily in order to understand how successive administrations applied the notion of ‘value added’ to evidence about standards, progress and effectiveness; and to attempt some evaluation of how far the national value added system, in which the debate has in one sense culminated, is likely to improve the quality of that evidence. The chapter concludes by suggesting that, whilst the introduction of a national value added system has much to recommend it, there may still be crucial work for policy-makers to encourage and support if value added indicators are to have a beneficent impact on school processes.

#### **4.2 Changes in Official Policy on Value Added**

It is well known that the Conservative administration took some time to come to terms with the idea that it might be helpful to take account of background factors in assessing how effective schools were. This is the then Parliamentary Under-Secretary (Michael Fallon) (Fallon, 1991) responding, in a radio broadcast in 1991, to issues raised by the analysis of 1990 GCSE examination results for the Association of Metropolitan Authorities (Nuttall *et al.*, 1992):

[Interviewer: People would say that gender and social class and one-parent families and so forth, these are very important backgrounds and should be considered when school tables are published.]

Well I’m sure parents will consider them but what we are not going to allow all the educational experts to do is to dress up these tables with a whole lot of sociology so that parents can’t actually make valid comparisons from one school to another... All these issues will be debated I’m quite sure as the Bill progresses in Parliament, but we don’t underestimate parents. We think the important thing is that parents should be the interpreters of this information and the last thing we are going to do is let the education industry in on the act to start glossing up all this clean data

with all kinds of educational sociology which we simply don't think is valid.

The Secretary of State in 1993-94 (John Patten) remained adamant that he wanted no truck with what he called 'cooked' as distinct from 'raw' data on schools' results. He was replaced in July 1994 by Gillian Shephard. A signal that the government might now want to take notice of 'cooked' data was contained in a press notice launching the government's new initiatives on teacher training; one of the innovations was to be the placement of pupil teachers in urban schools which were doing well despite appearances. Over the next few months, there were accumulating signs that the Secretary of State was attracted to what value added analyses seemed to offer: a fairer way of comparing schools' performance.

### **4.3 The Value Added National Project**

In fact, such a change of policy had already been well prepared for by Sir Ron Dearing's recommendation, contained in the interim report on the National Curriculum and its Assessment (Dearing, 1993), that work be commissioned into value added performance indicators. This work was undertaken during 1994 by a working party convened by the then School Curriculum and Assessment Authority (SCAA) – later to become the Qualifications and Curriculum Authority (QCA) – and chaired by John Marks. The SCAA working party produced its report and recommendations in October 1994 (SCAA, 1994). Sir Ron Dearing, in commending the report to the Secretary of State, was explicitly ambitious about its implications: '*... it should be possible for us to make proposals for the development of a national system.*'

The 25 detailed recommendations include the principal recommendation that three simple models for value added analyses should be developed, '*the essential feature of these models [being] that they provide a basis for making like-with-like comparisons*'. The summary concludes (p.7) that:

the use of value-added indicators in developing policies for school improvement should be encouraged on a national basis by developing a culture of openness in connection with performance indicators and methods for school improvement, by using the inspection process to encourage schools to pursue further the use of quantitative performance indicators, and by making available to all schools, via the DFE, OFSTED and SCAA, information

concerning methods of analysing performance indicators which have been found to be helpful.

One of the tasks of the Secretary of State during her first six months in office was to welcome the report; in so doing she demonstrated the 180 degree turn-around in Department policy (GB.DFE, 1994):

I am firmly committed to the development of robust national measures of the value-added by schools to children's education... It is imperative to ensure that any measures of value-added are straightforward and are intelligible to parents... I commission SCAA to take this work forward...

An important feature of her response was to reiterate Dearing's acknowledgement of the useful work – much of which had been funded by local education authorities and other influential bodies – already being done outside national government circles:

I also welcome the impetus which the report gives to the use of value-added measures by schools and LEAs as indicators of their own effectiveness. Some useful work of this kind is, I know, already going on... The wider use of such indicators will be facilitated if SCAA, OFSTED and the Department are as open as possible about the development of work in this area... I shall expect the Authority for its part to have regard to the approaches being tried out locally in taking forward its work. (*ibid.*)

It then took little time for a national value added project, supported and funded by SCAA, to be initiated and a contract was set out in early 1995 to run from March of that year to December 1996. The objective of the project was '*to advise the Secretary of State on the development of a national system of value added reporting for schools based on prior attainment, which will be statistically valid and readily understood.*' (Fitz-Gibbon 1995, p.31.) The focus was to be on results at Key Stages 2 and 4, that is, broadly speaking, the outcomes from primary and from secondary schooling respectively. The contract was awarded to the Curriculum, Evaluation and Management Centre, then housed at Newcastle University and later relocated to Durham University. A series of technical and overview reports was produced, and the final report appeared in 1997 (see Fitz-Gibbon and Tymms, 1995; Fitz-Gibbon, 1997; Tymms, 1997a and b; Vincent, 1997). The chart overleaf represents some of the key 'moments' in the evolution of the national value added system.

The change of administration from Conservative to New Labour in May 1997 made no difference to the aims and momentum of the project; the policy transition was seamless. As the Qualifications and Curriculum Authority put it, *'the White Paper Excellence in Schools [GB.HoC 1997] stressed the need to provide parents and others with information about the progress pupils make relative to other pupils. These are usually referred to as "value added" measures.'* (QCA 1998, p.1). At the time of writing (1998-99), the Standards and Effectiveness Unit at the Department for Education and Employment and the Office for Standards in Education, as well as the Qualifications and Curriculum Authority, are taking a lead in the production of national value added and other contextual data on school and local education authority (LEA) performance. In part, this is intended to aid the setting of 'challenging but realistic targets' for schools and LEAs (see DfEE Circular 11/98). One might even argue that it is this statutory requirement to set targets which will bring value added analyses closer to the everyday life of schools and teachers.

#### **EVOLUTION OF THE NATIONAL VALUE ADDED SYSTEM**

- 1993 Sir Ron Dearing's interim report on the National Curriculum and its Assessment recommends that work be commissioned into value added performance indicators for schools
- 1994 SCAA working party, chaired by John Marks, reports to Secretary of State with 25 recommendations for taking value added measurements forward based on public examination and National Curriculum assessments
- 1995 Contract for a feasibility study of the procedures necessary to set up a national Value Added System awarded to the Curriculum, Evaluation and Management Centre, University of Durham: statistical studies on available datasets to be completed by September 1995 and two new pilot projects – KS1-2 and KS3-4 – to run from September 1995 to December 1996
- 1997 Series of project reports published, advising that *'it is possible and desirable to set up a national system to provide schools with value added indicators of their performance'* and providing detailed findings from the project
- 1997-8 Consultations take place with schools and other interested bodies on the best ways of calculating and presenting value added indicators for schools and value added measures in school performance tables
- 1998 The first national value added analyses, covering KS1-2, KS2-3 and KS3-4, published in autumn term in the form of i) lines representing the achievement of pupils at the median and quartiles of the distribution of outcomes for each possible input score, and ii) of chances graphs showing the percentages of pupils achieving different outcomes from the same starting points

Secondary school performance tables contain 'progress measures' based on aggregate data (as distinct from true value added, which is based on individual pupil-level data)

2000	Secondary school performance tables to include value added measures
2002	Primary school performance tables to include value added measures

As for Shephard's wish that SCAA should take due note of developments already happening, it is noteworthy that SCAA's successor, the Qualifications and Curriculum Authority, set up a series of meetings with value added service providers – of which the NFER was one – at which updates on the progress of the national value added system were given, opinions taken from those already involved in value added and problems frankly aired.

One could legitimately ask at this point whether there is any evidence that the DfEE made efforts to draw on the experiences of other countries. Part of the problem was that very few education administrations had established statutory value added systems at the time the English system was being developed. However, a working party from the DfEE made a visit to the USA to find out about the implementation of a statutory value added system that had been set up in one state to promote accountability, especially of individual teachers.

A confidential report (not referenced) on the visit has been made available to the author, in which amongst other things the working party notes that the system had become a highly politicised issue '*so much so that public professional debate on its merits or otherwise was difficult*'. Furthermore:

public understanding of the data produced by the system was poor ... Teacher support had not been secured. Indeed many teachers felt that [the system] was a high stakes exercise (as a legal requirement) adding to pressure on them. Many were also nervous about... a system which did not include qualitative as well as quantitative data, and did not refer in any way to socio-economic circumstances. Despite these views, teachers were keen to receive training in using value added data...

The report identified a number of issues relevant to the development of policy and practice in England; these included the need to maintain the distinction between accountability and school management in terms of what is collected and reported; the importance of presenting data clearly to parents and teachers whilst ensuring its technical soundness; the need to provide systematic teacher training and guidance, 'crucial to successful implementation'; and the need to keep focused on absolute standards as well as relative effectiveness. It is not known to the author how widely the report was circulated and how much

influence it actually had within the DfEE; one could argue that the value added system we have now has attempted to go at least some way to address all four issues, though without satisfying some of the main criticisms made (see Section 4.4 below). At the same time it is highly likely that political considerations – such as the need to establish the promised national value added system as soon as possible whilst keeping costs down – held sway; for example, the provision of training has so far been limited to written guidance as part of the ‘Autumn package’.

Other governments may be trying to learn some lessons from the English experience, however: an international e-mail group discussing ethical as well as technical issues involved in value added was initiated by Dutch evaluators in spring 1999; they had been commissioned by the Dutch education ministry to explore how to set up a valid and acceptable national system. Various academics based in institutions in England, France and the Netherlands took part in a very detailed and wide-ranging discussion, which culminated in a set of recommendations being made to the Dutch government (see Visscher *et al.*, 2000).

To return to the English developments: in the period between the policy declaration in favour of a national system and its implementation, various other reports on value added prepared for or on behalf of a policy interest appeared, including the following:

- GCSE to GCE A/AS Value Added: Briefing for Schools and Colleges (GB.DFE, 1995a)
- Value Added in Education: A Briefing Paper from the Department for Education (GB.DFE, 1995b)
- Value Added and School Improvement (Society of Education Officers, 1995)
- Adding Value and Improving Performance (Elkin, 1995, for the City Technology Colleges Trust)
- Value Added Measures of School GCSE Performance (Jesson, 1996, for the DfEE).

Undoubtedly these publications helped to establish and promote the debate amongst a wider public concerned with education than would normally have access to academic journals and technical papers. The reports written by Jesson and the Society of Education Officers (SEO) are particularly clear and helpful, in taking on some difficult technical and statistical issues but at the same time always relating these to the educational purpose(s) of attempting to measure value added. As well as discussions of the relative merits of regression and hierarchical (i.e. multilevel) models, and of prior attainment and socio-economic data as ‘predictors’ of performance, for example, the SEO report contains sections concerned with what might be called value added in practice: for example, ‘making sense of the outcomes’, ‘the link with school improvement’, ‘confidentiality versus sharing the outcomes’, ‘the role of the education authority’ and ‘the case for local value added initiatives’. It concludes with the observation that:

For most education officers and schools, the analysis is not an end in itself. Much work remains to be done to develop the insight it provides. Collaboration of this kind [a partnership between LEAs and schools, individually and collectively] probably offers the best prospect of school improvement currently available in the service of all children and their parents. (Society of Education Officers 1995, pp.38-9)

Of course, one might say, this is a predictable position for the report’s authors to take, but the general proposition remains true, as is argued in Chapter Five of this thesis. Some of the other material, by contrast, served to perpetuate unhelpful ambiguities and misunderstandings. I would want to take issue with several of the definitions of value added given in these various documents, for example, as being unenlightening or inaccurate or both; but since this review began with the argument that the term was dependent on ‘functional ambiguities’, it would be tedious to labour the point with further examples.

On more specific points of contention, Elkin, in her section discussing factors other than prior attainment which may affect performance (1995, pp.30-1), apparently does not distinguish between those over which schools have little or no control (such as percentage of students with a statement of special educational need) and those which are directly the result of school policy and management (such as amount of homework). This suggests that she has not understood the ‘eliminative’ nature of the conceptual model used by all the

quantitative approaches she comments on. Furthermore, in discussing socio-economic factors, Elkin claims that:

the idea of “making allowances” is a seductive one. Yet... it is only too easy to say that because “our” school has large numbers of students from low income families, or whatever, that the pupils cannot be expected to achieve much. (p.32)

She thus makes the common mistake (and one which perhaps has an ideological component) of confusing ‘making allowances for’ with ‘allowing or controlling for (certain specified and empirically established factors)’. (Both of these general issues were discussed in Chapter Three above.)

More significantly for the development of a national system, Fitz-Gibbon (1997, p.48) seems, on the face of it, to adopt a similar position in writing up the findings of the national value added project. No surprise then to find a comparable confusion in the DFE Briefing Paper on value added (GB.DFE 1995b), the main purpose of which was to give some background, from the Department’s perspective, to the national value added project and to describe briefly the work in hand. The paper points out (p.2) the ‘*major technical difficulties in assessing the scale of any such effect [of adverse socio-economic factors]*’ and goes on to claim that ‘*adjustments to value added measures on these grounds which were not sufficiently rigorous could justify poor performance or legitimise low expectations.*’ Again, this appears to have some political rhetoric at the back of it. Adjustments for socio-economic factors should indeed be based on rigorous and empirically replicable data; but problems of non-rigorous interpretation may still remain. Legitimising poor performance is a failure of understanding, interpretation and, ultimately, imagination rather than a result of poor data. One way round this issue is to present school-level residuals at three stages of adjustment: unadjusted, adjusted for prior attainment and adjusted for prior attainment together with socio-economic context, in order that insights can be obtained from examining the *differences* between a school’s results at successive stages of analysis (as does the NFER’s QUASE service). This was also discussed in Chapter Three above.

The DFE paper also, rightly, discusses how to arrive at the right balance between paying attention to ‘technical niceties’ and enabling ‘intuitively attractive’ summary measures to be derived. However, the ensuing recommendation (p.8) that ‘*summary measures, e.g. for use in comparative performance tables, need to be few in number, and must be consistent and fair between institutions... [whereas] value added measures for use by school and college management will typically be more disaggregated*’ seems either to duck the issue or else to imply that summary measures of value added which are to be used for publicly comparing schools can afford to forgo (some of – which?) the technical niceties. Whereas, on the basis of evidence presented in Chapter Three above, many academics think that the more ‘high stakes’ the value added analyses are – and public comparisons of institutional performance are pretty high stakes – the more crucial it is that the technical niceties are observed.

Naturally, much of the debate has been concerned with how to strike the right balance between what is ideal, analytically speaking, and what is feasible. The remarks of the SCAA working party report (SCAA, 1994) are worth quoting, even though they are hidden away in Appendix 9:

... decisions about which variables to include inevitably depend on views about what is important in the real-world situation we are attempting to model... Such decisions are therefore, almost inevitably, open to question and debate... It therefore needs to be recognised that there is no single correct method of analysing a complex social situation... It also needs to be recognised that all corrections and adjustments to primary data are approximate and that there will always be arguments about their validity. (p.85)

The paragraphs in the DFE document could be regarded as no more nor less than a gloss on what SCAA had argued, but something has been lost in the translation, as they say.

Interestingly, the SCAA report goes on immediately to say:

There is thus a strong case for using simple methods of analysis alongside complex ones and for presenting analyses in ways which retain as much contact as possible with the primary data... (*ibid.*)

– which could, without too much risk of over-interpretation, be read as a belated argument in support of raw data.

But both the DFE/DfEE and SCAA/QCA recognised that value added analyses may be put to a variety of different uses; a basic distinction can be drawn between accountability requirements and the contribution of performance data to school improvement. Both uses were apparently supported by a majority of respondents to the SCAA/QCA consultation exercises (SCAA, 1997a and b; QCA, 1998); although in the final project report (Fitz-Gibbon 1997, pp.25-9), the sample population of headteachers were found to be overwhelmingly in favour of value added data for internal use but very negative towards the idea of publication of such data. Are there methodological or technical implications of trying to put the data to different uses? The DfEE/DFE's position on this has been outlined on the previous page; according to QCA (1998, p.1), the distinction means only that *'the information and measures are both arrived at by the same method, but their purpose, although complementary, is very different.'*

Well, in one sense, the important thing is that schools and LEAs are no longer operating in a policy void as far as value added is concerned: the salient facts to bear in mind are, first, that one of the main objections to published performance tables as originally conceived, that they gave no indication of relative progress in differing pupil populations, has been met at least in principle. Secondly, all schools, whether or not they choose to supplement national data with other sources of value added analysis, will have access to quantitative performance analyses which could in theory be a substantial contribution to management development, institutional review and target-setting.

Why, then, qualify these desirable and desiderated moves with phrases like 'in theory' and 'in principle'? At this point, it may be helpful to see to what extent the policy requirements and intentions can be reconciled with rigorous academic criteria of the kind discussed in the previous chapter of this thesis.

#### 4.4 Reconciling Academic Criteria with Policy Requirements

In constructing a brief for the national value added project, SCAA specified that a national system would need to be:

- readily understandable;
- statistically valid;
- not an undue burden on schools;
- cost-effective.

(Fitz-Gibbon, 1995, p.2)

More specifically, it is argued (QCA, 1998, p.1) that value added measures included in published tables of performance need to be:

- applied consistently across all schools;
- based on robust statistical methods;
- sufficiently straightforward for parents and others to understand, and for schools to check.

As policy requirements, these are perfectly legitimate, one could say minimal. To what extent are they capable of being fulfilled? One problem may be, given the evidence marshalled in Chapter Three from the academic literature, that it will not be possible to have value added measures which are both statistically valid and readily understandable (*pace* Fitz-Gibbon, 1997, p.14: even with her '*very simple and accessible models*', one wonders with what sort of sample of ordinary people the graphs and tables were trialled), and also give better insights into schools' performance than raw scores.

In Chapter Three above, a general consensus was noted amongst writers in the school effectiveness field that measuring value added with accuracy depended on the following criteria:

- using individual pupil data;
- taking account of prior attainment;
- taking account of other pupil background factors (gender, ethnicity, SES);
- including outcome measures which reflect the whole range of achievement;
- analysing the data with a multilevel model.

Criteria on which there is not a full consensus but for which a cogent case can nonetheless be made included:

- the need to take account of school context (as well as individual pupil level) factors;
- the need to use finely differentiated input and outcome measures, not merely broad bands or levels of achievement;
- the need to avoid ‘high stakes’ measures as input variables; a good example of this would be using the publicly available outcomes from a previous stage of education such as National Curriculum test scores as the measure of prior attainment or ‘input’ for the next stage.

These criteria are concerned with questions of accuracy and robustness – although problems of cost effectiveness and comprehensibility for a lay audience have not been entirely ignored in academic literature. (The extent to which widespread comprehensibility should continue to be a preoccupation may be a moot point, although it has superficial appeal. How far parents use performance tables as the determining criterion in choosing a school has yet to be definitively established, but the exercise of parental choice seems to be dependent on both parental class and local circumstance.) A reading of the SCAA/QCA reports on the national value added project (Fitz-Gibbon and Tymms, 1995; Fitz-Gibbon, 1997; Tymms, 1997a and b; Vincent, 1997) shows that they contain as thorough a presentation of methodological and technical problems as could reasonably be expected of a principal agency of a government committed to introducing a national system of value added as soon as possible. In the final report, there is discussion of, *inter alia*:

- the need to base calculations on individual pupil data;
- the need to have more finely differentiated performance data than that provided by assessments according to national curriculum levels;
- the need to use three year rolling averages;
- whether and how to account for factors like student mobility and attendance;
- the relevance of socio-economic factors;
- relative merits of different statistical models;
- presentation of results;
- questions of quality control and data-checking by schools;
- training and support for school staff in understanding value added data;
- the potential contribution of value added data to school improvement;
- the possibly harmful effects of a national system.

This last issue is discussed in the report's penultimate and thought-provoking chapter (pp.87-93), about which not a great deal has been said (aside from Kellner, 1997). Fitz-Gibbon quotes Smith (1995) in identifying unintended and presumably unwelcome behavioural consequences of publishing performance data. '*Eight problems associated with non-effective or counter-productive systems*' are listed, as follows:

- tunnel vision
- sub-optimisation
- myopia
- measure fixation
- misinterpretation
- misrepresentation
- gaming
- ossification.

Fitz-Gibbon comments on this list:

With the exception of ossification every one of these possibilities was commented on by headteachers in open-ended items in the questionnaires. These are not theoretical but actual, already-perceived problems.

As was discussed in Chapter Three, a reputable school of thought is now arguing that value added measurements – far from making the public reporting of comparative performance more accurate and informative – underline the fact that performance tables are impossible to construct with any validity. Fitz-Gibbon's critique strongly suggests that the usefulness of published value added measurements '*in attempts to improve efficiency and equity*' may also be compromised. This is the nub of a public relations problem for New Labour: a commitment to continuing the Conservative policy of publishing performance data has led inexorably to their trying – like the Conservatives – to make that data fairer and fuller, and value added measures offer the most plausible approach. The alternative, of ceasing to produce performance tables, seems unthinkable at this point. An independent evaluation of the effects over time on schools and pupils of the national system is surely called for. In the mean time, the evidence in Chapter Three would appear to lead to the view that value added data in the public domain is most useful as a screening device

for identifying those relatively few schools which are ‘outliers’ in statistical terms, rather than for providing definitive assessments of all schools.

What of the use of value added for institutional review and school improvement? Well, some similar strictures must apply; that is to say, if it is true that value added measures:

- are only as good as the data they are based on;
- deal in correlations, not causes;
- contain an irreducible degree of statistical uncertainty;
- are based on a normative and retrospective model which may not provide any helpful information about levels of performance in future;

then it needs to be said that value added measures function as only one instrument of evaluation. Even so, the information on relative performance provided by value added analyses is more robust than that characteristically given by other methods in the past. A commonly-heard phrase in schools these days is: ‘*There’s no hiding place now*’ (which can be taken to mean something like: teachers need to demonstrate rather than merely claim that their pupils’ performance is as good as it could be). Most importantly, explanations of poor performance which rely on assertions or assumptions about the inadequacy of the pupils can more readily be questioned. Value added measures make it a great deal easier to identify the root causes, as well as the underlying trends and patterns, of under-achievement.

So it seems reasonable to make the modest claim that, given the right culture and conditions, value added can help to pose better and more focused questions about the way a school or LEA has performed with its pupils and to stimulate more informed discussion amongst school staff about the way they organise and deliver their teaching. How far ‘the right culture and conditions’ exist, what needs to be done to nurture them and whether the introduction of the national value added system will help their growth are key questions to which national policy makers will need to seek answers. The following chapters in this thesis discuss ways in which some aspects of these questions can be explored.

## ***CHAPTER FIVE***

### **GETTING BEYOND THE NUMBERS GAME: THE EVIDENCE ON HOW SCHOOLS USE VALUE ADDED DATA**

#### **5.1 Introduction**

Previous chapters have discussed the development of value added measures of performance in methodological and analytical terms, as well as the growth of their political acceptability to the point where, in the autumn of 1998, all schools in the primary and secondary sectors received a package of materials from the DfEE containing ‘value added’ information which purported to help them analyse their own performance and progress. Various kinds of analyses are also being generated by schools themselves or LEAs as well as the DfEE and QCA and there is now an unprecedented amount of assessment and performance information circulating in the system.

It might therefore be excusable to assume that school and LEA staff have acquired along the way an understanding of how to use such data in a valid manner. There is no doubt that the tightening-up of the school improvement agenda by the DfEE’s Standards and Effectiveness Unit has focused attention on the respective roles and responsibilities of LEA officers and advisers, school managers, governors and teachers in a new and unprecedented way for using sophisticated quantitative data. But the material discussed earlier in this thesis has shown just how challenging the technical and conceptual issues are in relation to this kind of data; I hope it has also demonstrated how unhelpful it is to conflate the accountability and development aims associated with the data.

Moreover, it is far from clear that the user end of the process, so to speak – of how exactly school staff do interpret, manage and utilise such data – has been sufficiently well researched and documented. The remainder of this chapter accordingly presents and discusses the evidence base in this area and identifies the key issues for further research.

## 5.2 Schools' Use of Value Added Data: the Relationship between 'School Effectiveness' and 'School Improvement'?

The fact that the evidence base is so far quite modest is probably due to a number of factors. First, it needs to be remembered that of course it is only quite recently that the majority of schools had access to anything that could be called value added data. Secondly, many people – including policy-makers – may not have thought there was a problem which warranted investigation. The idea seems more or less to have been: the data shows such-and-such about your school, so you now have the information about what you need to do to raise standards. Or, to put it another way, making school effectiveness data available is what is needed to bring about school improvement.

In exploring why we may not simply assume that value added data will work this way, there is a need to distinguish carefully between school effectiveness and school improvement research and their respective outcomes and perspectives (which many non-academics assume are more or less identical). 'School effectiveness' is now very well developed as a quasi-scientific body of research – that is, one which attempts to build up a coherent knowledge base through the replication of rigorously-designed sequential studies. These are quantitative and methodologically complex; they typically require the application of statistical techniques such as multilevel modelling to large datasets acquired through longitudinal data collection. Value added analysis is a variant of such studies (see Chapter Three above for a fuller discussion of this point).

By contrast, the field of 'school improvement' continues to be characterised by a wide range of qualitative studies which have not, on the whole, been intended to replicate each other but rather to explicate the highly context-specific nature of 'improvement'. Such studies have contributed to *inter alia*:

- the conceptualisation of the *management of change*, plus an understanding of the role and function of 'change agents' in education (see, e.g., Fullan, 1992, 1993)
- frameworks and models for understanding the actualities of improvement as a *process* at system, school and classroom levels: (see, e.g., Hopkins *et al.*, 1997b)
- insights into the different *experiences, views and needs of the key players*: principals/headteachers, teachers, pupils, parents (see, e.g., MacBeath *et al.*, 1996; Rudduck, 1996; Rudduck *et al.*, 1996; MacGilchrist *et al.*, 1997; MacBeath, 1998)

- in-depth descriptive analyses of in-school *cultures and power relationships* (see, e.g., Ball, 1987; Hargreaves, 1994; Hargreaves, 1995; MacGilchrist *et al.*, 1997; Stoll, 1998; Bishop and Mulford, 1999; Freiberg, 1999)
- *evaluations* of individual improvement initiatives (see, e.g., Stoll and Fink, 1992; Myers, 1995; Hopkins *et al.*, 1997a; Slavin *et al.*, 1992).

On the face of it, these kinds of understandings could be very important to the study of how value added analyses are dealt with once they reach schools – especially if the ostensible purpose of the data is to assist school improvement. However, it is only a slight overstatement to say that the broad tendency in school improvement research has been to assume that performance data is of limited practical interest. And although several books published in the last few years have attempted to integrate the school effectiveness and school improvement traditions – see, for example, Gray *et al.* 1996a; Reynolds *et al.* 1996; Barber and White 1997; Gray *et al.* 1999 – this has chiefly been from the point of view of using effectiveness data to test or ‘prove’ improvement strategies. Few if any research studies in school improvement have focused specifically on the uses of performance data as an aspect of school practice.

### **5.3 The Evidence on Schools’ Use of Value Added Data**

This being so, what kinds of evidence do we have on the ways in which school staff use value added data and what does it show?

In an article about teachers’ use of quantitative data (Dudley, 1999b, discussed in greater detail in the final chapters of this thesis), an LEA senior adviser saw value added data in these unambiguously positive terms:

[Value added data] is, undeniably, uniquely powerful in the way teachers and managers are prepared to trust the data and act on them. Value added is therefore valid in a number of ways. First, because people are prepared to act on the basis of the data, it... has a strong ‘validity of “consequence”’ [Messick, 1989]. Second, in measuring pupil progress rather than merely learning outcomes, value added data... has a strong ‘construct validity’ [Messick, 1989]... I would argue that value added data enjoy this validity and currency almost entirely because of such psychological *affective* factors as trust and perceived fairness. (Dudley, 1999b, p. 113.)

However, we perhaps need to be aware that not all school staff may have been able to acquire an accurate understanding of value added. Maw (1999) argues that the press (including the *Times Educational Supplement*) coverage of 'value added' has largely been in the context of publishing the controversial tables of school performance. Moreover, the coverage of key issues concerning the measurement and presentation of value added has not given a clear and coherent message: 'there appears to be some confusion in the press' is Maw's conclusion from the collection of contradictory or ambiguous statements she quotes.

Furthermore, as we saw from the previous chapter, Fitz-Gibbon (1997) gave some warnings in the last chapter of the final report on the national value added project about the perverse ways in which value added measures might be used and the negative consequences which would follow. These warnings were based on situations and circumstances reported to her from the field during the piloting of the national system – which themselves perhaps partly arose from the lack of clear messages noted above, as well as from the politicised context in which the value added agenda was being developed.

Over the last ten years, there have been a number of published studies relating to schools' use of value added data *via* a number of 'services' – including ALIS (University of Durham), the Hampshire project (University of London Institute of Education) and QUASE (National Foundation for Educational Research) – from which some empirical evidence can be garnered.

Data from ALIS (A-level information system) was available to schools over a decade ago, and the providers undertook to find out something about how schools were making use of what was a relatively novel system. *'The picture that emerged was of a very peripheral level of awareness, a tendency not to attend meetings, to read reports sparingly and to take no action on reports once read.'* (Williamson and Fitz-Gibbon, 1990, p. 40.) A little later, the picture was reported to have changed for the better (Williamson *et al.*, 1992), but the following points emerged from interviews conducted in eight schools:

- levels of awareness of ALIS data varied considerably in these schools, with managers reporting a higher degree of engagement than heads of English departments;
- for several heads of English interviewed, the impact on practice had been minimal;
- the statistical nature of the data was a serious obstacle;
- there was a general lack of within-school systems to ensure that the data was acted upon.

Evidence from interviews was supplemented by questionnaire survey data, although the response rate was low (between one-quarter and one-third). The responses indicated that teachers of mathematics were more likely than their English colleagues to value ALIS. Other information collected for the study – although rather limited because of small samples and/or low response rates – suggested that in-service training provision related to the data had some impact on teachers’ use of ALIS to inform their teaching; and also that short summary reports were more valued by English teachers than the normal full-length detailed reports. The study concluded that, ideally, each school needed to establish an ALIS coordinator who would be externally trained so as to acquire ‘a full understanding of the system’ and given authority and dedicated time in order to ‘disseminate findings in a user-friendly way’. ALIS is, of course, a particular kind of system, established at a time when the use of performance indicators for monitoring and feedback was at an early stage of development; it would be risky to assume that these findings would hold good in a different context. Nonetheless, they plainly suggest that the active use of data supplied to schools, even when they are paying for the service, is not something which can be taken for granted.

Over the years since that study, a few other pieces of evidence of different kinds have emerged. We can again note the strictures highlighted by Fitz-Gibbon (1997) in the final report on the national value added project (discussed in the previous chapter). One may assume that the national study commissioned by QCA in autumn 1997 to examine schools’ use of value added data was a specific response to these warnings. However, the outcomes from that study have not, at the time of writing, been published.

In one of the few studies discussing longitudinal work undertaken with primary (rather than secondary) schools – in Hampshire LEA (Yang *et al.*, 1999<sup>7</sup>) – the report’s authors say that *‘the research sought ways of communicating the results to head teachers in ways which were meaningful without destroying the underlying complexity of the relationships uncovered... [and] in ways which can assist the process of school improvement’* (p. 469). The paper cites the three criteria used to support these requirements (pp. 476-7):

- the data is confidential to the school and LEA;
- the system is an evolving one, taking account of feedback from users and of new national developments;
- value added estimates are understood as having uncertainty attached to them.

The authors say that a programme of in-service training has been implemented for headteachers and advisers *‘to familiarise users with interpretations, including the limitations of these results’* (p. 479). Additionally, it is reported that there will be an evaluation of the use of the results. So, although the paper states that *‘it has been shown that complex multilevel models can be presented in ways that are accessible to potential users without sacrificing the essential components which emerge from the analyses’* (p. 479), a descriptive assessment of the actual usage of the results is not yet available.

As it happens, much of the remaining evidence available for this present study on the use schools make of value added data is linked directly with QUASE (Quantitative Analysis for Self-Evaluation), the NFER’s value added service for secondary schools<sup>8</sup>. This latter evidence is derived from:

- a published research and development study undertaken in the early ’90s study with a consortium of secondary schools (Harris *et al.*, 1997; Wikeley, 1998);
- information supplied by schools subscribing to QUASE in their annual feedback questionnaires;
- a published case-study of one school’s use of data, including QUASE, for raising pupils’ attainment (West and Moore, 1998);

---

<sup>7</sup> Interestingly, one of the article’s authors is Goldstein, who has argued in previous papers (e.g. Goldstein and Thomas, 1995) against the use of value added data in such a context.

<sup>8</sup> For a description of how this value added service operates and the kinds of analysis it provides, see Appendix A.

- the first-hand experience – through providing seminars and in-service training – of the author.

### **5.3.1 Research and development study**

The first of these sources comprises findings from a research and development study involving a particular cluster of urban secondary schools which were trying to find ways of working together on generating and using value added data. It was undertaken at a time when competition between schools for pupils and the published performance tables were contributing to the vulnerability of schools which could not show high standards on ‘raw’ scores. The project was intentionally building on, and using the final phase of funding from, the TVEI (Technical and Vocational Initiative) Extension, one of whose objectives was to strengthen consortium working. The project used QUASE data, first, to identify whether and where there was ‘added value’ in schools in the cluster; secondly, in conjunction with a team from a local university, to link this quantitative data with qualitative evidence gathered at departmental level, in order to discover whether there were any common organisational and/or pedagogical characteristics in departments identified as adding value which could be used to assist other departments.

According to Harris *et al.* (1997), a range of such characteristics were found, as follows:

- Good organisation/management, from schemes of work to planning and utilisation of resources;
- Systematic scrutiny of exam and test results;
- Sense of vision (especially of subject discipline) conveyed by department head;
- Change and innovation accepted, if in accordance with departmental vision;
- Effective internal communication within department;
- Delegation of tasks to colleagues (based on sense of professional trust);
- Active involvement of pupils in review of and reflection upon their own work;
- Consistent approach to assessment which motivates, reinforces and builds confidence;
- Clear structure for lessons which also integrates feedback to/from students;
- Consistent homework policy;
- Use of reward, in preference to punishment, to modify behaviour;

- Normal range of experience, capability, motivation, etc., amongst departmental staff, but within a context of low staff turnover.

These are interesting findings in their own right, because it is at the departmental level where the evidence from school effectiveness studies suggests a great deal of intra-school difference lies. The thesis will return to this issue in later chapters: the point to be made here is that the project generated local qualitative data which could be used alongside the quantitative data.

Despite the necessary limitations of the study (for example, no work was undertaken with departments *not* adding value, to test the validity of the above list), it was intended that the quantitative and qualitative findings should together be used as an aid to self-evaluation and internal improvement within the cluster of schools participating in the study. This phase of the work was also assisted by an external researcher, who wrote up the findings (Wikeley, 1998). Her study revealed some intriguing problems in disseminating and sharing so-called ‘effectiveness’:

- staff felt the dissemination of the ‘effectiveness’ findings had been imposed from above by their senior management team/an external agency;
- staff in departments which had ‘added value’ felt that being identified in this way was divisive;
- some claimed that the research ‘told us nothing new’;
- senior management’s perceptions about effective departments seemed to differ from other staff’s;
- there was a tendency to distrust external quantitative data as opposed to instinctive judgement; people tended to be much more critical of the basis for the statistical analysis than they were of the basis for so-called professional judgement;
- there was, on the other hand, a lack of other robust kinds of data to confirm or refute the effectiveness findings;
- there was a lack of follow-up to the work *via* self-evaluation in departments; staff tended to focus on the need for further resources instead.

These problems give some important clues that, in investigating what happens to value added data when it is introduced into schools, we need to pay attention to how data is perceived and accommodated by managers and by teachers; so Wikeley’s findings can be seen as setting at least part of the agenda for further empirical research. (They are also relevant to the

dissemination and use of research-based evidence beyond the particular example of 'value added' data.)

### **5.3.2 Feedback from QUASE users**

In some contrast with this 'real-life' view is the information provided by QUASE schools in their feedback questionnaires on the service provided by NFER; these were issued each year to users of the service when they received their annual school Profile. After questions relating to the clarity, usefulness and value for money of the analyses, questions were asked about the uses to which QUASE data were being put in the school, as follows:

- Do you propose to use the QUASE Profile in helping you to plan changes within the school?
- If you are planning to use QUASE in this way, what mechanisms are you using?
- Do you think the QUASE results will contribute to positive changes within the school?

Management reports for the internal steering group were compiled for each of the first four years of the service on the basis of the school returns, from which the following aggregate information has been extracted. Response rates to the proforma never exceeded around 66 per cent and sometimes no more than a third of participating schools replied, so the evidence must be treated with caution. Even so, according to the returns, under half the respondents in the first year or two planned to use QUASE to plan changes in the school, compared with over three-quarters in later years. The most frequently cited purpose was school development planning, but over time departmental reviewing and planning, and target-setting, became equally cited. The likelihood of QUASE contributing to positive changes was noted by half of correspondents in the early years; the proportion rose to four-fifths in later years. (The schools participating, and *a fortiori* the schools responding, were not a stable population; thus the changes over time noted do not necessarily represent changes in practice within the same schools.)

This information, in the absence of any evidence collected *in situ*, would suggest that the use of QUASE analyses became increasingly well embedded in general management practice in schools.

### 5.3.3 Case study of one school

The most recent piece of published evidence on QUASE the researcher has encountered is a written case study of a secondary school which has used QUASE data (West and Moore, 1998). The case study was compiled independently of the NFER and the author of this thesis; but it would appear to confirm, and elaborate on, the impressions collated from the feedback questionnaires noted immediately above. The authors claim several benefits arising from the school's use of value added data and the developments to which this has led in respect of monitoring pupils' performance and progress. These can be summarised as follows:

- Staff attitudes have changed from a '*minimalist understanding*' and even a feeling of unease to a situation where '*measurement data is seen more as a tool to aid increased pupil achievement.*'
- '*There is more openness surrounding the audience for the data and far more discussion about the issues that the data highlights. Many staff now use data to inform their curriculum planning at both departmental and classroom level.*' Specific developments have been introduced to enhance the quality of learning.
- Pupils have benefited from a growing self-awareness about their progress (or lack of it).
- '*The most valuable outcome, paradoxically, probably belongs to that aspect of school effectiveness that remains elusive, defying quantification. It is to be found in the increased levels of awareness of the teachers, the day-to-day adjustments, the small experiments with new approaches, the time devoted to reflection or just thinking about the relationships between teaching and learning.*'

It has to be said, however, that much of this paper is written from an implicitly management perspective. Some of the claims might be seen differently by other staff than senior managers, of which the following is a good illustration when contrasted with the Wikeley findings above:

Some curriculum areas/departments are now identified as better performing than others and this is known throughout the school. This could be interpreted as creating a difficulty, but it is rather seen as an opportunity for differentiated support to be provided for those departments/areas that need it. This enables the effective targeting of resources for improvement (West and Moore, 1998).

On the face of it, this sounds rather more like a management development aim than a description of institutional reality, and reinforces the idea that we should try to hear more than one side of the story.

#### **5.3.4 Anecdotal experience of the author**

At this point, it is worth making it clear that, prior to the present study, my professional work had already led me to have strong convictions about the potential uses of value added for school improvement; indeed, much of my work in the 1990s aimed to set up the conditions in which productive conversations about value added could happen in schools and LEAs<sup>9</sup>. I would characterise this work as the ‘development’ dimension of research-and-development. The reason for undertaking the present study was that, by the middle of the decade, I had reached a point where I wished to address the ‘research’ dimension in more depth and detail than my normal professional work permitted. From my own experience, I thought it might well be the case that many schools no longer evince a lack of belief in externally-generated data. Participants’ responses at the various seminars I have given for senior managers suggest that taking a ‘lies, damned lies and statistics – I trust my gut instincts’ view of performance data is no longer the default option. Many school and LEA senior managers are or have become highly numerate and some are well aware of what different kinds of value added analyses can tell them (and, just as importantly, what they cannot reveal). On the other hand, my experience also suggests that there may still be many in the profession, at management as well as classroom level, for whom such data presents difficulties of acceptance, understanding and/or utilisation. Thus the purpose of the thesis is to provide a more detached investigation of an area with which I have already become deeply familiar, and in which I readily acknowledge I have a vested interest.

#### **5.4 Issues for Further Research**

Taking all the above material into consideration leads me to the conclusion that we do not yet possess a coherent body of evidence about the uses of value added data. First, we still need a better description of how schools – or rather the various staff within them – are actually using value added data and for what purposes. We also need a better grasp of what the difficulties are in

---

<sup>9</sup> My dissemination and in-service activities in the value added field over the period 1991-1999 were considerable, and included numerous school- and LEA-based workshops, several regional and national conferences and some international seminars.

using value added data and where they arise from. And we need an assessment of the outcomes of using such data, both intended and unintended, for managers, staff and pupils. As Brown *et al.* (1995) argue, the notion of a top-down, management-oriented approach to school improvement based on school effectiveness data may be '*simple-minded*' if it fails to address the question of how the findings are, in real life, to be integrated into the thinking and practice of those who are seen as needing to improve. Furthermore, if such findings are exclusively based on policy-makers' and/or researchers' own constructs, they may be insufficiently valid and/or replicable.

Any empirical exploration should, of course, be informed by considerations of which the existing evidence – or lack of it – suggests we should take account. The issues identified by Wikeley strongly indicate, for example, that there are acculturation processes which schools and teachers need to go through before external data can be accommodated; and even then data may not be acceptable to some teachers for a variety of reasons. Lack of attention to such issues has, in my view, restricted the usefulness of recent field-based work commissioned by DfEE on schools' use of data for target-setting and raising standards.

My experience also suggests that there may also be ethical questions to address in how value added data is used. In some cases, senior managers have talked about value added data in terms that suggest it may be used as a lever for change in questionable ways; other staff then pick up on these ideas and feel concerned or even intimidated by them. As I reported a while back (Saunders, 1997), the 'military metaphor' is not uncommon in managers' conversations about value added analyses: whilst we should not over-play the importance of this kind of discourse, we should also not ignore the fact that there is often talk of data being used as '*ammunition*', or a '*weapon*', or a '*bomb to put under so-and-so*'.

That these areas of ethos and values, as well as of technical knowledge, are of concern to other researchers in the value added field has been demonstrated by the continuing interest shown in a working group whose initial meeting was hosted by NFER in early 1998. Key figures in value added and school effectiveness research are exchanging views for the purpose of discussing these wider issues, including possible protocols for interpreting and using value added: a joint article was published as a result of the first two meetings

(Saunders and Thomas, 1998). One of the contributors has been Professor Harvey Goldstein who, with another colleague active in the school improvement field, has written a paper outlining a possible code of ethics for performance indicators more generally (Goldstein and Myers, 1996).

One might argue that the political and ethical questions are even more important to be addressed wherever value added data is being used for school review and improvement than where it is produced for external accountability<sup>10</sup>. Because ultimately – given the enormous national investment in assessment and performance data – the challenge to those generating value added analyses must be: Is the effort worthwhile? Do value-added approaches help to improve education in practice? Under what conditions and with what prerequisites? How do we know?

All of this is crying out to be explored empirically, and the best way would seem to be by building up a systematic evidence base, through qualitative fieldwork mediated by rigorous conceptual frameworks. (As was noted earlier on in this chapter, much has recently been written by leading academics about the failure of school effectiveness studies to relate directly to the ways in which schools can, in practice, change and improve – and about the need to undertake better-specified research which will close the gap. See, for example, Reynolds *et al.* 1993; Brown *et al.* 1995; Reynolds *et al.* 1996; Gray *et al.* 1996b; Barber and White 1997; Lauder *et al.* 1998; Mortimore 1998.) Even from a policy point of view it must be important to get beneath the management rhetoric, not in order to put a negative slant on headteachers' intentions nor to undermine the principle that quantitative data has a role to play in informing and improving pupils' progress, but in order to find out how best to understand and address any difficulties which may arise in the process.

The fieldwork exercise described in the following chapter was designed to contribute to this empirical investigation.

---

<sup>10</sup> However, this seems to depend on national policy context. See Chapter Four above for a reference to the Dutch initiative exploring how to set up a valid and acceptable national system.

**CHAPTER SIX**  
**FRAMEWORK, METHODOLOGY AND PROTOCOL FOR**  
**CONDUCTING THE FIELD-BASED RESEARCH**

**6.1 Introduction**

This chapter describes how the fieldwork dimension of the study was set up and conducted. It starts with the aims and objectives for the fieldwork, and then discusses the research issues to be investigated. It goes on to present the methodological framework for selecting the case study schools and to describe how the fieldwork was conducted: the timescale, fieldwork methods, and form of recording. It concludes by highlighting the main methodological problems and how these were addressed.

**6.2 Fieldwork Aims and Objectives**

The aim for this qualitative study was to explore empirically and to evaluate the ways in which selected schools were using value added analyses. Because of the author's involvement with the QUASE service developed at the National Foundation for Educational Research (NFER), it seemed sensible to choose the case studies from schools subscribing to QUASE. This had the advantage that the author was familiar with the technical aspects of the data and knew what conclusions could and could not accurately be drawn from it. A counter-argument might be that QUASE schools are self-selecting and possibly untypical of the majority of schools, in that they have actively chosen to use value added data. This could be construed as another advantage, however, for this study: schools using QUASE were subscribers to a service for which they had to pay a fee and could therefore be expected to have given at least some thought to the use of the analyses they had commissioned.

A full description of QUASE, together with sample data, is given at Appendices A and B, but it may be useful to give an outline in this chapter: see overleaf for the publicity information sent to schools.

## THE QUASE SERVICE, 1996

QUASE (Quantitative Analysis for Self-Evaluation) is a 'value-added' service provided by the NFER to secondary schools. It assesses the progress schools make with their pupils between joining the school and taking GCSE examinations, by allowing for key factors which have been shown to affect academic performance. The service is completely confidential and is based on the belief that school managers and teachers need high-quality quantitative information to set alongside other forms of evidence about their school's performance to inform their development planning.

A QUASE Profile for each school is produced annually, which generates results – together with a detailed commentary – on **seven GCSE-related performance indicators and fifteen individual subject areas**. Each set of results is presented numerically, graphically and in words, to aid understanding and interpretation. A rolling three-year analysis is provided as soon as school data permits.

QUASE Profiles have been shown to be useful management tools in helping to identify a school's strengths and weaknesses with different groups of pupils and in different departments; and can therefore help school managers to make appropriate changes in school organisation, curricula and/or teaching/learning strategies.

### ***FURTHER OPTIONS***

#### *Analyses of Previous Years' Data*

Analyses of previous years' performance (as far back as the Year 11 cohort 1993-94) for schools participating for the first time.

#### *Predicted vs Actual GCSE Results for Individual Pupils*

#### *Questionnaire Surveys of Students and Parents*

Attitudinal surveys of students and/or parents covering:

- ◆ experiences and views of school generally
- ◆ discipline, rules, bullying, school environment, homework
- ◆ careers education and guidance
- ◆ preparation for post-16 choices.

#### *Information Sheet for Governors*

Summary sheet for the school's governors explaining QUASE and summarising the school's results in straightforward language.

*After-Care, e.g. Helpline; Feedback Sessions; Brokerage with Other QUASE Schools*

The fieldwork project was designed with the intention of providing a small-scale but indicative empirical evaluation of the following issues:

- what the common needs and gaps in technical knowledge are, especially for senior and middle managers: *value added as an innovation*;
- the uses to which value added data is actually being put in schools: *value added as a lever for change*;
- the managerial and political issues this raises: *value added as a management instrument*;
- the kinds of school culture or ethos most likely to make use of value added for improvement: *value added as a cultural artefact*.

Accordingly, the objectives of the study were to arrive at an understanding of the ways in which staff at different levels within schools' hierarchies were (or were not) given access to, and enabled to use, sophisticated analyses of performance data which were potentially revealing about under-achievement in different subject areas and amongst different groups of pupils. The study also sought to answer the questions: how well are the new expectations and new uncertainties which value added approaches bring with them being managed? What are the main or common unintended consequences – including ethical and professional ones – of introducing and using value added analyses?

### **6.3 Research Issues**

From the objectives listed above and using the discussion in the foregoing chapters as a basis, I was able to generate a number of specific research issues which I thought relevant to a study based on qualitative fieldwork. I expressed these as follows:

- how do school senior managers understand and respond to value added data? How do they integrate it with review and development planning processes? With whom do they share value added information and for what purpose? Can a range of functions and uses of value added data as a management tool be identified? What are the main constraints and barriers to understanding and use?
- how do middle managers understand and respond to value added data? How do they integrate it with their review and target-setting procedures? What are the main constraints and barriers to their understanding and use?
- how do teachers understand and respond to value added data? How do they integrate it with their diagnostic, assessment and curriculum planning

procedures? What are the main constraints and barriers to understanding and use at this level?

- are staff in schools with poor value added results surprised to learn this, or does it confirm what they knew? If the latter, what reasons – managerial, pedagogical, curricular, resource-related – would they advance for the results?
- do poor value added results on one or some key performance indicators mean that any given school has lower-than-expected results on all performance indicators? Have the staff identified precisely where their strengths and weaknesses lay? Are there any pupil groups or subject areas in these schools whose performance has been satisfactory or even higher-than-expected?
- turning to schools with good value added results, again, does this confirm what staff feel they knew? What reasons would they advance for the results? Are there any pupil groups or subject areas whose performance has not been so good, in value-added terms? What evidence, if any, is there of consistent improvement over time?
- how well are school staff able to use performance data interactively with their own professional judgements about effective teaching and learning, at departmental/ year group, at classroom and at individual pupil levels? How do school managers and teachers use their results to set realistic targets for individuals and for groups?
- how far do staff at different levels consider value added data helps to raise pupils' attainment? What evidence can they demonstrate for this?

It became clear that these research issues were too ambitious for the current study, but I am leaving them *in situ* because I think they are worth articulating for any subsequent work. What the list does demonstrate is that there is much complexity and inter-relatedness in the kinds of questions to be explored; and thus that a case study approach, that is, one based on interviews and on-site observation rather than a mailed self-completion questionnaire, is the most appropriate method.

The remainder of this chapter addresses the selection of schools and the ensuing conduct of the case studies.

## 6.4 Design of the Study

### 6.4.1 Source of the data

As background to this study, it needs to be noted that, during the period of this study, the author worked at the NFER where, as Head of the School Improvement Research Centre, she had responsibility for running the QUASE (Quantitative Analysis for Self-Evaluation) service which provides confidential value added analyses to subscribing secondary schools. The NFER funded the field-based study described in this chapter in order to acquire further intelligence to help with the development and promotion of the service. Schools selected to participate in the study (see Sections 6.4.3 – 6.4.5 below for methods of selection) were all past or current subscribers to the service; there was a wealth of performance and other data available at the individual school level to inform and shape the study. **Appendices A and B** consist of an anonymised set of tables and diagrams, together with commentary, for a QUASE school: this is the kind of information each case-study school would have received.

An issue which arose early in the course of designing this study concerned the use of QUASE schools and their data. Two quite different dimensions of this issue were explored:

- That of whether, and how, to keep the research – as an impartial, exploratory and descriptive exercise – distinct from any guidance and consultancy input which the school managers might have wanted or expected from the author in her professional capacity. After considerable thought, I took the decision to adopt the role of the ‘impartial observer’ so as to conduct the project in a way I thought appropriate to a doctoral study (see Chapter Five above – Section 5.3 – for a discussion of the context and Section 6.6 below for a note of how it was handled in practice).
- That of whether to widen the scope of the study to include schools using other value added data services than QUASE. I rejected this as an inappropriate course of action for someone employed by one of the institutions providing such services, since it might not have been seen as an impartial study<sup>11</sup>. As it turned out, some of the case study schools were using more than one source of data (this was before the implementation of

---

<sup>11</sup> Around that time, as a matter of fact, I, together with other colleagues active in the field, wrote to DfEE and QCA suggesting that they should commission a comparative evaluation of different value added systems, with recommendations for the further development of the national system. None of the signatories felt they could, for obvious ethical reasons, engage directly in such an evaluation. The reply, when it came, was informal and indicated that this was not a policy priority.

the national value added system) – although school staff often did not or could not distinguish between one system and another.

Thus the study was originally designed on the basis of a detached investigation (rather than a research-and-development project) using only QUASE schools. In the event, however, a number of complicating factors came to light which, whilst not undermining these decisions, meant that in order to give a good account of schools' use of value added data I needed to make use of the more 'engaged' understanding I had acquired through my professional work. This theme and its implications is explored in the latter chapters of the thesis (and see Sections 6.7 and 6.8 below).

#### **6.4.2 Methodological rationale**

As was intimated in Section 6.3 above, to investigate the research issues with any degree of success would entail acquiring subjective as well as factual information from a number of sources and with more than one inquiry method. It would have been possible to acquire some information by mailing out a questionnaire survey but it was considered that qualitative fieldwork conducted *in situ* with a small number of schools had a greater capacity to provide insight into the processes and 'narratives' of the management of value added analyses.

Some general support for this view can be found in a number of sources (for example, Marshall and Rossman, 1995, p.146; Frankfort-Nachmias and Nachmias, 1996, p.299). However, the present study differs from more traditional qualitative research in the following ways:

- it is explicitly related to current government policy, and therefore has an agenda which is not open-ended;
- it is based on, and also seeking in a particular way to elucidate, complex quantitative data;
- it does not pretend to 'share a commitment to naturally-occurring data' (Silverman, 1993, p.23) in the ethnographic, anthropological or symbolic transactionist conventions; in intention, it is probably closest to what Atkinson, Delamont and Hammersley (1993, p.17) call 'qualitative evaluation'.

The fieldwork in the chosen schools (see below for selection criteria) was accordingly designed to include the following activities:

- scrutiny of school policy and strategy documents pertaining to value added in the context of raising attainment;
- programme of interviews with key personnel (senior management team, heads of department), using semi-structured interview schedules;
- programme of interviews with teaching staff, again using semi-structured interview schedules;
- occasional attendance at relevant staff meetings, INSET sessions, etc.

It was planned that such activities would take place at intervals over a period of nine months, in order to observe the impact of value added in the life of a school.

#### **6.4.3. Selecting the initial sample**

As explained above, schools participating in the NFER's QUASE service formed the potential sample. In March 1998 a total of 104 schools were approached by letter (see **Appendix C**); this number represents the majority of schools which had participated in QUASE the previous year, i.e. 1997. (A handful of QUASE schools were excluded from the mail-out because their QUASE results were still being scrutinised and/or corrected.) Some 33 positive and 14 negative responses were received in answer to this mail-out. A holding letter was accordingly sent in early April to the former, explaining that a sample would need to be drawn, as a result of which they would be contacted near the beginning of the summer term about their subsequent involvement (see **Appendix D**).

A Research Information Sheet was also sent to these schools, so that they could have some general preparation for what would be required of them (see **Appendix E**).

#### **6.4.4 Devising criteria for selecting case-study sample**

It was decided that a sample of ten schools – suitably selected – would probably provide a good range of priorities, problems and issues relating to the use of value added measures, whilst still being manageable within the given timescale and budget. In order to coordinate the NFER-funded project (see Section 6.4.1) alongside the present doctoral study, the author took responsibility for the research design of the fieldwork (and for all the discussion which appears in this and subsequent chapters, except where explicitly noted) and for actual fieldwork in four of the ten schools.

There is plainly no requirement that a qualitative sample be representative, since its purpose is to provide evidence of *issues*. Nonetheless, I thought it important to identify obvious sources of bias in the selection which might have the consequence of failing to encompass some key issues. In the context of this study, the main need was to select a group of schools which were sufficiently differentiated in one or more of the following ways:

- they were not all situated in only one or two local education authorities (which might have their own histories relating to value added analyses);
- at least some had more than minimal familiarity with value added principles;
- at least some had provided information which showed evidence of pupils' 'starting points' on entering secondary school (the QUASE service is offered to schools which cannot provide prior attainment data);
- they did not all occupy a similar position in the published tables of performance (either low- or high-performing).

This last point is a crucial one, in that a major working hypothesis of the study, as mentioned above, was that schools with different 'starting points' would be confronting different issues and dealing with them in different ways. The following framework was accordingly used to assist in the progressive selection of ten case-study schools from amongst the 33 which would, at least *prima facie*, present an interesting and insightful range of problems and priorities concerned with pupils' performance. (Since part of the study was contracted under the auspices of NFER, I submitted this methodology to the internal steering group, and they approved it.)

#### *Stage 1. Overall profile*

Using QUASE Table 1.0 in the QUASE profile (**Appendix A**), all 33 schools were scrutinised to see how far they fitted the national profile of schools. The findings from this exercise are given in Table 6.1 below.

**Table 6.1 Comparison of QUASE schools, potential case study schools and the national population of secondary schools, England and Wales<sup>12</sup>**

	<i>QUASE schools</i>		<i>Case study sample</i>		<i>National popn.</i>	
	<i>%</i>	<i>Number</i>	<i>%</i>	<i>Number</i>	<i>%</i>	<i>Number</i>
<i>Total</i>	100	104	100	33	100	3374
<i>Type of school</i>						
Comprehensive to 16	36	37	41	13	38	1297
Comprehensive to 18	46	48	44	15	51	1715
Selective / independent	18	19	16	5	11	362
<i>Type of LEA</i>						
Metropolitan	39	41	44	15	33	1130
Non-metropolitan	61	63	56	18	67	2244
<i>Region</i>						
North	30	31	25	9	29	972
Midlands	13	14	19	6	23	790
South	55	57	53	17	41	1391
Wales	2	2	3	1	7	221
<i>1997 GCSE data (%5+A*-C)</i>						
< 25	34	35	29	10	21	724
26-35	19	20	19	6	18	593
36-45	15	16	16	5	17	584
46-55	14	15	29	9	17	565
>55	15	16	6	2	20	677
Not available	2	2	3	1	7	231

The table shows a bias amongst the potential case study schools in favour of a location in the south of England and in metropolitan LEAs; comprehensive schools with sixth forms and schools in the highest GCSE achievement category were under-represented. Redressing this bias in the selection of case study schools was not thought a priority, however, compared with the other selection criteria.

<sup>12</sup> 'Secondary' schools are defined here as schools having a Year 11 cohort, i.e. excluding middle-deemed-secondary schools and sixth-form colleges. Also excluded, for statistical reasons, are schools with a Year 11 cohort of fewer than five pupils.

Some additional pieces of politico-geographical information were available, including whether schools were LEA- or grant-maintained and in which particular local education authorities the schools were situated. Eight of the 32 sample schools were grant-maintained. The individual LEAs represented were as follows:

- 1 inner London borough (Southwark)
- 3 outer London boroughs (Barnet, Croydon and Ealing)
- 9 metropolitan boroughs (Dudley, Sandwell, Knowsley, Bolton, Tameside, Wigan, Barnsley, Bradford, Newcastle)
- 1 Welsh authority (Denbighshire)
- 7 shire counties (Berkshire, Devon, Essex, Hertfordshire, Kent, Lincolnshire, Surrey)
- 6 unitary authorities (Bath and NE Somerset, Hull, Luton, Derby, Poole, Southampton).

Despite the observable biases noted above, the spread of sample schools was considered to be relatively good – especially amongst different kinds of LEA.

#### *Stage 2. Length of participation in QUASE*

It was decided to distinguish between schools with only one year's worth of QUASE data and those with more than one year, on the following grounds:

- it might be expected that schools with more than one year's worth of data would have a different view of – and probably more familiarity with – value added data;
- schools with more than one year's data would also provide an indication to the researcher about trends in their performance (one year's worth of data would give only a 'snapshot' view).

A total of 11 schools had only one year's worth of data, eight two years' and 14 three or more years'.

#### *Stage 3. Availability of prior attainment data*

It was finally considered necessary to distinguish between schools which had data on the prior attainment of their pupils on entry to secondary school and those which did not. As mentioned above, value added analyses are carried out for QUASE schools without this data, but the assessment of relative

effectiveness is more rigorous if it is available. Analyses of this kind are also arguably more useful to schools themselves, in that the relative progress of a given cohort of pupils from a known starting position can also be assessed.

For various reasons, even if prior attainment data exists for the majority of pupils in any given Year 11 cohort, such data may not be available for the entire cohort: a pupil may move to a different school between Year 7 and Year 11; or (prior to end of Key Stage 2 tests) his/her primary school may not have conducted assessments and/or passed on the results to the receiving secondary school; and so forth. For the purposes of QUASE analyses, intake data is considered to be a reasonably good indication of a cohort's prior attainment if it is available for 75 per cent or more of the cohort.

Of the 33 schools in the sample, 20 had intake data for 75 per cent or more of each cohort under study; seven had partial data (less than 75 per cent for one or more cohorts); and six had no intake data for any cohort.

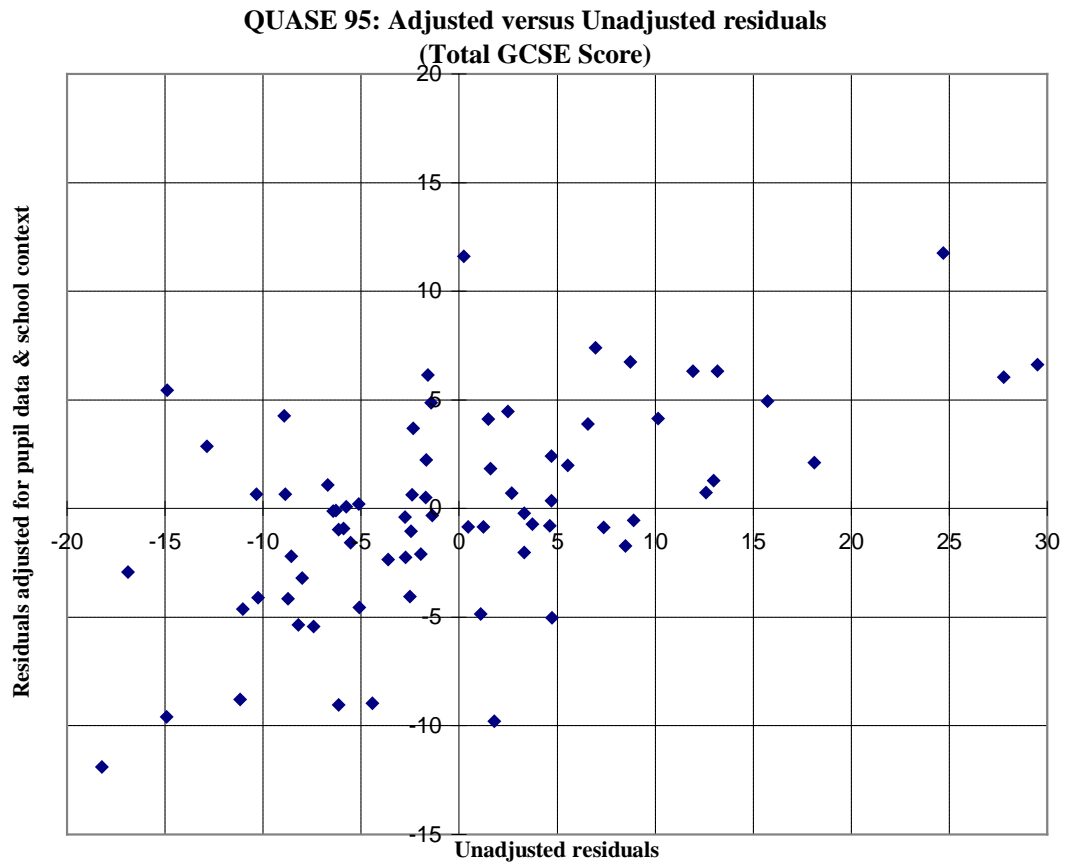
#### **6.4.5 Selecting the case-study sample**

Given all the above, it was decided to proceed by *excluding* from the outset the 13 schools with no or partial intake data, on the grounds that without pupils' intake data a full and rigorous 'value added' analysis of performance is not possible.

Three other factors were then taken into account in choosing the 10 case study schools:

- i. the desirability of conducting case study work with low-, medium- and high-performing schools as reflected in their raw scores at GCSE, on the supposition that their managerial and pedagogical priorities would be different;
- ii. the further insights which might be gained by including schools with positive and with negative value added scores, irrespective of their raw scores. The quadrant diagram below (Figure 6.1), based on earlier QUASE data, illustrates this principle. Schools in the upper right-hand quadrant are those which performed well on raw scores and showed better than expected performance ('value added') even when the analysis took into account a whole range of factors which might have affected their pupils' performance. Conversely, schools in the lower left-hand quadrant are those showed below average performance on both raw and adjusted scores, that is, even after every adjustment had been made for pupils' prior attainment, socio-economic context, etc.;

**Figure 6.1** Quadrant diagram illustrating different ways of grouping schools in terms of their ‘value added’



- iii. the necessity of checking that schools were not selected from only one or two LEAs and/or geographical regions; included some grant-maintained schools; included schools participating in QUASE for one year, two years and three-plus years.

*Step 1.*

The 20 schools with prior attainment data were categorised according to their levels of performance over the whole period for which there was QUASE data available for them (one, two or three-plus years). Instead of using the %5+A\*-C grades indicator (as in Table 1) which is rather limited and not very reliable<sup>13</sup>, the QUASE analytical model using seven separate GCSE-related indicators<sup>14</sup> was deployed. Schools were divided into three groups according to whether four or more of these GCSE indicators at the first stage of analysis (i.e. their 'raw' scores) were, statistically speaking, significantly above average, no different from average or significantly below average<sup>15</sup>.

On this basis, six schools had the majority of their indicators categorised as 'above average', nine schools 'average', and five schools 'below average'.

*Step 2.*

These three groups of schools were further scrutinised in terms of their 'value added'<sup>16</sup>.

**Group 1** (above average performance on raw scores): three of the schools maintained their above average performance on the majority of performance indicators when pupils' prior attainment was taken into account, whilst three showed only average performance. (No schools in this group showed below average performance at this stage of adjustment.) In common educational parlance, the first three schools could be classified as 'highly effective', the second three as 'complacent' or 'coasting'. There are arguments for selecting

---

<sup>13</sup> This indicator represents the performance of only the highest academic ability pupils; it is therefore limited as a means of assessing the quality of performance of all pupils in a school. Moreover, minor differences in the *number* of pupils in this category can manifest as a much larger change in the *percentage* of such pupils, especially if cohort numbers are small.

<sup>14</sup> These are: mean total GCSE score, mean average GCSE score, mean mathematics, English and total science GCSE scores, mean number of A-C grades, mean number of A-G grades.

<sup>15</sup> Statistical significance was calculated at the 5 per cent level in order to make assessments of 'above average', 'average' and 'below average' results, relative to the whole QUASE population. For more information about QUASE statistical models, please refer to Appendices A and B.

<sup>16</sup> The headteacher of one potential case study school withdrew from the project at this point because of threatened industrial action by teaching staff protesting against increasing administrative and bureaucratic duties.

at least one school from both of these categories, in order to address the questions:

- How does a ‘highly effective’ school identify what it is doing well? How does it seek to maintain these levels of performance over time? Are there any concealed areas of underperformance (e.g. amongst specific groups of pupils or in particular subjects) which need to be addressed?
- How does a ‘complacent’ school react to the idea that its underlying performance is less good than the published tables of performance would suggest? How does it identify underperformance and what actions are taken to address this?

**Group 2** (average performance on raw scores): eight of these schools continued to show ‘average’ levels of performance on most of the GCSE indicators when pupils’ prior attainment was taken into account. Choosing at least one school from this category would be justified, on the grounds that the questions for this group of schools have more to do with the interpretation of the data. Even though both their raw and adjusted results were not statistically different from the average, the detailed data for these schools would have contained several clues and possibilities for further exploration. How far were schools able to undertake this more sophisticated investigation? Only one school in this group had adjusted results which revealed a different – more negative – picture than its raw results. This school would also be an useful one to include because of its apparently substantial underlying underachievement.

**Group 3** (below average performance on raw scores): this group is potentially the most vulnerable in terms of the schools’ likely position in the published tables of performance. It was therefore interesting to find that three schools in this group had average levels of performance, once their pupils’ prior attainment was taken into account; indeed, one school showed above average performance on one indicator at this stage of adjustment. However, the other two schools apparently had substantial underlying underperformance, since their adjusted scores continued to show below average performance. Choosing at least one school from each of these groups would also seem to offer potential insight into the questions:

- did schools with below average performance on raw scores just assume that ‘value added’ analyses would reveal that they had in fact added value?
- what was the effect on staff in schools with below average performance on both raw and adjusted scores of knowing this? How far was it considered possible to identify and address the extent of underperformance?
- had the levels of socio-economic disadvantage<sup>17</sup> experienced by the two schools’ pupils diminished teachers’ expectations of their academic potential?

*Step 3.*

Combining this information on performance with the earlier information on location and length of participation in QUASE, ten schools were initially chosen as follows:

1. School A (south-east England); grant-maintained, coeducational, 1299 on roll; three years’ data; below average raw scores, below average adjusted scores.
2. School B (north-east England); LEA-maintained, coeducational, 1113 on roll; three years’ data; below average raw scores, average adjusted scores.
3. School C (north-west England); LEA-maintained, coeducational, 689 on roll; three years’ data; below average raw scores, average adjusted scores.
4. School D (Wales); LEA-maintained, coeducational, 736 on roll; three years’ data; average raw scores, below average adjusted scores.
5. School E (outer London); grant-maintained, coeducational, 1499 on roll; one year’s data; above average raw scores, average adjusted scores.
6. School F (west Midlands); grant-maintained, coeducational, 684 on roll; one year’s data; below average raw scores, average adjusted scores.
7. School G (south-east England); LEA-maintained, coeducational, 1448 on roll; one year’s data; above average raw scores, average adjusted scores.
8. School H (outer London); grant-maintained, single-sex (boys), 926 on roll; three years’ data; above average raw scores, above average adjusted scores.

---

<sup>17</sup> Data relating to free school meals, school catchment area and parental attendance at parents’ evenings was available on each school. For more details of how the data was used in QUASE calculations, please see Appendices A and B.

9. School I (south-east England); independent, single-sex (boys), 284 on roll; two years' data; above average raw scores, above average adjusted scores.
10. School J (north-west England); LEA-maintained, coeducational, 1052 on roll; two years' data; average raw scores, average adjusted scores.

These schools were duly sent a letter notifying them that they had been selected to participate and asking them to confirm that they were happy to go ahead (**Appendix F**). The remaining 23 schools were contacted by letter to notify them that they were not amongst the schools selected for case study (**Appendix G**). The author discussed with her team which schools she would undertake to visit personally.

## **6.5 Setting Up the Study**

### **6.5.1 Instruments**

Semi-structured interview schedules were drafted for use in case study schools (**Appendix H**), in order to address the research issues identified in Section 6.3 above. It was intended that the headteacher and the heads of the three core subject departments – mathematics, English and science – should be interviewed, together with any other staff who had a responsibility for disseminating or strategically using QUASE data. These latter included, in different schools, the head of ICT, the assessment coordinator and the coordinator for special educational needs. Semi-structured schedules were devised both because of the open-ended nature of the study and because of the relative seniority of the staff (who could be expected to have strongly-held and/or individually-elaborated views about a number of issues implicated in the study); a structured schedule would have been inappropriate in the circumstances.

Whilst these interview schedules were constructed to give shape and structure to the programme of interviews, it was expected that different schools might well have different working environments in which the issues raised by the researchers could be explored. A series of initial visits was therefore set up, in order to introduce the project and to gain some impression of the kinds of areas which might need most probing over the next few months. The nature of these initial meetings – on school territory – was very much in the hands of the headteacher and sometimes included a tour of the school.

### **6.5.2 Timetable of activities**

Visits to schools were arranged primarily to suit the schools, and so the timetable was constructed on an *ad hoc* basis: the author was conscious of the many administrative pressures on school staff. The fieldwork mainly took place over the period July 1998 to March 1999, although a few interviews were conducted in the summer term of 1999, in order to check out observations made in earlier visits around which there was some important uncertainty. A schedule of visits conducted in each school is given in Chapter Seven.

In the event, several issues arose during the course of fieldwork which modified the nature of the study and its outcomes. These are alluded to in Section 6.7 below and further discussed in the following chapters.

### **6.5.3 Recording and writing up**

Each visit and interview within the visit was recorded manually, according to the exemplar set out overleaf. A tape-recorder was not used because, in the author's experience, some interviewees tend to feel inhibited by this. Handwritten notes were taken at the time, and any further details added as soon as possible after the interview was completed.

Detailed write-ups of each interview were included in the record, using verbatim quotations where recorded. Issues to be followed up in later visits were also recorded.

*[Front page summary]*

**SCHOOLS' USE OF VALUE ADDED DATA: AN EXPLORATORY  
STUDY**

**CASE STUDY NOTES**

<b>School name</b>	<b>SCHOOL G</b>
<b>Interviewee(s)</b>	<b>Head of IT; Head of Science; Head of Maths</b>
<b>Observation(s)</b>	<b>Departmental Reviews of: English; Modern Foreign Languages</b>
<b>Date</b>	<b>12.10.98</b>
<b>Interviewer</b>	<b>Lesley Saunders</b>
<b>Documents obtained</b>	1998 GCSE results by gender; print-out from spreadsheet for individual pupil assessments; analysis of MFL results
<b>Documents requested</b>	1998 Examination Results Whole School Analysis (sent by post)
<b>Overview</b>	Conversations with Heads of three Departments; attendance at two Departmental Reviews
<b>Follow up</b>	Headteacher's secretary to arrange dates/times for follow-up interviews with the above members of staff

*A full write-up of each interview conducted then follows, and the record  
concludes with issues to be pursued in the next visit.*

## 6.6 Research Protocol

Although the Research Information Sheet (q.v.) promised that schools would receive informal feedback as a *quid pro quo* for their participation in the project, it was important for the author (and other members of the research team) to define the boundaries of this.

### 6.6.1 Issues of confidentiality

The NFER Code of Practice (NFER 1996a and b) – by which the author was bound – makes it clear that information given in confidence must not be divulged outside the research team without being anonymised and/or aggregated. There is no ambiguity about this<sup>18</sup>, which means that any piece of information given to the researcher in the course of school-based interviews or observation cannot be passed on to the school's management team. There is usually no problem with such strictures, and an experienced researcher learns how to provide satisfactory and useful feedback to a head teacher at the end of fieldwork which preserves confidentiality.

### 6.6.2 Issues of research versus development

The particular problem for this study was to clarify for participating schools the distinction between research investigation and development/guidance (see Section 6.4 above for a discussion of why this distinction was built into the study design). The following form of words was used in introducing the project to school managers and staff:

*Because this is essentially a research project, it is very important that the researchers maintain impartiality throughout the fieldwork phase. There is an urgent need for policy-makers to understand how schools respond to and use value added data, and it is our responsibility to make every effort not to 'contaminate' this important piece of work by giving you advice or guidance as we carry out the fieldwork. We hope you will understand this. We shall be glad to give you feedback at the end of this project, but in any case would remind you that you are always welcome to seek advice from the QUASE team at the annual clinics on questions of interpretation, etc.*

The penultimate section in this chapter outlines the key issues and problems encountered in implementing the fieldwork.

---

<sup>18</sup> Except in extreme and exceptional cases where pupils' welfare was at stake, for example, when advice would be sought from NFER's senior management.

## **6.7 Main Methodological Problems Encountered**

It is to be expected that a research design will not always be able to be implemented in precisely the ways envisaged, and this study was no exception (as was outlined in Section 6.4 above). The problems encountered, and how they were addressed if it was possible to do so, are discussed briefly in this section.

Two areas gave rise to immediate problems or issues: the selection of suitable case study schools (and conduct of some interviews therein) and – once that problem had been resolved – the unanticipated lack of information about and/or use of QUASE and value added data amongst staff in the case study schools. This last problem meant that I felt bound to utilise knowledge I had acquired from outside the study (see Section 6.8).

### **6.7.1 Selection of case study schools**

As noted in Section 6.4 above, the QUASE database was used to identify suitable candidates for case study work after an initial invitation to QUASE schools to participate. It had been assumed that schools would wish to participate only if they were still actively involved in QUASE. In the event – that is to say, on the initial fieldwork visit – it was discovered that three schools (Schools E, F and J) were not using QUASE in the current year. This meant that, although questions could have been asked about previous years' practice in these schools, the issues could not have been explored on the same 'live' basis as in other schools and so a decision was made not to work with these schools any further.

The fact that this was not discovered until the fieldwork had started should not perhaps have been a surprise: in secondary schools, the contact person (often a senior teacher or deputy headteacher) may well be an enthusiast for QUASE, but decisions about what expenditure to commit are made 'in another part of the forest', e.g. by the finance committee of the governing body, and are not necessarily conveyed to the appropriate person until later in the year.

Using the same criteria as before, two substitute schools were selected (rather than three – by that time the timetable for fieldwork was becoming rather tight): these became new Schools E and F. In place of working with a third substitute school, a one-day visit was made to the school written up by West

and Moore (1998) and discussed in Chapter Five. The author considered that the issues posed by this case study were pertinent to the thesis.

### **6.7.2 Difficulty of accessing interviewees**

The many administrative pressures on school staff mean that hosting research studies is now a favour which schools grant rather than a responsibility they welcome. In most schools, staff were generous with their time and insights. But in one of the author's case study schools it proved difficult to get access to any other member of staff apart from the deputy headteacher, on the grounds that QUASE was in its early days in the school and no dissemination of the data had been undertaken. In another school, the interviews with the headteacher and chair of governors proved hard to arrange; the interview with the headteacher did not happen until summer term 1999, and the planned interview with the chair of governors did not take place at all.

### **6.7.3 Level of basic knowledge of QUASE**

It had been assumed from the outset of the study that the fact that these schools had invested time and money in commissioning value added analyses meant that school staff would be well briefed about QUASE and how it operated. In fact, the fieldwork visits raised several unanticipated questions about the level of basic understanding of QUASE in schools. This meant there were problems in attempting to conduct the fieldwork according to the original research brief.

These problems are discussed and investigated in the following chapters, which undertake three distinct but related kinds of activity: presentation of the fieldwork evidence on a school-by-school basis (Chapter Seven), followed by further exploration of the database information on each of the case study schools in the light of that evidence (Chapter Eight), and then by a more speculative analysis of participants' responses (Chapter Nine). The intention is, by using an iterative approach, to work towards a framework for understanding the behaviours and practices in the different schools.

## **6.8 Subsidiary Evidence**

As has been said before, in my professional role I act as a researcher and adviser on value added. Coincident with analysing the fieldwork evidence from this study, I was involved in providing QUASE performance feedback to school staff in seven secondary schools in an LEA which had commissioned the analyses. Because the fieldwork was delivering some unexpected outcomes, I took the opportunity to see whether the commissioned in-service sessions had anything different or additional to offer by way of evidence. Notes from the meetings held in the seven schools were accordingly taken, and some of the work is presented and discussed in Chapter Ten; the implications of this for the study itself are discussed in Chapter Eleven.

**CHAPTER SEVEN**  
**ANALYSIS OF CASE STUDIES I: A DESCRIPTIVE FRAMEWORK**

**7.1 Introduction**

So far as the fieldwork dimension of the study was concerned, it was first of all important to devise a descriptive framework with which to organise the different aspects of the evidence and the preliminary findings. The author and her colleagues shared fieldwork notes and observations throughout the fieldwork, and the author tested out her emerging perceptions and assessments in team meetings. In order to impose some commonality on the diversity encountered, both of school context and of policy and practice with regard to value added data, it eventually seemed useful to provide a descriptive analysis of each of the nine schools, written up as a brief case study with the following dimensions:

- the school and its context;
- the programme of visits and interviews;
- an overview of how value added data seemed to be perceived and used, by the senior management team, heads of core departments and, where applicable, other teachers and governors (distilled from a total of 80+ pages of interview notes).

The programme of visits and interviews for each school was developed in response to the different issues identified: it was therefore considered more useful to give information about the visits in this chapter than in the preceding one.

The chapter thus provides only a basic overview of the fieldwork data; it concludes with an overview of the issues which the evidence seems to raise and which consequently need to be explored in further chapters.

**7.2 The Case Studies**

Each school was visited over a period of about nine months (summer 1998 to spring 1999); interviews were set up and conducted in each school with the headteacher and the heads of the three core departments (mathematics, English and science) or, alternatively, of the departments most closely involved in using value added data. Further interviews were sometimes subsequently

arranged as the fieldwork progressed with school personnel who were either directly recommended by other interviewees or whose names had been mentioned in interview in such a way as to suggest to the researchers that they might have a professional involvement/interest in value added data. Interviews with members of staff other than senior managers or heads of department, although originally part of the study design (Chapter Six, section 6.4.2), did not take place to the extent envisaged. If it emerged – as it sometimes did – from interviews with heads of department or senior management that little work was happening around QUASE data even at middle management level in the core departments, then other staff were not interviewed.

Visits and interviews were recorded using the proforma given in Chapter Six (section 6.5.3). Interviews followed the broad headings of the semi-structured schedule (see Chapter Six, section 6.5.1). Interviews tended to take between 45 and 60 minutes (sometimes longer), and this was true even when the central topic turned out not to be QUASE/value added data. The interviewer did not abandon interviews in those cases, but instead recorded the issues and topics which each interviewee seemed most keen to discuss; typically, these related in some way to the use of performance and assessment data.

In some schools, e.g. School I, it was relatively easy to establish how QUASE data was being used; in other schools, e.g. School G, it took several visits before the interviewer was able to grasp how QUASE data was perceived and used (or not) by different heads of department.

Sections 7.2.1 – 7.2.4 below cover the schools visited by the author (i.e. Schools A, F, G, I); sections 7.2.5 – 7.2.9 deal with those visited by her colleague, for which the visits and interviews had been recorded in a similar format (Schools B, C, D, E, H).

## 7.2.1 SCHOOL A

### PROGRAMME OF VISITS AND INTERVIEWS

Three visits were made to the school, as follows:

1. Extended initial interview with member of senior management team with responsibility for performance data.
2. Separate interviews with head of science, head and deputy head of English, head of mathematics, and two members of the senior management team; joint interview with heads of technology, modern foreign languages and performing arts.
3. Interview with headteacher.

It was also intended that the interviewer should interview the chair of governors and attend at least one of the middle management training sessions on the use of performance data (planned for autumn 1998 and spring 1999 terms); but neither of these possibilities materialised.

### THE SCHOOL AND ITS CONTEXT

*Basic information (for 1998):*

- grant-maintained upper school
- coeducational
- 1068 on roll
- no pupils with English as a second language
- over 25 per cent of pupils eligible for/taking free school meals
- around 30 per cent of pupils classified as having special educational needs

The school is situated on the margin of a large local education authority, in an area which is geographically and economically quite isolated. Key points from the initial interview with the member of senior management can be summarised as follows:

- The dockyard closed in 1967 and since then the local economy has been depressed: 25 per cent of males in the 18-25 age group are unemployed, for example, and other indicators like single parent families, multiple occupancy housing, lack of inside bathroom, are all high.
- The interviewee characterised the young people as '*shrewd rather than streetwise*'. An additional feature is the high proportion of fostered children in the area – every year the school has around ten students on roll at the beginning of Year 9, officially fostered locally but whom staff never or rarely see. This acts as another small downward drag on the school's results.

- The area has a population which is skewed towards both the very young and the post-retirement groups. This, coupled with the high unemployment rates, means that the locality contains a large number of people dependent on welfare. There are almost no professional or managerial parents, since the managers of the area's businesses commute in from more well-to-do towns some twenty miles or more away. This has also made it hard to recruit governors with the requisite skills. The main economic activities are transport-oriented or tourism-oriented; and the 'black economy' thrives.
- The school itself is a 13-18 upper school with a 16-form entry (cohort size in the region of 350; full-time staffing equivalent 95+). It was founded in 1970 as a 'community school', but this idea was never quite realised. Before that, there were two secondary modern schools and a technical school in the locality, but these were amalgamated to make one large upper 'wide ability' school. Because there are two single-sex 11-18 grammar schools within travelling distance, these take at least the top 25 per cent of pupils per year, academically speaking; there has been a steady drift of able pupils out of the area's schools, particularly since the introduction of open enrolment. School A now has a very skewed intake, with only a bare handful of students falling in the upper quartile of academic ability. The school has more boys than girls and one-third of the intake each year has some form of special educational need, although few are actually stated.
- School A is 'fed' by three middle schools, one of which is a denominational (C of E) school, one is grant-maintained and the third LEA-maintained. These tend to lose their brightest pupils at 11+ to the selective schools; thus pupils who stay for the last two years of middle school already sense themselves to be failures. The girls' grammar school within travelling distance holds places available at 13+ as well, so has the capacity to 'cream off' even more pupils.
- The school was inspected by OFSTED in 1998, when it narrowly avoided being put on special measures, partly because of the above factors and partly because the senior management team could say they were aware of the problems of 'chronic underachievement' (for example, only about 15% 5+A\*-C grades at GCSE) and were trying to put structures and systems in place to address them.
- Despite a deficit budget, the governors made a decision to 'over-staff' the school so as to give a 14:1 pupil-teacher ratio, partly in order to provide more support for learning difficulties but partly because – being an upper school – there are only two year groups, Years 10 and 12, where there are no public examinations at the year end.
- According to the interviewee, a key aspect of the school's problems is the 'stagnation' of the staffing body, of which the salient features are these:
  - Staff are mostly at the top of the pay scale and therefore expensive to maintain.

- The staff tends towards middle age and turnover has been relatively low – about ten per cent a year – mainly amongst young or newly-qualified teachers. The average age/length of service in the school means that many of them, particularly heads of department, are resistant to change; they are particularly cynical about numerical data and prefer ‘professional judgement’.
- The very large cohort/year group size means that departments are disproportionately large, allowing not just diversity but unilateralism and fragmentation to take hold.

## OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

The school joined QUASE in 1995. In the initial interview, the director of finance said that the school had not really used QUASE data so far – the senior management team wanted to build up a three-year picture before setting about making radical changes. The senior management team had been reconstituted by the headteacher on taking up appointment eighteen months ago, and consisted of the headteacher and five directors, of finance, curriculum, personnel (who was also the assistant headteacher) main school and post-16 respectively. It is possible that these changes were taking time to become institutionalised. The interviewee said the main need at the moment was to ensure that policies on raising achievement were consistently implemented in each department and year group. Even so, he said that senior management considered QUASE to be a potentially powerful lever for change in identifying the extent of under-achievement.

The other interviews conducted confirmed what was said during this initial visit. So far as the prime object of the current study was concerned – the use of QUASE value added data – little was garnered; it was difficult to see how QUASE data could be used a lever for change in the way the first interviewee said senior management wished, since most heads of department seemed unaware of QUASE, and therefore of what the analyses revealed.

The interview with the headteacher did not add much to these impressions since, although the purpose of the interview had been made clear in writing beforehand and the headteacher had agreed to the interview, he began by saying that he passed QUASE and other performance data immediately to the other members of the senior management team to deal with.

There seemed not to be a clear rationale for delegation of data-related activities. According to another member of senior management, target-setting interviews with individual pupils had up to now been conducted by form tutors, whereas CAT (Cognitive Abilities Test) prediction data was given to heads of department. The interviewee himself believed that the follow-up system for monitoring progress, covering pupils’ attendance and homework, should operate through subject departments as well as form tutors. Currently, target-setting interviews for Year 10 pupils were being conducted by senior staff, because of ‘*the dangers of getting it wrong*’ if left to form tutors.

Overall, therefore, it seems as if QUASE analyses were currently ‘trapped’ in the layer of senior management below the headteacher, although this may not have been obvious. Of the people in this team, the person who seemed to understand and be the strongest advocate of the data and its implications was the director of finance. It is questionable, however, how much executive authority and/or influence he had, given the *de facto* autonomy of heads of department, to ensure a coherent strategy of dissemination of the data throughout middle management and to integrate its utilisation with whole-school strategies for raising attainment.

## 7.2.2 SCHOOL F

### PROGRAMME OF VISITS AND INTERVIEWS

Three visits were made to the school, as follows:

1. Initial joint interview with headteacher and the head of mathematics as the member of staff with responsibility for performance (including value added) data.
2. Separate interviews with head of mathematics, head of science and head of ICT/assessment coordinator.
3. Separate interviews with head of English, member of mathematics department and chair of governors (plus a final brief meeting with the headteacher).

### THE SCHOOL AND ITS CONTEXT

*Basic information (for 1998):*

- LEA-maintained comprehensive
- coeducational
- 1125 on roll
- no pupils with English as a second language
- under ten per cent of pupils eligible for/taking free school meals
- under ten per cent of pupils classified as having special educational needs

School F is in the centre of a small conurbation in the West Midlands. The school was founded 60 years ago as two single sex secondary modern schools. It became a mixed comprehensive in the 1970s, but never managed to establish a good reputation for results or behaviour, according to the present headteacher, who was appointed in 1983 with a remit to turn the school round. Her success in doing so she attributes to three main factors: strategic deployment of TVEI resources, new appointments of key staff, and development of relationships with parents; but the strength of her personality is probably a salient additional factor. The school has won a grant (through the national lottery) for a community arts and education centre.

The headteacher described the school as serving three council estates together with localities of home ownership, and the pupils as covering the whole ability range. The school has doubled in size since the mid-1980s to 1122 currently on roll. The expansion is apparently due to the fact that upwardly mobile families in the neighbouring LEA want to send their children 'over the border'. Very few pupils come from minority ethnic backgrounds, which the head said simply reflects the local population. The fact that boys have been underachieving relative to girls was mentioned: the average figure of 57% A\*-C grades at GCSE in 1998 was made up of 69% girls and 42% boys.

## OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

The school joined QUASE in 1995, so has had some years' experience of using QUASE and also other value added (YELLIS) data. The headteacher takes a strong interest in the monitoring of school and departmental performance and has promoted the use of value added approaches in that context. The head of mathematics, in his capacity as data manager, appears to have taken on the role of 'champion' of value added analyses for the senior management team and for other heads of department. He prepares an overview each year of what the different analyses show and gives a worked example from his own department to demonstrate what the analysis implies and how it can be used. Heads of department are expected to respond to the analyses by following up apparent areas of under-achievement and by using them to set challenging targets for pupils' performance.

However – as was clear from the interview with the head of English, for example – a certain *de facto* leeway in whether and how heads of department use the analyses is tolerated, if not necessarily welcomed. It emerged from one or two interviews that, when asked by senior management how they had followed up the QUASE analysis, some had not done so, pleading pressure of teaching load.

It was the head of mathematics – perhaps not surprisingly, given his role as 'value added champion' – who gave the most clear and convincing demonstration of the practical use of QUASE data in his department. The approach was verified by the member of his department interviewed at a later date, although she evidently deferred to, and had perhaps not wholly internalised, his views and rationale. The example is written up in Chapter Ten below.

The head of English held particularly strong views about the uses of data and these views are discussed further in Chapter Nine below. The head of science was also less convinced by the usefulness of this kind of data than his counterpart in mathematics, because he too had strong views about performance data analysis and target-setting, though for quite different reasons from those expressed by the head of English. For example, he considered that departmental meetings were '*too crowded to discuss statistical data*'. He was quite impatient with the '*complicated statistics*' he had to use and interpret

from different sources; as an ex-production engineer, he saw himself (and other heads of department) as operational managers and felt it was management's job, not his, to set targets.

The head of ICT, who was also the school's assessment coordinator, indicated that the principle of QUASE and YELLIS analyses was understood and accepted by heads of department as a group, and the data was used to raise interesting questions in meetings; however, the active use of data within departments was quite variable and '*good intentions do not always get put into practice*'. How the use of data might be promoted and developed by him in his role as assessment coordinator and/or the head of mathematics as data manager was not altogether clear. In his own department, the analyses had led to some experimentation with setting, for example, into single sex groups. This strategy had not been evaluated in terms of results, however.

The chair of governors was relatively new in post at the time of interview, and did not give the impression of understanding the intricacies of performance data analysis in such a way as to exert his authority to change what the school was already doing: in relation to target-setting, for example, he said '*[The headteacher] sets the standard and talks it through with me*'.

In summary, then, this school had (i) a headteacher who believed in, and was convinced of the management value of, value added performance data and (ii) a 'champion' of the data (in the head of mathematics) who both knew how to use the data and made efforts to help his colleagues do the same. Even so, the school cannot be said to have developed a whole-school approach which was agreed with, and implemented by, all departments equally.

### **7.2.3 SCHOOL G**

#### **PROGRAMME OF VISITS AND INTERVIEWS**

Six visits were made to the school, as follows:

1. Initial interview with headteacher.
2. Joint interview with head of mathematics, head of science and head of ICT; observation of English and modern foreign languages departmental reviews.
3. Separate interviews with head of mathematics (plus colleagues), head of English, head of science, head of ICT, head of modern foreign languages.
4. Joint interview with chair of governors and chair of curriculum committee.
5. Interview with senior teacher with responsibility for teaching.
6. Final interview with headteacher.

## THE SCHOOL AND ITS CONTEXT

*Basic information (for 1998):*

- LEA-maintained comprehensive
- coeducational
- 1329 on roll
- three per cent of pupils with English as a second language
- three per cent of pupils eligible for/taking free school meals
- 15 per cent of pupils classified as having special educational needs

The school is situated on the edge of a suburban sprawl north of the river Thames; the area was once a series of small villages which are now joined up by modern residential estates and shopping centres within easy reach of the M4 in 'Silicon Valley'. The school is on a clean and litter-free site and, beyond, plenty of tennis courts and acres of playing fields near a pleasant wooded area.

The headteacher had taken up post two years ago. His style of leadership appeared managerial rather than collegial, charismatic or authoritarian. The one colourful phrase he used was '*robber barons*' about the heads of department. He said that the school was achieving above average results, had a good reputation, and indeed had been oversubscribed for the last three years: 380 applicants for 240 places. The school caters largely for middle-class families in the immediate locality consisting of '*concerned*' parents and enjoyed a '*stable*' teaching staff. However, he said that there was '*some complacency*' amongst the staff which he wished to tackle. He supported this by saying that the school was not doing as well as it could in terms of 5+A\*-C GCSEs; it had maintained a steady 55 per cent or so over the last five years on this indicator, whereas its neighbour schools had made conspicuous improvements.

Another issue of concern was the increasing gap between boys' and girls' achievements at GCSE; in 1998, girls achieved 20 per cent more on 5+ A\*-Cs than boys at GCSE, and attained more A\*-Cs in every subject except PE and economics. Conversely, boys were obtaining 20 per cent more D-Gs than girls. However, the interviewer noted that the verbal reasoning scores in Year 7 and Cognitive Abilities Test scores in Year 10 had not so far been analysed by gender. Thus there was no evidence to indicate whether the GCSE results reflected a change over time in entry characteristics (due to some local demographic factor, for example) or were an indication of pedagogical issues within the school, or indeed both. The headteacher said that it seemed to be a matter of boys' attitudes rather than their ability, although he readily acknowledged that the school had only anecdotal and impressionistic evidence to go on at the moment.

With regard to data analysis and dissemination, the headteacher said that the reports for senior management team and governors had been simplified in 1997-98, because the great amount of material produced in previous years (by the member of staff with responsibility for statistical analysis) were not enlightening and did not lead to action.

#### OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

The school joined QUASE in 1997, but essentially, despite impressions to the contrary given by some of the interviewees, '*QUASE is not really used below senior management team level*'. This was said explicitly by one of the heads of department interviewed, and one or two of the others eventually implied something similar. However, because some interviewees either did not want to acknowledge this straight away or wished to talk about their use of performance/assessment data generally, it took some while for the interviewer to work out how restricted the dissemination of QUASE data was. This is why more interviews were conducted in this school than in the others.

So, for example, one head of department (ICT) spent a great deal of time and effort producing in-house data on departmental residuals for GCSE results which he disseminated to colleagues, and which he demonstrated to the interviewer. It is possible that he felt motivated to undertake this work because of the inadequacy of the 'official' performance information system (see above). But as his efforts were seemingly not at the instigation of the senior management team (and did not feature in conversations with the heads of English and modern foreign languages, nor in their departmental reviews), the extent to which the data was used seemed to depend entirely on individuals' initiative. The heads of mathematics, science and ICT were all keen to develop or discover a system of data analysis which could help them with target-setting and 'predicted' grades at pupil level. A great deal of work around pupil assessment was happening in each of these departments, but it was not clear to the interviewer how this was fed back into curriculum design and delivery.

By contrast, the head of modern foreign languages – in the presentation he made for the departmental review – had a clear method and rationale for linking the GCSE performance data with an incisive commentary on its meaning and implications for teaching. For example, the presentation disaggregated the different skill areas in the examination papers, compared boys' and girls' results in each, and suggested strategies for changing teaching approaches for the next cohort. Even so, none of the data used was 'value added' in a recognised sense.

Thus these interviews, like those in other case study schools, shed much light for the interviewer on heads of department's views of the role of assessment and performance data, but little on their perceptions and use of value added information. In fact, it was not until the very last interview in the school – with a member of staff who had recently come from another school which had been using QUASE – that the situation in this school became more

understandable. This person had the post of senior teacher ‘with responsibility for teaching’. The interview with her was arranged because the Chair of Governors had mentioned her name and role: no member of senior or middle management had suggested that it might be useful to speak to her.

In interview, this senior teacher said that she had been somewhat taken aback by the lack of consistent and diagnostic use of value added data in the school compared with the last school she had worked in. She was in the process of trying to get heads of department to work more closely on QUASE data but felt she was perhaps trying to move some people along too fast; and it seemed that although she had a formal remit for the work she was finding it hard to implement. (The author surmised aloud that this appointment had been made in order to try to bridge the gap between the senior management team and the heads of departments; the interviewee said that this was indeed the case.)

In the final interview with the headteacher, the interviewer raised the issue of the apparently different attitudes towards, and uses of, performance data by different heads of department; and asked what strategies the school was now proposing in the light of the under-achievement identified in the latest analysis of examination results. This analysis contained extracts from the QUASE Profile, but the QUASE results had not been circulated to staff in their entirety. The headteacher responded that he did not want to impose a standard approach on departments, but he acknowledged that *‘departments do get away with doing their own thing ... [and] some people have got to be moved on radically’*.

Thus the current situation may be explained by the fact that the heads of subject have in the past had no immediate incentive or compulsion to question their performance – as the headteacher said, the school has been over-subscribed and has had a good reputation with local parents; and the headteacher said about himself that he was a good starter of initiatives, but did not always follow them through. But the signs from the data – including but not restricted to QUASE – were inescapable that it was definitely under-performing relative to pupils’ attainment on entry. The headteacher appeared to be beginning to wish to tackle this more strongly. The key issue seemed to be whether the senior teacher’s experience and understanding of QUASE (from her previous school) could be recognised and harnessed, in order for the data to have a stronger role to play in both whole-school and departmental development.

## 7.2.4 SCHOOL I

### PROGRAMME OF VISITS AND INTERVIEWS

Two visits were made to the school, as follows:

1. Initial interviews with headteacher and deputy headteacher as the member of staff with responsibility for performance data.
2. Extended interview with deputy headteacher.

### THE SCHOOL AND ITS CONTEXT

*Basic information (for 1998):*

- independent selective
- single-sex (boys)
- 243 on roll
- one per cent of pupils with English as a second language
- no pupils eligible for/taking free school meals
- under one per cent of pupils classified as having special educational needs

School I at first resembles a large house rather than a school; it is situated in a well-to-do part of the home counties near the river Thames. Although the school is independent, the academic honours board (which the interviewer had a chance to inspect before the interview) started in 1985, had relatively few names on it, and hardly any Oxbridge places. In interview, the headmaster said he had taken up post in 1985 at a time when the school was, in his words, *'the last chance saloon'* for parents who wanted to send their boys to a fee-paying school. He claimed that the academic quality of the intake had never been high: *'an independent secondary modern... [for] 'sons of second-hand car dealers'* was how he described it. The recession in the late 1980s was particularly difficult for the school's potential parents and the school roll fell quite suddenly. The headmaster considers his job has been to raise expectations and the school's reputation so that it becomes a first-choice school. So the fact that two boys obtained places at Oxford University last year and a growing handful of others have gone on to other university places over the last ten years was a matter for celebration. The headmaster was delighted that QUASE data has so far vindicated his view that the school is doing very well with the boys who come here. The interviewer observed a quiet and friendly working atmosphere in the classrooms visited, although she also had the impression that the school was not used to being visited by external professionals (as distinct from parents and helpers).

According to the deputy headteacher, the school has good specialist staff with excellent academic qualifications and only one or two who are not also 'good teachers'. There is an emphasis on good relationships and individual support. There have been worries about performance in modern foreign languages,

particularly French (although in fact the QUASE data does not indicate grave problems here). The deputy headteacher talked of teachers being individualistic, identifying with the '*[School I] ethos*' but not being part of a departmental team in the modern professional sense. The '*[School I] ethos*' or the phrase '*a [School I] boy*' was mentioned many times in conversation; according to both the headmaster and deputy, the school is ostensibly concerned to instil positive relationships, a team spirit and a love of learning at least as much as to achieve good academic results.

With regard to curriculum development and school development planning, there is little formal documentation and decision-making in evidence. According to the deputy headteacher, '*the Headmaster is a charismatic leader*' and the development of the school is all '*in his head*'. There is the beginning of departmental target-setting, but the deputy headteacher claimed that they were not pushing this except in the case of modern foreign languages.

#### OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

The school joined QUASE in 1997 and, according to the deputy headteacher, the senior management team is still thinking about how best to use the analyses. QUASE data has not been disseminated to any staff outside senior management and the headmaster admitted that he does not take too readily to sophisticated numerical data; his main interest in QUASE has been as a promotional tool, i.e. to enable him to 'prove' to governors and parents that the school is doing well<sup>19</sup>. It was the deputy headteacher who has the professional interest in value added and perhaps this was really her agenda rather than anyone else's. However, because she was relatively new to the school, she did not consider it appropriate to push value added or target-setting too hard or too fast with colleagues.

The school does not participate in SATs, and although there are signs of a recognisable national curriculum in the curricular provision (and the deputy head spoke in terms of 'key stages'), there are differences as well: for example, Latin is compulsory for all boys. The school is moving to the use of the NFER-Nelson Cognitive Ability Test (CAT), in order to be able to explore progress and value added, especially from Year 9 to Year 11. However, there is already a very organised and intensive system for monitoring pupil progress within each subject and in terms of overall performance. It looks old-fashioned, being entirely paper-based and individualised by each teacher, but it seems to work in terms of giving early warnings about pupils' under-achievement. It is clear why the headmaster and deputy felt they needed to think very carefully about how to graft value added analyses on to such a well-established system.

---

<sup>19</sup> In fact, the QUASE Profile (in common with outputs from some other similar services) is accompanied by a request to schools not to quote text or data from it in any brochure or promotional literature.

On the first visit, it seemed possible that the deputy head – about whom the signs were that she had been appointed, amongst other things, to modernise the school’s approach to management – might be able to take a strategic lead on value added and other systematic ways of planning school development. However, the second visit suggested that she was still moving very cautiously; the most recent QUASE analyses had not been disseminated and she had formed no strong view as to how to take the agenda forward.

**The next five schools described were visited by the author’s research colleague; the descriptions were written by the author on the basis of the colleague’s fieldnotes and then checked with him for accuracy of fact and interpretation.**

*Note:* the author’s one-off visit to the school described by West and Moore (1998) – see Chapter Five – largely confirmed the findings outlined in Section 7.3 below.

## **7.2.5 SCHOOL B**

### **PROGRAMME OF VISITS AND INTERVIEWS**

Two visits were made to the school, as follows:

1. Initial joint interview with headteacher and assistant headteacher.
2. Separate interviews with headteacher, assistant headteacher, head of science and head of learning support.

### **THE SCHOOL AND ITS CONTEXT**

*Basic information:*

- LEA-maintained comprehensive
- coeducational
- 1113 on roll
- no pupils with English as a second language
- 43 per cent of pupils eligible for/taking free school meals
- 20 per cent of pupils classified as having special educational needs

The school has many of the characteristics of an inner city school, although one way in which the school differs from other inner city schools is that it has very few children from ethnic minorities. There are high levels of unemployment and social deprivation, especially in the area towards the river where the docks used to be. According to the assistant headteacher, some families in the area have experienced deprivation for generation after generation; there are quite a number of children at the school with severe behavioural problems and the staff have a reputation for dealing well with such difficulties. He said that children tend to enter the school with very low reading ages and the school has had a drive (prior to the national initiative) to improve literacy levels. The assistant head considered that the special educational needs coordinator (SENCO) and her colleagues had been remarkably successful in improving the reading ages of pupils.

The assistant head also said that there had been some limited expansion of the school's catchment into owner-occupied territory to the north over recent years and this was part of the reason why the school had been able to make a gradual, but sustained, improvement in GCSE results. The slow expansion northwards had brought in more pupils who were capable of achieving A\*-C grades and the culture of achievement had spread to pupils from the more deprived parts of the catchment area to the south. The school's percentage of students achieving five or more higher grade GCSEs had improved to 21 per cent in the previous year, a steady increase from only one per cent in 1990; although the most recent results were down again, to 14 per cent. The school had been bottom of the LEA's 'league table' for some time, but was now three or four places away from the bottom, which the senior management team regarded as a considerable achievement. One of the ways in which the school hoped to drive up examination results even further was through the use of Study Support, which provides facilities in out-of-school hours, to cultivate study skills in pupils and encourage further parental involvement.

At the time of visiting, the school had a youngish and apparently dedicated teaching staff, who seemed very friendly and lively. Teamwork was said to be a particularly important part of the school ethos and there was evidence that staff shared ideas, information and good practice, as well as social events.

#### OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

The headteacher and assistant head held a strong belief that the figures for examination results as presented in the 'league tables' did not reflect the efforts staff had been making, nor did they show the success of the school in improving reading levels, raising Key Stage results and improving pupil attainment generally. This concern had been the main driving force for using some kind of value added analysis: not only to demonstrate the perceived successes but also to boost staff morale: *'no one can work any harder than our staff do'*. The school had been participating in QUASE (and also ALIS and YELLIS analyses) for three years, as well as other assessment data; the headteacher was enthusiastic and knowledgeable about value added and school effectiveness issues, and was undertaking a doctoral study on the

subject under the guidance of an internationally-known professor at the local university.

The school's senior management team used a wide range of complex datasets with confidence – even though the headteacher at one point acknowledged, *'we are awash with data, in fact we have data overload'* – but it would seem that they were particularly pleased with the way in which QUASE data had enabled them to go beyond raw scores and crude league tables. It was evident that QUASE results were an important ingredient in the process of trying to maintain and boost staff morale. Selected comments from the QUASE Governors' Reports were included in the school's Examination Analysis documents for 1996 and 1997. The headteacher also mentioned the use of QUASE in identifying underachievement, which could be an issue for the school given the problems in the catchment area.

The headteacher claimed that QUASE analyses were *'very definitely'* used in development plans and in target setting; they had a *'direct influence'* and were an *'essential ingredient in school improvement'*. He said that performance analyses were useful in several respects:

- for exposing the school's results relative to neighbouring schools;
- for exposing subject differences;
- for exposing gender differences;
- for exposing the classroom performance of a teacher.

Active understanding and use of QUASE seemed to be the province of the two senior managers rather than of heads of department. Even so, it was not clear that they always made the distinction between QUASE and other kinds of data analysis. For example, the headteacher said, as if *apropos* of QUASE, that *'the best meetings we have are where heads of department talk about the individual child – why is he in set 3 for science, but 1 for maths?'* Since QUASE analyses do not give this kind of information, it seems likely that he was talking about assessment data on current pupils as well as (or as distinct from?) aggregate performance analyses.

The two senior managers' enthusiasm for QUASE was evident, but perhaps led the assistant head to over-estimate to some extent the heads' of department involvement with the data. There were indications that whilst heads of department viewed QUASE data favourably they were not making as full use of it as they might. For example, according to the assistant head, the QUASE data the previous year had revealed that the science department's results were not as good as could have been expected. He contacted the NFER to use the brokerage service in order to find an example of a school with a 'successful' science department. As a result, he said, the head of science and another teacher in the department were able to visit the school and to incorporate what they learnt into their own department's development planning.

The head of department's account had a rather less committed feel to it: '*I think this visit reassured us that we did a lot of things right*'. With regard to the use of QUASE data generally, she said:

*'[the assistant head] photocopies the relevant pages [of the QUASE report] and then gives them to heads of department. I don't use it very much – I probably should use it more... I don't pass it on to the department – you get so much information... To be honest I'm not sure what to do with all the information that you get... The school needs a policy, a common way of handling it.'*

In this context, it was interesting that she reported that the heads of department of science, mathematics and English met every month to compare pupils' predicted grades and progress: '*These meetings are really useful – we follow up inconsistencies*'.

The headteacher himself acknowledged that '*there is an urgent need for staff development in the use of performance data*', which would help them to build upon their professional knowledge of teacher effectiveness. But overall it seemed evident that this was a school which was evolving quickly in relation to the use of data; given the great deal of goodwill and sharing of information and strategies amongst staff, it would not be surprising to find the use of value added analyses much extended in future years.

## **7.2.6 SCHOOL C**

### **PROGRAMME OF VISITS AND INTERVIEWS**

Two visits were made to the school, as follows:

1. Initial joint interview with headteacher and first deputy headteacher.
2. Separate interviews with first and second deputy headteacher, head of English and head of geography.

### **THE SCHOOL AND ITS CONTEXT**

*Basic information:*

- LEA-maintained comprehensive
- coeducational
- 689 on roll
- no pupils with English as a second language
- 60 per cent of pupils eligible for/taking free school meals
- 35 per cent of pupils classified as having special educational needs

The borough has one of the worst unemployment rates in the country and the southern half of the borough, where School C is located, suffers more than the northern part. It is estimated that 50 to 60 per cent of families in the locality are single parent families. The school has three main primary feeder schools (with additional intake from half a dozen more). According to one of the deputy headteachers, an arrangement with other secondary schools in the area ensures that there are boundaries to the competition for 11-year-olds: this also makes for a degree of predictability in terms of the next year's intake. Relations with the LEA were said to be good. A number of disabled students are integrated into the school.

Interviewees were positive and upbeat about their school. One said that the staff as a whole refused to be negative and, although the impact of social deprivation was clearly considerable, it was not seen as an 'excuse for underachievement'. The school had recently achieved Sports College status which meant higher status and an increase of funding. The school had also received a number of local 'Quality Awards' for aspects of its work and was recently awarded *Investors in People* status.

Not surprisingly, the school had not fared well in the local and national league tables, though senior managers were quick to say that examination results were '*on the up*'. Senior staff spoke in terms of the school being an 'improving' school: this phrase was used several times during the interviews. The roll had decreased in the late 1980s and early 1990s, due to the lack of young families moving in to the area (back in 1974 the school had 2000 pupils). In the early to mid '90s there was some local speculation about the possibility of the school closing, but after the arrival of the new headteacher (who, interestingly, holds a degree in statistics) the situation had improved. The most recent OFSTED report was '*good*'. This improvement, coupled with new housing developments in the locality, were said to have contributed to a situation where the roll was increasing over the foreseeable future.

Levels of pupil attainment upon entry to the school were said to be below average on the basis of Cognitive Abilities Test (CAT), reading age and Key Stage 2 data, but were '*definitely improving*'.

#### OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

The school had been participating in QUASE for five years (as well as using CAT, MIDYIS and internal assessments). The senior managers said that the use of value added performance data, particularly QUASE, had '*definitely*' contributed to the process of school improvement. The headteacher remarked that the QUASE report '*provides an external view*' and was also a useful supplement to OFSTED information for the school. She added that, although there was some distrust of performance data at first, staff were now more than happy to use factual data. She mentioned that some Heads of Department were '*seriously into*' target setting and analysed data in detail themselves. QUASE analyses were reported to have been useful in a number of different ways, the scatterplots being particularly useful for heads of department. '*They*

*enjoy comparing like with like*'. However, it was suggested that, with all the data and expertise now available in-house or supported by the LEA, it was possible that the school would no longer need QUASE in future.

One of the senior management team described how QUASE was used thus far:

*'The three of us [headteacher and two deputy heads] look at the QUASE Report in detail. The head, together with one of us two deputies, then sees each head of department to review the results... QUASE provides a useful comparison of results... Heads of department are not threatened by this – QUASE is part of a wider picture. There is a good climate and staff want to be informed: the more information the better... The head is very good with statistics and uses them in a positive way'. She went on to say: 'The use of performance data and target setting has contributed a lot to our success. We're using a lot of information which is giving us a clearer picture of where we are going. It is a very positive school with a very positive climate of support. People are not threatened with information'.*

The head of the geography department confirmed much of this by saying:

*'[The headteacher is] very strong, very caring and very concerned that the children should have every opportunity to achieve their potential... Heads of department work in a very positive climate... We have frequent heads of department gatherings and we all observe each other's results... [statistics] are now part of the school culture [in the sense that] data is valuable to us as teachers, but not directly for raising attainment – it is useful as a kind of backup, a basis for us to set our own standards. On the whole we get about the right level of information'.*

This middle manager described her use of QUASE as follows:

- I look at the scatterplot*
- I try to understand it – I'm not statistically minded*
- I compare my department with other departments*
- I make the comparison with all QUASE schools*
- I compare it with other information and set the relevant targets.*

*I find that the statistics are useful, though I struggle with the presentations. The first time I went through it in detail with the headteacher... We can identify underachievement and try not to let it happen'.*

The head of English gave a very specific example of how QUASE data had been used to good effect:

*‘[The head and deputy] explain everything. It is very clear, we work together. We get copies of the QUASE report and she [the Head] shows us the scatterplots. I find the scatterplots useful. We have a review meeting and discuss the next targets. The idea is to keep improving... A weakness in English literature had been spotted from the QUASE scatterplot last year. When we looked at the breakdown of examination marks that made up these results, we found that the weak area was English Heritage Poets. The QUASE data started the ball rolling, and the specific information was in the examination mark scheme. This helped us to formulate where to target. We had extra homework, tests and even a mock paper on the English Heritage Poets. A specialist vocabulary was developed and we all worked hard on the English poets. The improvement shows in this year’s results, when QUASE suggested that that the English literature results were better than expected’.*

This school was clearly developing its use of data of various kinds, and not only at senior/middle management levels; the head of English said that the use of statistics was being ‘cascaded’ to other members of her department: ‘Some were reluctant at first, but they’re getting better now’. In the light of comparisons with many of the other schools in the study, the consistent mentions by interviewees of the headteacher’s leadership, the collegial climate and the staff’s perceived dedication to improvement are possibly key to this degree of development in the use of data.

## **7.2.7 SCHOOL D**

### **PROGRAMME OF VISITS AND INTERVIEWS**

Two visits were made to the school, as follows:

1. Initial joint interview with headteacher and deputy headteacher.
2. Separate interviews with headteacher, deputy headteacher, head of mathematics and head of science.

### **THE SCHOOL AND ITS CONTEXT**

*Basic information:*

- LEA-maintained comprehensive
- coeducational
- 736 on roll
- no pupils with English as a second language

- 17 per cent of pupils eligible for/taking free school meals
- 10 per cent of pupils classified as having special educational needs

The school serves the market town in which it is situated plus a large rural catchment area, together with some out-of-county children. The school has technology college status and has also received National Lottery money for new sports buildings. The LEA is a unitary authority in which school-LEA relations were said by the headteacher to be '*superb*'. Pupil intake ability levels were described by the headteacher as 'a mixed bag'. Quite a number of children came from farming families and these had suffered hard times recently. The school has a reputation for being good with Special Educational Needs pupils: at the time of visiting, there were reportedly 114 pupils with SEN at various stages of whom 40+ had full statements. '*We don't turn them away. We have pupils at all levels of ability*'.

#### OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

The deputy headteacher said that the school was using a number of different types of numerical data, including:

- QUASE (over three years or more – but the only school in the LEA using QUASE)
- Keele University survey of pupil attitudes
- national value added analyses and subject reports
- the school's own GCSE/A-level examination analyses (carried out by the deputy headteacher himself)
- CAT test scores and 'predictions'.

This led the interviewee to claim that '*we have more data than most schools*'. The interviewee described QUASE as '*straightforward and open*', although he was now looking for ways of tracking pupils' progress across Key Stages 2, 3 and 4 and felt the school could probably make better use of its Key Stage data. He mentioned that the school had done some 'contextual analysis' work with a renowned academic at the University of Sheffield.

The headteacher and deputy agreed that, four or five years ago, heads of department had not been involved with the use of performance data and staff had not been comfortable in terms of 'opening up' to the implications of performance data. In particular, they seemed to feel that QUASE analyses meant that their subject/department was being directly compared with other subjects/departments. Now, however, the feeling was that staff were very much involved in analysing and discussing data, and the headteacher said she was expecting even more progress over the next twelve months since they were aiming to '*create a climate in which we are able to look at the numbers*'.

The QUASE Report, according to the deputy headteacher, had two main uses: (i) to demonstrate any added value; (ii) to facilitate interdepartmental comparisons. It should be noted that the school had not had sufficiently accurate intake data for QUASE purposes, hence the view that *'Words are useful, even if the stats are less useful... When they come through, CAT scores will give us a much clearer picture, [with] more credence'*.

The deputy headteacher used the QUASE Report to compile an internal 'comparability study' (of departments) and to produce subject residuals on an annual basis. The comparability study – of which QUASE data formed only one small part – was the focus for discussion with heads of department of value added and target-setting issues. The interviewee reported that such data helped the heads of department *'to start talking – they open up the discussion'*.

The two heads of department interviewed were somewhat cautious about the use of data, seeming to express, in their very hesitancy, the tension between making use of performance data for internal development and being externally judged by it. The head of mathematics said:

*'it's all useful stuff – you need to look at trends and at the overall picture... you do have to be careful, though – you have to use both – your professional judgement and the statistics... I tend to keep away from residuals as much as possible – we're always going to come out with a negative score... I've always looked closely at pupils' results – we've been doing some kind of analysis ever since we've had computers... statistics can put you under tremendous pressure'*.

This head of department mentioned QUASE directly only once, despite prompting. The head of science, when asked directly about QUASE, said:

*'I've heard of it, but I'm not sure what it is. There are so many different ones. We mainly use CAT [scores]'. She went on, 'we have an annual meeting of heads of department. [The deputy head] produces the graphs. Some people feel threatened and try to justify their subject [results]... Statistics can be used to prove almost anything'. She argued that, 'you have different priorities. Marking might be more important than looking at a statistical report. Statistics go to the bottom of the pile'*.

Talking specifically of her QUASE science scatterplot, the interviewee acknowledged, *'I've not done anything with it'*. Even so, she claimed that the three core departments *'make good use of statistics... everybody is asked to justify their results... this is good – the more this happens, the more they will get used to it... The more we use [value added data], the less we feel threatened [by it]'*.

The evidence here seems to indicate a marked difference of view between senior and middle management over the school's stage of development in the use of value added analyses.

## 7.2.8 SCHOOL E

### PROGRAMME OF VISITS AND INTERVIEWS

Two visits were made to the school, as follows:

1. Initial interview with deputy headteacher.
2. Separate interviews with head of Year 10, head of science, head of upper school.

### THE SCHOOL AND ITS CONTEXT

*Basic information:*

- LEA-maintained comprehensive
- denominational (C of E)
- coeducational
- 825 on roll
- Seven per cent of pupils with English as a second language
- 11 per cent of pupils eligible for/taking free school meals
- 25 per cent of pupils classified as having special educational needs

The school's catchment area, like those of many denominational schools, was very broad geographically and pupils came to the school from over 50 different primary schools. The deputy headteacher described the intake as '*skewed towards lower ability pupils*' and said that the school had a reputation for being 'caring' towards children with special educational needs. At the same time, it had been performing at the lower end of the 'league tables' and experiencing falling rolls, and this was something the senior management wanted to change. How to do so was a challenge, given that the school had been under threat of merger/closure twice in the last ten years and was '*surrounded by high-achieving schools*'. The appointment of a new headteacher with a clear and conspicuous commitment to the Anglican identity and ethos had, in the deputy head's opinion, changed the perception and, as a result, the intake of the school.

The deputy head also talked about a need to '*shift the culture*' internally, in terms of building up a bank of information about pupils' progress, for which heads of year were now responsible; this had created certain tensions with the role of departments. He had personally taken on the task, since his appointment, of encouraging staff to raise their expectations of pupils and use data about pupils' progress/performance more actively – although without a clear view at the beginning of how this might be achieved: he had found some staff '*quite resentful*' about the increasing rigour being demanded of them. There had also been a variety of other new initiatives, for which it had been important for staff to gain a sense of ownership.

## OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

As part of the move towards more rigorous monitoring of achievement, the senior management team was promoting the use of various kinds of performance and assessment analyses. At the time of visiting, the school had two years' of QUASE data and was also using YELLIS, ALIS and NFER-NELSON tests, including CAT (Cognitive Abilities Test). The interviewer was told that the senior management team had provided in-service training on the use of QUASE to heads of department, so that they could disseminate the QUASE graphs to their department teams. The deputy head said that QUASE analyses were attractive because they were understandable (unlike the national value added indicators, which *'looked so crude'*). In the event, however, some staff apparently found the analyses *'threatening'* and several had experienced problems with interpretation. Part of the problem here seemed to be that there was no significant added value in the school's results, and some individual departments came off badly – which went counter to what staff had been hoping or expecting.

It was planned that in the following year the head of post-16 provision would take responsibility for dealing with QUASE analyses (presumably in a more proactively managed way): the member of staff was studying for what he called an MA in 'value added methodology'. This interviewee had interesting views on the relative usefulness of the QUASE and YELLIS datasets in the school context, saying, *'I wanted a polarisation of performance... I need the data to kick people up the backside – but it [QUASE] wasn't actually strong enough'*. The sense of there being different datasets around which could be telling different stories was reinforced by the fact that the head of science said he had created his own system for value added. He added that he had seen the QUASE science scatterplots and thought they provided a useful discussion point, but had not received any in-service training.

Generally, it seemed that QUASE data was being used alongside various other data by senior management and the head of upper school as a means of stimulating the discussion about expectations and pupil performance amongst heads of department. The head of Year 10 claimed, *'it [performance data] actually forces people to think'*, although he also observed, *'the data doesn't show you everything – management need to consider what's really going on in the classroom'*.

The challenge for this school, with so much information being created and disseminated in contrast to what had only recently been the case, was to make the wealth of data which was becoming available have a clear enough purpose for staff to find it manageable, credible and useful.

## 7.2.9 SCHOOL H

### PROGRAMME OF VISITS AND INTERVIEWS

Two visits were made to the school, as follows:

1. Initial joint interview with headteacher and deputy headteacher.
2. Separate interviews with headteacher, deputy headteacher, head of science, head of physics and value added coordinator, head of mathematics, coordinator for special educational needs and head of Year 7.

### THE SCHOOL AND ITS CONTEXT

*Basic information:*

- grant-maintained
- denominational (Roman Catholic)
- single-sex (boys)
- 926 on roll
- Very few/no pupils with English as a second language
- 16 per cent of pupils eligible for/taking free school meals
- 17 per cent of pupils classified as having special educational needs

The school is apparently situated very favourably, both socially and educationally, but the reality is much more complicated. The school has been a Funding Agency for Schools case study and has also been identified as a 'successful school' by the DfEE. There is also an *OFSTED Oscar* on display: there are plenty of accolades for the school in evidence. The school would appear to operate in a very competitive environment. It is in a suburban location, yet over 50 per cent of the school's intake comes from inner city areas. Once a grammar school itself, it was claimed that the school does not have its '*full share*' of pupils at the top academic ability end because of competition from (i) other Catholic schools; (ii) independent schools; (iii) a number of highly selective schools in the borough. The results from CAT scores in Year 7 were 'slightly skewed towards the bottom end'. Nonetheless, the senior managers said that the school had a strong local reputation and the benefits of a stable and committed staff. Many of the staff were said to be ambitious for the pupils and the school.

### OVERVIEW OF HOW VALUE ADDED DATA IS PERCEIVED AND USED

Although the school had been participating in QUASE for at least three years, the deputy headteacher admitted, '*We have not used QUASE as effectively as we could have done: you tend to get overtaken by events*'. It was also evident that he tended to equate the phrase 'value added' with CAT prediction data.

The interviewer was informed that the head of physics had been designated the school's *value added coordinator*, mainly working on ALIS data, but also responsible for digesting the QUASE report. During the course of interview, it was said that the school would probably stop using ALIS that year because much of the information had never been used. The deputy head commented on the importance in this context of *'customising your data'*, partly because, at least initially, people were *'frightened'* by statistics. The interviewee said that he and the headteacher had been actively trying to *'shift the culture'*, so that staff could use performance data in a constructive and non-threatening way: *'staff now know it's not a stick to beat people with'*.

From interviews with staff who were not heads of department, it seemed evident that the main impetus for using data was perceived to come from the headteacher and senior management. These interviewees tended to know more about ALIS and YELLIS than about QUASE (even though the school had had an in-service training session on QUASE); but they were even more interested in and knowledgeable about CAT data – no doubt because of the particular posts they held, as head of Year 7 and head of special educational needs respectively. They felt that that the senior management team, with a strong lead from the headteacher, were *'training us as staff to be more aware of using data'*; but one of them said, *'I wish we had more time to use it'*.

So far as middle managers were concerned, the head of the mathematics department was particularly keen to stress the importance of using value added data *'sensitively and in conjunction with teachers' professional judgement'*. She expressed concern about *'drawing too many conclusions from one or two years' data. We need expertise and years of data'*. Even so, this interviewee said that *'underachievement in Years 10 and 11' was 'a front line issue for staff at Department meetings'*. YELLIS and QUASE and GCSE results were used to assess how well the Department had performed in relation to national results and also within the school compared with other departments, although this was said not to be competitive; *'[to know] whether value has been added is always useful'*. Interestingly, the interviewee was of the opinion that *'mathematicians, scientists and geographers feel happier with data [than staff in other disciplines]'*. By contrast, the head of science claimed that he had *'no direct contact'* with data from QUASE, YELLIS or ALIS and that he relied on the value added coordinator: *'it's nice to have someone there who's on top of it... the information is all there'*.

The value added coordinator confirmed this view when he said *'I'm an information source [for staff] more than anything; I attend the Institute of Education [University of London] courses to pick up the latest ideas... I keep a library of useful value added information'*. He averred that the school was not using value added in a *'formative'* way, but went on to say that they looked at previous cohorts' performance, generally and individually (which is not at all incompatible with using the data formatively). He said that it was important to read the whole of the QUASE Profile in order to get the full benefit from it, including the attitudinal data. But he noted that the full report was seen by the senior management team only; the heads of department were given the Governors' Information Sheet (see Appendix I for example). An

issue was the varying levels of knowledge and confidence amongst heads of department: *'those who are not particularly confident will perhaps pay lip-service to [the data]'*; he had set up a more structured approach to help heads of department get to grips with their data.

It was clearly the interest in value added he had expressed as an individual which had resulted in his formal role; but he felt there were some tensions within this:

*'how much should I be allowed to say, for example, when I am producing a précis of a document? I'm a funnel, the data comes down from on high... Most of the people who have access to such data are senior managers, so there is an issue of confidentiality – I've been uneasy with knowing which departments are underperforming in QUASE...'*

This school was at the stage of disseminating the data in a controlled way to heads of department, and had identified a member of staff who had the enthusiasm, knowledge and credibility with colleagues to take on this role. Unsurprisingly, perhaps, there were still some issues to be resolved in how exactly this should be handled, but the school would appear to be eminently capable of taking the uses of value added (and other) data forward in a constructive way.

### **7.3 Overview of Preliminary Findings and Issues Raised**

#### **7.3.1 Some problems in the basis for fieldwork**

The main findings from the evidence presented in this chapter are, to some extent, concerned with the basic knowledge and use of QUASE data amongst case study schools. As noted in Chapter Six, it had been assumed from the outset of the study that the fact that these schools had invested time (in collecting the data) and money (in the form of the QUASE subscription) in commissioning value added analyses would mean that school staff would be well briefed about QUASE and how it operated. Furthermore, the schools had volunteered to take part in the study which was explicitly about investigating their use of QUASE data: it had therefore seemed fair to assume that there would be a widespread understanding and utilisation of the QUASE analyses in these schools.

In fact, during the time of this particular study (1998–99), case study schools' active use of value added data turned out to be quite limited and the fieldwork visits raised several unanticipated questions about the basic understanding of QUASE in schools. The case study schools were, on the whole, still getting to grips with the data at senior management level – what it looked like, what it

meant, how the results had been arrived at. Sometimes, however, this was not made clear by interviewees, and the researchers were left to infer what was *not* being done rather than discovering what was being done.

Problems which arose in attempting to conduct the fieldwork according to the original research brief can be grouped as follows:

### ***Confusion/conflation of terms***

- At a general level, some interviewees were evidently not familiar with the quasi-technical use of the term ‘value added’ – a commonly-occurring phrase was ‘we are all adding value’; alternatively, it emerged in the course of conversation that some people took *value added analysis* to mean ‘predictions of (current) pupils’ GCSE results’.
- More specifically, it became evident that some interviewees had confused or conflated QUASE analyses (the service provided by NFER) with Cognitive Abilities Test (CAT) data (a test-based service, with optional analyses including GCSE Indicators, provided by NFER-NELSON, a commercial publishing company in which NFER holds shares).
- In this context, several interviewees gave the impression of intending to use Year 7 data chiefly to be able to demonstrate their (assumed) added value, without apparently realising that value added analyses usually work to produce losers as well as winners. Others wanted primarily to ‘predict’ pupils’ GCSE results from such baseline data, rather than, for example, to inform school organisation (e.g. setting), curriculum design and/or pedagogy.
- An impression was given by some staff that, with so much performance and other data (e.g. PANDAs, PICSIs, YELLIS, LEA-generated data and, more recently, the DfEE ‘Autumn Package’) circulating in the school, they did not always or consistently differentiate between one set of statistics and another. This is not necessarily due to negligence or lack of interest on the part of staff. Although QUASE Profiles were designed and written by the NFER team with the intention that they should be self-explanatory and able to be circulated amongst all staff, this is not how the Profiles were invariably used. In at least one school, results given in the Profiles were summarised, anonymised and included in a report alongside other performance data written by the headteacher or other member of senior management. The effect of this was to ‘mask’ any specific QUASE branding of data.

### ***Lack of dissemination of QUASE data within the school***

- In a few schools, QUASE analyses seemed not to be seen, let alone used, by anyone outside the senior management team. The rationales offered for this were, variously, pragmatic (there was only one year’s worth of data for the school) or managerialist (‘staff wouldn’t understand the data’).

### ***Lack of clear purpose for collection/analysis of performance/assessment data***

- Some interviewees were keen to demonstrate how much assessment data they had collected (several years' worth of notebooks of data, in one case), without giving a clear rationale – either in terms of departmental accountability or individual pupil progress monitoring – for doing so.

These features, separately or in combination, made it difficult to talk with at least some interviewees in most schools about the use and impact of QUASE, particularly about how QUASE might have been used to bring about changes in curriculum or pedagogy.

On reflection, one would have to admit that, although services like NFER's QUASE and the University of Durham's YELLIS have been available for some years, there is no compulsion on participating schools to follow a prescribed, or any, course of action in response to the analyses they generate. This lack of statutory force could be a factor in the variability observed.

### **7.3.2 Some preliminary findings and observations**

To summarise the schools' positions at this initial stage of the analysis, it looked as if:

- Schools B, C and F were more advanced than the others in their use of QUASE data for school development – particularly in terms of its impact on monitoring and self-evaluation – although there was still not a fully-fledged strategy that was equally operational throughout all departments;
- In Schools E and H there appeared to be a well-developed understanding amongst the senior management team of how the data could be used in reviewing pupils' performance;
- Schools A, D and G were beginning to find uses for the data in terms of overall monitoring;
- School I seemed to be least advanced in the sense that the data were not being used for school development but rather for promotional purposes.

On the basis of the descriptions there are three general, if unsurprising, observations to be made:

- The visits revealed that schools differed considerably from one another in the degree to which they were making use of QUASE data for developmental purposes. On the other hand, none of them seemed to have fully established, even at heads of department level, a collective strategy for using the analyses to their full capacity – although this was not invariably recognised by the senior managers. I have argued (see Chapter

Three) that the two distinct functions of ‘value added’ are for constructing fairer league tables, or at least a truer picture of a school’s achievements, and for diagnosing pedagogical strengths and weaknesses within the school. On the whole, and in so far as they made the distinction, it was the first of these functions which tended to be uppermost in senior staff’s consciousness at the time of the study. The design of the study had assumed that the second function would be more recognised and developed than it was.

- In the schools where QUASE data was being actively utilised (whether for basic monitoring or in a more developmental way as well), the headteacher was invariably leading, or at least seen to be supporting, its use. Additionally, there was usually a ‘champion’ of value added at senior level who took charge of making sense of the data for senior and/or middle management colleagues (sometimes this person may have taken on the role informally, as a result, for example, of his/her own professional development or academic study). The data was seen as part of the school’s intelligence-gathering rather than a formal accountability exercise.
- The third general observation of note was the variability between departments in the same school, that is, that what was observed of one department in a school was quite often not observable in another, in terms of policies, perceptions and practice around the use of value added (or any other kind of) data. This finding supports the idea – already addressed by, for example, Sammons *et al.* (1997) and Harris (1998) – that variability in departmental cultures and practices within secondary schools may be the norm rather than the exception. There were certainly indications in several of the schools of a fragmentation or else unacknowledged overlap of responsibilities in respect of the related tasks of collecting, analysing and disseminating performance data and target-setting – of course these are large and complex areas of activity which cannot (and arguably should not) be managed exclusively by one person. But the sense in several schools was of a lack of clear policy on handling data and of a rationale for who was doing what: people sometimes seemed to pick up a role through personal interest and commitment. Sometimes this resulted in the individual becoming the ‘champion’ of (value added) data for other staff and sometimes it did not.

### **7.3.3 Revisiting the hypotheses for the study**

In other words, given that one of the main aims of this study was to provide a good empirical description of what happens *vis-à-vis* the deployment of value added data in schools committed to using such instruments, the evidence so far raised a number of quite challenging questions about the perceptions and meanings of (value added) performance data within these schools, and how to understand them; and indeed these could be said to undermine some of the basic assumptions and hypotheses which had informed the design of the study.

Of course, the author's own involvement in the construction of value added analyses for schools was an informing factor in the hypotheses underlying the study. The fact that the author acts in her professional role as an advocate for value added performance data may have led her to expect that school staff would, all other things being equal, share similar views about its uses and benefits – whereas the evidence seemed to hint that for some staff, at least, the data had other functions.

What seemed to be needed, then, were further analytical exercises that would (i) go back to the details held on the NFER's database to see whether there were any clues to be had from scrutinising management information; (ii) generate some conceptual frameworks which, although speculative, might help in understanding what had been observed. The following two chapters undertake these distinct kinds of exercise respectively.

## ***CHAPTER EIGHT***

### **ANALYSIS OF CASE STUDIES II: FURTHER EXPLORATION OF QUASE DATABASE INFORMATION**

#### **8.1 Introduction: Issues for Further Exploration**

The descriptive analysis undertaken in Chapter Seven suggested three areas for further exploration, as follows:

- the relative lack of use of QUASE data for school development and improvement in the schools selected for investigation, compared with what had been anticipated;
- the differences in level and type of use of QUASE by different schools in the study;
- the differences between level and type of use in different departments in the same school.

This chapter explores how far the various kinds of management information about the schools held on the NFER's QUASE database could provide some explanation of these issues. First, background data about the schools was collated which the author surmised might have some explanatory power, as follows:

- school type;
- length of time of schools' participation in QUASE;
- schools' previous take-up, if any, of guidance from the NFER's QUASE team.

These pieces of information were gathered into a school record, presented and discussed in Section 8.2.

Secondly, an exercise was carried out which drew on the expertise of NFER statisticians to analyse QUASE performance data in such a way as to construct a model of schools' 'relative effectiveness'. The author wished to explore whether the existence or not of an added value element in schools' actual performance seemed in any way related to their use of the data.

Thirdly, performance at individual departmental level was examined with the same intention.

## 8.2 Construction and Analysis of School Records

### 8.2.1 The school record

QUASE and other NFER administrative databases were used to collate various pieces of information about each school which have been collated and reproduced here. The record comprises the following information (some of this information has already appeared in Chapters Six and Seven.):

- The letter refers to the school code used in Chapter Seven, and also on the ‘relative effectiveness’ diagram reproduced in section 8.3 below.
- Some basic information about each school was supplied by NFER Field Research Services from NFER’s schools database and was recorded in Chapter Seven; only location, type and size are repeated here.
- QUASE records were searched by a member of NFER Field Research Services for information on whether or not the school had at some time in the past commissioned a Governors’ Information Sheet (example supplied at **Appendix I**), a confidential ‘clinic’ held at NFER for school senior managers and/or an in-service training session held on the school site (usually for senior and middle managers). For the sake of completeness, schools’ additional QUASE options for the 1998 data are also given, although the fieldwork began before these options would have been taken up.

The record for each school yielded the following information:

1. **School A** (south-east England); grant-maintained, coeducational, 1068 on roll; three years’ QUASE data.

Level/type of any additional input from QUASE team:

	1995	1996	1997	1998
Governors’ info sheet	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Clinic at NFER	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NFER visit to school for INSET	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. **School B** (north-east England); LEA-maintained, coeducational, 1113 on roll; three years’ QUASE data.

Level/type of any additional input from QUASE team:

	1995	1996	1997	1998
Governors’ info sheet	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clinic at NFER	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NFER visit to school for INSET	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3. **School C** (north-west England); LEA-maintained, coeducational, 689 on roll; three years' QUASE data.

Level/type of any additional input from QUASE team:

	1995	1996	1997	1998
Governors' info sheet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clinic at NFER	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NFER visit to school for INSET	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. **School D** (Wales); LEA-maintained, coeducational, 736 on roll; three years' QUASE data.

Level/type of any additional input from QUASE team:

	1995	1996	1997	1998
Governors' info sheet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clinic at NFER	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NFER visit to school for INSET	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. **School E** (south-east England); LEA-maintained, denominational (C of E), coeducational, 825 on roll; one year's QUASE data.

Level/type of any additional input from QUASE team:

	1995	1996	1997	1998
Governors' info sheet	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	
Clinic at NFER	<input type="checkbox"/>		<input type="checkbox"/>	
NFER visit to school for INSET	<input checked="" type="checkbox"/>		<input type="checkbox"/>	

6. **School F** (west Midlands); LEA-maintained, coeducational, 1125 on roll; three years' QUASE data.

Level/type of any additional input from QUASE team:

	1995	1996	1997	1998
Governors' info sheet	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Clinic at NFER	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NFER visit to school for INSET	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

*Note:* The school was part of an LEA-wide QUASE analysis in 1995 and 1996; schools were invited to a seminar led by NFER discussing the 1995 results. During fieldwork, it was ascertained that both the head teacher and the head of mathematics attended that meeting, although they did not remember this without prompting.

7. **School G** (south-east England); LEA-maintained, coeducational, 1329 on roll; one year's QUASE data.

Level/type of any additional input from QUASE team:

	1997	1998
Governors' info sheet	<input type="checkbox"/>	<input type="checkbox"/>
Clinic at NFER	<input type="checkbox"/>	<input type="checkbox"/>
NFER visit to school for INSET	<input type="checkbox"/>	<input type="checkbox"/>

8. **School H** (outer London); grant-maintained, denominational (RC), single-sex (boys), 926 on roll; three years' QUASE data.

Level/type of any additional input from QUASE team:

	1995	1996	1997	1998
Governors' info sheet	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Clinic at NFER	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NFER visit to school for INSET	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9. **School I** (south-east England); independent, single-sex (boys), 243 on roll; two years' QUASE data.

Level/type of any additional input from QUASE team:

	1997	1998
Governors' info sheet	<input type="checkbox"/>	<input type="checkbox"/>
Clinic at NFER	<input type="checkbox"/>	<input type="checkbox"/>
NFER visit to school for INSET	<input type="checkbox"/>	<input type="checkbox"/>

### 8.2.2 Discussion

The correlation between use of QUASE data and **school location/size/type** is not strong, although Schools B, C, E and F happen to be LEA-maintained rather than grant-maintained or independent. This may be purely coincidental, since there was no indication from the interviews with school staff that these local education authorities were felt to be giving a strong lead on value added.

There does seem to be some correlation between use of QUASE and **length of time in QUASE**: the three schools which were most advanced in their use of QUASE had been subscribing for at least three years. Conversely, schools with only one or two years' participation were unlikely to be exhibiting widespread use. This is not surprising, since the data is complex and several

staff acknowledged that it had taken them some time to learn how to interpret the analyses and to understand what their implications might be. Even so, the mere fact of subscribing to QUASE for three or more years seems to offer no guarantee of more established use.

With regard to **guidance and support from NFER**, relatively little of this provision had been taken up by these schools. (The precise pattern of take-up by all other QUASE schools over the years is not available, but NFER colleagues' impression is that the case study schools were not atypical in this respect.) Even so, the information indicates that those case study schools which showed greater levels of understanding, at least amongst some staff, were often those where some form of support with interpretation or in-service training had been given in the past; conversely, those which had not had any such input were those which showed least developed use of QUASE. This finding was also reported anecdotally by Williamson and Fitz-Gibbon (1990). Of course, the fact that some schools had commissioned extra input (at additional expense) could indicate a greater level of interest in and commitment to the data in any case. Interestingly, it emerged in interview that two schools which had not had much input from NFER – Schools B and D – had been working with a university (different in each case) on issues to do with value added and contextualised data.

It seems, then, that the database information on the schools' background had some explanatory power with regard to two issues (even allowing for the small number of case study schools involved): that is, the relative overall lack of use of QUASE data and the difference between schools in their use of data.

### **8.3 Exploration of Schools' 'Relative Effectiveness'**

The possibility of a difference between schools in their use of QUASE data had been one of the working hypotheses of the study: Chapter Six (section 6.4.4) argued that schools with different 'starting points' in terms of performance might be confronting different issues with regard to the use of data. The second area of investigation using database information was accordingly the relationship between what the performance data itself showed for each of the case-study schools and the schools' use of data.

Some initial 'value added' information had already been derived by the author from QUASE analyses in selecting schools for case study work. In Chapter Six, it was explained that potential case study schools had been divided into three groups according to whether four or more of their seven QUASE indicators, before adjusting for pupils' prior attainment, were above average, close to average or below average (based on their results in the 1997 analysis); ten case study schools were originally selected to give a spread of both 'raw' and 'value added' results. The results for the final set of nine participating case study schools were as follows:

**School A:** below average raw scores, below average adjusted scores.

**School B:** below average raw scores, average adjusted scores.

**School C:** below average raw scores, average adjusted scores.

**School D:** average raw scores, below average adjusted scores.

**School E:** average raw scores, average adjusted scores.

**School F:** above average raw scores, average adjusted scores.

**School G:** above average raw scores, average adjusted scores.

**School H:** above average raw scores, above average adjusted scores.

**School I:** above average raw scores, above average adjusted scores.

However, this information does not reveal schools' *relative* effectiveness, which the author believed might be a useful analytical tool. Some additional work from the NFER's statistical service was therefore commissioned with the intention of constructing a model of schools' performance in terms of total GCSE score on 'raw' and on adjusted residuals. It might then be possible to discern a pattern linking schools' 'relative effectiveness' as measured in this way with their capacity to make use of data.

Figure 8.1 below was devised by NFER Statistics Department and shows each of the case study schools distributed in a quadrant diagram. The diagram relates 'unadjusted' or 'raw' scores, using total GCSE score as the indicator, to scores 'adjusted' (i.e. statistically controlled) for key factors correlated with performance<sup>20</sup>.

---

<sup>20</sup> In this case, pupils' prior attainment as measured by standardised tests taken on intake to

In the diagram, the triangles represent the case study schools – identified by their code letter – and the circles represent other QUASE schools.

**Figure 8.1 Case study schools (A-I): ‘relative effectiveness’**

This figure could not be converted into pdf format. Please consult the MS Word version of this chapter to view the figure.

An earlier version of this diagram (i.e. without individual schools identified on it), together with an explanation of its general implications, was given in Chapter Six. It may be helpful to elaborate on those implications here:

- Schools in the top right hand quadrant are those performing well on raw scores (and therefore probably in the published performance tables) *and also* on scores adjusted for pupil and school factors listed above. They could be called ‘(highly) effective’ schools. Schools in the top left hand quadrant are schools whose reasonable or even good performance is probably being obscured by performance tables, since they are getting below average raw scores but above average adjusted scores. Schools in the bottom right hand quadrant could be termed ‘complacent’ schools, because although their raw scores suggest above-average performance their adjusted scores tell the opposite story. Schools in the bottom left hand quadrant may well be ‘ineffective’ or even ‘failing’ schools because, whatever factors are taken into account, their pupils are obtaining below average results.

---

the school, their sex, ethnic group and date of birth, plus the overall percentage at school-level of pupils eligible for free school meals, an indication of catchment area and the size of the Year 11 cohort(s). All of these empirical factors have been found consistently to correlate with performance at GCSE: see Chapters Three and Six above.

- One could go further and postulate that schools in the top right hand quadrant first need to identify, and disseminate among staff, what it is they are doing well, what factors have contributed to this and whether it is sustainable. Schools in the top left hand quadrant probably need to do some justifiable morale-boosting, not only with governors and (prospective) parents, but with their own staff. Schools in the bottom right hand quadrant need to take pre-emptive steps to identify and remediate the under-achievement which the data suggests is present. Schools in the bottom left hand quadrant may need much by way of support and intervention: they may have become dysfunctional as learning institutions.
- A general question this raises is what the incentives might be for schools at these different stages of development (if that is an appropriate way to describe them) to make use of value added data. For schools which are already served well by league tables it might be supposed that the incentives are rather weak for making extensive use of value added indicators; but that a particularly committed senior manager might wish to find ways of further increasing the quality/consistency of provision in the school. 'Effective' schools which have below-average raw scores, on the other hand, might be presumed to have urgent reasons for studying value added data closely, in order to be able to demonstrate what the staff may long have believed to be the case, that they are doing well by the pupils they have. For 'complacent' schools, it is only 'value added', as opposed to raw, results which provide any incentive or means for identifying weaknesses: but one might expect that this will not be perceived by all staff as equally valid or valuable. One could predict that, in this context, 'value added' data might be seen as a lever for change by senior staff acting very much in a managerial capacity. For 'ineffective' or 'failing' schools, on the other hand, the hypothesis might be that staff defensiveness might be quite pronounced, and that support from an external agent (the local education authority, for example) would be needed to create and manage a receptive context for the data.

It can readily be seen from Figure 8.1 that the case study schools are distributed over all four quadrants. The following assessment of relative effectiveness at the whole-school level – in terms of total GCSE score only – can be made:

- A Not effective on raw *or* value added score
- B Effective only on value added score
- C Effective only on value added score
- D Effective only on raw score
- E Effective on raw *and* value added scores
- F Effective on raw *and* value added scores
- G Effective only on raw score

H Effective on raw *and* value added scores

I Effective on raw *and* value added scores

Now, it is quite true that the scores in each case may not be statistically significant (the individual schools' tables – not given here – provide such information), and such assessments should not be used as indicators for accountability purposes. Initially, the author looked only at the statistically significant scores and these seemed not to permit any relationship to be posited between schools' QUASE performance and the apparent use they made of QUASE data. However, schools were advised by the QUASE booklet that they could treat non-significant data as potentially indicative from a diagnostic perspective and so the author decided to look at non-significant data as well. When schools' QUASE performance data (both statistically significant and non-significant) on total GCSE score was set alongside the fieldwork evidence, it *was* possible to perceive some positive relationship between these broad assessments and observable school policies and practices. What one would have predicted from the hypotheses suggested above is that Schools B and C would have the strongest incentive for using value added analyses and would be 'ahead of the game'; that Schools D and G would have the weakest incentives and thus slower at developing the use of value added data, unless driven by a strong senior management agenda; that Schools E, F, H and I would need a senior management team highly committed to using the data for improving quality even further; and that School A might be having some difficulty finding a way into the use of value added data at all.

As can be seen from the data presented in Chapter Seven, this predicted pattern is not unrecognisable in the picture given by the fieldwork data; for Schools E, F and H, for example, the relationship holds fairly well:

- School E had a senior management team who were well-briefed and enthusiastic about QUASE. In School F, the headteacher took a strong interest in the monitoring of school and departmental performance and was promoting the use of value added approaches. In School H, the senior management team had been actively trying to change the culture, so that staff could use performance data in a constructive way for school improvement.

But the correlations are not exact nor complete, and the findings should certainly not be over-interpreted. An unequivocal and complete correlation might indeed have been surprising, for the following reasons:

- the fact that schools had been participating in QUASE for differing lengths of time must, as has been demonstrated, be taken into account as well;
- total GCSE score may be an insufficiently sensitive indicator of 'effectiveness': other or additional indicators of effectiveness might be needed, including those at departmental level, in order to explore the relationship between performance and practice;
- there may be a time-lag between observed practice and translation into pupil performance outcomes;
- we do not (yet) know the precise nature of the causal relationship between measured effectiveness and school practice: that is, we do not know if the assessment of effectiveness which can be construed from the performance data is the direct and deliberate result of identifiable 'school improvement' practices (or lack of them);
- and of course the use (however well established) of data (however sophisticated) may be necessary, but not sufficient, to produce good performance results. The correlation might typically be a weak one. The cogent question would be therefore be, what other practices which result in effective performance does the use of data lead to?

#### **8.4 Exploration of Departmental Effectiveness**

Given the finding that departments within the same school varied in their use of QUASE data, the QUASE database was also used to investigate whether there was any correlation between performance outcomes and the use of QUASE data at departmental level. (Examples of the kinds of tables, bar charts and scatterplot diagrams presenting department-by-department data which the schools received are given in Appendix A, and an explanation of the basis on which this information was compiled and presented is contained in the QUASE commentary, given in Appendix B.)

It should be noted that, contrary to what might have been expected, the author and her colleague did not use this subject-level data in their interviews with heads of department in the case study schools because:

- the author initially assumed interviewees would be familiar with the data (since the QUASE commentary contained guidance notes on how headteachers could use the data with middle management colleagues);
- in cases where it became evident that the data had not been disseminated to heads of department, the author did not think it was appropriate for an outsider to take the initiative in showing them the data;
- in cases where the data had been seen but apparently not understood, the author did not wish to be drawn into an explanation or justification of the data (this being a different professional role).

The information held on the database concerning the subject-level performance in the case study schools showed the following:

### **School A**

On the basis of three years' data, performance was *better* than 'expected' in:

- mathematics
- science
- English

Performance was *worse* than 'expected' in:

- languages

### **School B**

On the basis of three years' data, performance was *better* than 'expected' in:

- geography
- history

Performance was *worse* than 'expected' in:

- science
- English

### **School C**

On the basis of three years' data, performance was *worse* than 'expected' in:

- geography
- mathematics

### **School D**

On the basis of three years' data, performance was *better* than 'expected' in:

- science
- computing and ICT

### **School E**

On the basis of three years' data, performance was *better* than 'expected' in:

- technology
- history
- humanities

### **School F**

On the basis of three years' data, performance was *worse* than 'expected' in:

- mathematics
- art
- physical education

### **School G**

On the basis of two years' data, performance was *better* than 'expected' in:

- mathematics
- computing and IT

### **School H**

On the basis of three years' data, performance was *worse* than 'expected' in:

- science
- technology
- geography
- physical education

### **School I**

On the basis of two years' data, performance was *better* than 'expected' in:

- mathematics
- English
- technology
- geography
- art

It perhaps needs to be explained that QUASE subject-level performance was calculated relative to pupils' total GCSE score (as well as gender, age and ethnic group); this gave the 'expected' level of performance. One of the consequences of using this model was that significantly low total GCSE scores overall would tend to give an artificially elevated view of individual subject performance and *vice versa*. The 'better than expected' performance of departments is thus remarkable in School I (which was already achieving higher than expected performance overall), but not so extraordinary in the other schools, particularly Schools A and D, which had a significantly low overall performance. The 'worse than expected' performance of several

departments in Schools F and H needs to be similarly contextualised: pupils' overall performance was higher than expected (although in School F's case not by a statistically significant amount), and so these departmental results should not be interpreted as signs of serious weakness. Of course, this gloss on the statistical data would need to be understood by school staff: although the QUASE commentary (Appendix B) explains such features, it emerged that few heads of department had worked through or had sight of the text.

So far as exploring a correlation between performance outcomes and use of data is concerned, the following observations can be made:

- The heads of the five departments in Schools A and G which had significantly better than expected QUASE results evinced enthusiastic use of performance and assessment (although not necessarily QUASE) data.
- However, two heads of department who exhibited relatively advanced use of QUASE data had significantly worse than expected QUASE results (geography in School C and mathematics in School F).
- It is not possible to make any observations about correlations between qualitative data and departmental performance in Schools B, E, H and I because no interviews were conducted with the relevant heads of department (on the grounds that they had not had sight of QUASE data).

This seems to indicate that there may be a tentative positive relationship between the use of performance data for school development and performance outcomes – although of course it is impossible to attribute cause and effect.

## **8.5 Summary and Conclusions**

The additional *post hoc* use of the QUASE database information to carry out an analysis of the qualitative data has arguably provided some partial explanation of what was being observed during fieldwork. The implications can be drawn out as follows:

- *In combination*, schools' 'relative effectiveness' (in terms of total GCSE score or overall academic performance), their length of participation in QUASE and some previous in-service training from NFER on the results seem to correlate with the fieldwork findings.
- The question of schools' different degree of incentive for using value added data for school improvement may be relevant to the policy agenda; the evidence could be taken to imply that further encouragement and challenge from, for example, the local education authority may be

necessary for some schools to use such data more fully (see Chapters Ten and Eleven for further discussion on this).

- Departmental differences in the use of QUASE data within the same school *may* be linked to differences in the performance of departments, but this is by no means a secure finding and attribution of cause and effect is hard to make.
- The fact that in-service training seems to have had some impact on schools' capacity to make use of QUASE data for school development is worth noting. One might consequently argue that value added analyses as exemplified by QUASE are not self-explanatory, despite the QUASE team's attempts to develop and refine the feedback reports over the years. These findings may also imply that data analyses generally, unless accompanied by explanations and support suited to the particular needs and circumstances of the staff involved, may not be used for school development purposes or used in ways which are not intended.

The next chapter tries to take the exploration of the qualitative evidence further, by developing and using a more speculative type of analysis.

## ***CHAPTER NINE***

### **ANALYSIS OF CASE STUDIES III: NARRATIVES OF SCHOOL AND DEPARTMENTAL LEADERSHIP**

#### **9.1 Introduction: the Task**

The three chapters which analyse the fieldwork data – of which this is the third – follow the chronological and logical route taken in trying to make sense of a somewhat problematic set of initial findings. The analyses presented in Chapters Seven and Eight have so far permitted five factors to be identified which seem, in combination, to be related to schools' utilisation of QUASE:

- active support for value added by senior management;
- 'championing' of QUASE data by a senior member of staff;
- length of time of the school's participation in QUASE;
- previous exposure of at least some staff to guidance on the interpretation of QUASE;
- the degree of incentive or impetus for action provided by the school's overall 'relative effectiveness'.

These are arguably useful conclusions to have reached, particularly from a policy perspective. But the analysis undertaken so far by no means exhausts the possibilities offered by fieldwork data. Thus, after an initial section (9.2) defining the task more closely, this chapter is concerned to devise an exploratory framework for better understanding interviewees' perceptions of and attitudes towards QUASE data (however they understood that nomenclature). The chapter argues that the fieldwork data allows their accounts, as given in interview, to be construed as constituting a series of 'narratives' around leadership in school development (in the case of the senior managers) and around leadership in teaching and learning (in the case of heads of department).

Because this turned out to be a more subjective exercise than the analysis undertaken in previous chapters (and thus uses the first person voice throughout), the chapter explores only the author's own case study schools: I did not have the confidence that I would be able to use a colleague's fieldnotes in the same way as my own.

## 9.2 Taking the Analysis Further: Some Issues and False Starts

Chapter Seven provided a descriptive analysis of the fieldwork data and Chapter Eight used NFER database information to elicit some explanatory factors in relation to the broad findings. Yet the data collected during fieldwork, and alluded to in Chapter Seven, was suggestive of deeper or more complex issues to be explored, including the fact that different staff in the same school could hold radically different views about the value and usefulness of the data. It will be remembered that Chapter Five argued that the issue of how data is perceived and accommodated by managers and teachers was relevant to an explanation of what happened in a cluster of schools where value added and other performance-related data was introduced; and it was suggested that such issues might need to be taken account of in this study, too. At that stage in the design, I had assumed that this would help to explain a range of active uses of QUASE data: whereas, in fact, the findings from the fieldwork mean that part of the purpose of this chapter is concerned to explain the rather variable and even low levels of use. (I take some comfort from the words of Patton [1990, p.143]: ‘Always be suspicious of data collection that goes according to plan’.) This is a somewhat different enterprise for which a more speculative approach may be better.

What insights might be gleaned from a more detailed, open-ended and disaggregated examination of the qualitative data? Of course, it is important in embarking on such an exploration to keep to what can reasonably and relevantly be inferred from the actual data, that is, interviewees’ accounts as I recorded them, together with my additional contemporaneous observations – and therein lies an immediate difficulty for the construction of an adequate explanatory framework.

On the one hand, I had a strong sense that what I was seeing and hearing had a great deal to do with the tensions and interactions between institutional and departmental cultures, individual teachers’ professional identities, and external pressures from the policy agenda on performance. So in the course of trying to find a more insightful route through the fieldwork data I consulted several texts in two areas: (i) the literature which discusses how teachers identify and function as professionals; (ii) the literature which investigates how schools function as micro-political organisations with distinct cultures and sub-cultures.

At the same time, however, the type of data we had collected imposed severe constraints on pursuing these lines of inquiry. The main focus of the study as originally conceived was, of course, the use of QUASE data for raising attainment. Thus, although I found it both necessary and desirable to draw on interview material which was *not* directly related to the use of QUASE data<sup>21</sup> – of which, as it has been explained, there was a substantial amount (see Chapter Seven, section 7.2) – the research instruments were not designed with a view to studying teacher attitudes in the round, still less school or departmental ‘culture’ either in general or in depth. In particular, as noted above, the main fieldwork instruments were interview schedules and the material available for analysis therefore consisted chiefly of interview data. In interview, people may say things they do not mean (that is, would not put into action), or for a particular effect, or in ways which are open to different interpretations. I recorded in the fieldnotes the provisional judgements I made about interviewees’ meanings from the normal range of conversational cues (such as irony, hostility, intensity, humour); and also noted where interviewees’ accounts differed from each other or left unanswered questions. Having said that, there was no capacity within the project time-frame to test or supplement this data by going back to the schools to gather in-depth observational or ethnographic data.

So, although the micro-political literature in particular seemed to have much explanatory potential in respect of the issues I suspected were raised by the fieldwork, fitting the data available to this study into such a conceptual framework turned out not to be feasible and the idea, to my regret, had to be abandoned. That body of literature is, however, discussed in Chapter Eleven below because of its fruitfulness for other possible studies in this field.

What has been attempted instead in this chapter is more modest, and more appropriate and closer to the actual data. The inspiration for the analysis was partly subliminal but can perhaps be explained as follows. For me, the experience of reading and contemplating the interview material over a period of time was like learning to tune in to waves of narratives which persisted, had their proper existence indeed, before and after the particular interview ‘event’.

---

<sup>21</sup> This means there are two connected issues to be explored: (i) the difference in the understanding and use of QUASE data between different schools and different departments in the same school; (ii) the wider question of the range of perceptions of and attitudes towards value added and other performance data, and how this might be understood.

By ‘narratives’ in this context I mean something like *‘the stories people go on telling and revising (but never complete) as if to an ideal (i.e. detached but sympathetic) audience of what it is they do (in life) and why they do it’*. It could be argued that an interview, like any other invitation to self-reflection, intervenes in, and therefore to an indeterminate extent changes, those stories.

The outcome of this was that it seemed to be possible to interpret the content of the interviews as (partially articulated) expressions of beliefs and values about a number of issues still being actively worked through at both personal and professional levels by the individuals concerned; and then to construct a tentative framework for differentiating, at least to some extent, the beliefs and values held by different (groups of) interviewees. The challenge throughout this chapter is thus to offer convincing insights into the uses of QUASE and other performance data analyses by going beyond or behind the immediate face value of the accounts, whilst not being tempted into an over-interpretation of the data. In the case of the middle managers – for whose interview data I constructed a schematic matrix (see Section 9.5 below) – I was careful to test whether all 15 cases fitted into such a framework.

In attempting to construe the interviewees’ accounts as kinds of narrative, I have assumed that it is important to differentiate between the different levels of management hierarchy, since the impression obtained in the fieldwork visits was that headteachers had a different perspective from heads of department on the significance of QUASE. The following sections accordingly organise the material by structural position, rather than (as in Chapter Seven) by individual school, as follows:

- headteachers (9.3);
- other senior managers (9.4);
- middle managers i.e. heads of core/key departments (9.5);
- other teaching staff (but bearing in mind the paucity of data here) (9.6);
- governors (9.7).

Summaries of the salient points in the interviewees’ accounts can be found in Chapter Seven, under Schools A, F, G and I respectively.

### **9.3 Headteachers: Narratives of Leadership and School Improvement**

The discussion in this section is based on accounts given in interview by four headteachers (three males, one female).

These interviews, taken as a whole, gave a strong sense of the powers and responsibilities headteachers are required to exercise in the cause of school development and measurable improvement. These headteachers seemed almost impelled to construct a 'narrative of school development' – in which they were often both narrator and protagonist – in order to make sense of these responsibilities, in particular the need to deliver measurable improvements in pupils' performance. In the course of the interviews, I found myself playing to the fact that each of these headteachers seemed to appreciate having an alternative kind of audience to the usual one(s) and being able to tell their story to someone new. Within this, there were some interesting themes and variations.

Each of the headteachers presented a construction of the school's recent history and development which arguably had a bearing on the use made of QUASE data. Unsurprisingly, in every case the senior managers' longer-term agenda was to identify and address 'under-achievement': sometimes this was expressed in terms of pupils (particularly boys) under-achieving and sometimes in terms of departments. The surprising thing – given that these schools were all subscribing to the NFER service and had volunteered to participate in the fieldwork study – was the relative lack of reference to specific QUASE data in headteachers' own accounts of tackling under-achievement.

This may be better understood in the context of the 'educational market' brought about by open enrolment and local management, where information about a school's performance which is in the public domain has to be actively managed; each of the headteachers clearly felt that s/he had a key public relations function to fulfil, which it seemed largely revolved around managing published performance data and OFSTED inspection reports. QUASE data could be of interest in such a context: the headmaster of School I, for example, considered that his job had been to raise expectations and the school's reputation so that it became a first-choice school. He was delighted that

QUASE data had so far vindicated his view that the school was doing very well with the boys who had chosen to come here.

More generally, in a market context which is also characterised by high visibility and accountability of schools, a headteacher's role to some extent, if only as a way of leading from the front and of maintaining staff and parent morale, involves 'talking the school up': and this propensity was manifested in the interviews (as can be seen in Chapter Seven). Since QUASE data in all but one of the schools revealed apparent under-performance in one or more departments, it may not be surprising that the headteachers steered clear at least initially of discussing the data in interview. In one case, I asked twice about teachers' expectations of pupils in view of what the data showed – in effect, whether expectations were appropriate/high enough: but the headteacher chose not to answer.

One strategy in the narratives of leadership of three of the headteachers (one female, two male) was to identify themselves and their decisions with the institution, in a spirit of *l'école c'est moi*: this can be seen in some of the comments reported in Chapter Seven. Another strategy (from two males) was to use the language of delegation and collegiality, in which large areas of policy and decision making as well as day-to-day management were said to be devolved to senior colleagues: although this could as easily be a way of 'buying staff into' the headteacher's thinking.

At least some of these headteachers found the autonomy of heads of department (the '*robber barons*', in one headteacher's words) a problem which they felt they needed to address in the light of the differential subject performance: this was an issue which usually had to be surmised from what was merely hinted at in interview. Whilst only one of them openly admitted that his heads of department were a problem he was having difficulty managing well, each of them had something non-specific but fairly critical to say about the differential performance of departments in their school and often alluded to aspects of their cultures as well. Although they probably did not think it appropriate to share the detail of their plans with me, almost all the headteachers were envisaging the need to take action *vis-à-vis* one or more heads of department, ranging from '*kicking butt*' to '*moving some people along quite radically*'.

An intriguing thing to emerge from the interviews – although again this was often something which had to be inferred rather than heard – was the range of evidence which headteachers indicated they were deploying in implementing their narratives of improvement; some of this information was ‘headline’ data (figures appearing in the published performance tables), but a substantial amount of the evidence at their disposal would have to be classified as professional hunch or instinct; only a relatively small proportion of what they seemed to find valuable as managers could be said to consist of rigorously-analysed, i.e. value added, performance data. For example, one headteacher claimed that the most appropriate use of performance data at this stage of the school’s evolution was only ‘*broad brush*’, that it was pointless getting into the fine detail. His main objective was to ‘*beat the drum*’, not to ‘*depress and overwhelm*’ staff with masses of data.

This discussion perhaps illustrates the need to understand better the purposes – which are not always about rational solutions to clearly-defined problems – for which headteachers need information as leaders. In particular, a clearer idea may need to be developed about what incentives there are for a headteacher in his or her particular circumstances, including the degree of (perceived) competition with other local schools, to make active and conspicuous use of value added as distinct from other data.

#### **9.4 Other Senior Managers: Narratives of Counterpoint?**

The discussion in this section is based on accounts given in interview by two senior managers in School A, the senior teacher in School G and the deputy headteacher in School I.

An intriguing impression given by these interviews was their tendency to encompass a commentary or gloss on the headteacher’s strategy, as if this was something which needed interpretation or mediation. This could be true for quite practical reasons – in a school where the headteacher is a somewhat remote or pre-occupied figure (for reasons discussed in the section above) other senior managers may well find themselves playing a mediating role for colleagues, in order that decisions which have been made or are planned are conveyed personally as well as in written minutes of meetings, for example. But there are other reasons why mediation and interpretation may come to be required; one of these is a sense by the headteacher that there is too great a

gap between middle and senior management. This seems to have been the rationale for the creation of senior teachers' posts in School G, although it was not clear that the gap was being narrowed as distinct from merely occupied.

Another reason may be a feeling on the part of senior staff that the headteacher's thinking which has, for example, informed school development planning has not been fully articulated or that staff have not fully understood the processes. As the deputy headteacher in School I said, the development of the school was all *'in his [the headmaster's] head'* and she took it upon herself to act as a kind of go-between – perhaps, it may be thought, to clarify the development processes for herself since she indicated that she did not always feel privy to the headteacher's plans. She nonetheless admired his approach to headship; she said he was *'the heartbeat of the school'* and characterised his role as being like that of the Robin Williams character in *Dead Poets' Society*; she thought that headship conceived like this in terms of 'leading creative individuals' was probably the opposite of data-driven managerialism: *'the school functions on freedom, not on telling individuals what to do'*. The potential of value added data to play a disruptive rather than constructive role in an environment like this – where the individual progress of boys is extremely carefully monitored in any case – was obviously at the forefront of this deputy headteacher's mind. She said that she still needed to decide *'what I am asking of the data'*.

An interesting example of differing narratives at senior management level was in School A, where two members of senior management (in addition to the headteacher) were interviewed. They both talked about how QUASE data could contribute to the strategy for raising standards. One of them claimed that the most valuable aspect of QUASE was the external, impartial evidence of under-achievement; in his view, QUASE was potentially the most powerful lever for change in the insularity of the school – the difficulty with this notion was that the analyses were not given to departments. He considered that the main need was to ensure that the various policies on raising achievement (including QUASE) were enacted consistently and coherently in each department: the major problem was fragmentation. But he characterised the headteacher's approach as *'Byzantine'* – i.e. letting departments continue to act with relative autonomy. The view of the other senior manager was that the school needed to establish the use of QUASE data for target-setting, as a way of talking about pupils' potential based on rigorous data: he said that

QUASE data was given to departments but it had been left up to them how they used it. The problems was that an internal seminar on value added data two years ago had left a lot of staff *'bemused'*; now, the danger was that staff were using data *'too rigidly'*. He said that *'statistics overload'* was a danger for senior staff who had been spending large amounts of time generating and interrogating data: the implication being that heads of department needed to take more of this work on. The accounts given by the two interviewees are not mutually exclusive (except for the odd disagreement about whether QUASE data was disseminated or not): indeed, they have points in common. But their relative emphases evince different 'readings' of the main issues and priorities for development.

The question relevant to this study (but incapable of being answered by the data) is: what role might these 'commentaries' on school development have been playing in creating, as well as reflecting, the micro-political context in which performance data was being used or not? It might be presumed that the institutional 'space' below the headteacher – and how far the power/influence associated with it is perceived as being occupied by merit or mere position – is one of the aspects of the environment which controls how heads of department and other middle managers operate when they meet or act together. This is then one area worthy of further exploration in future research.

### **9.5 Middle Managers: Narratives of Leadership and Teaching**

The discussion in this section is based on accounts given in interview by 15 heads of department in Schools A, F and G.

For people at middle management/head of department level, performance data seemed potentially to have a stronger role than at headteacher level, in that several heads of department talked at length about their use of such data, including what they called value added data. Since there was a tendency, as already indicated, to confuse QUASE data with other kinds of performance data analyses, it was not always possible to know with certainty whether the interviewee and author were discussing the same thing. Rather than focus exclusively on QUASE data, therefore, the following discussion encompasses what interviewees said about performance and assessment data more generally as well.

Analysing what emerged from these interviews was not straightforward for other reasons, too. Underlying the various descriptions given by different interviewees of their use of data seemed to be a range of feelings about the proper place of data in their own and their colleagues' teaching, which were only occasionally made explicit. (Indeed, a difficulty for researchers in this field is that much of what teachers believe about teaching may well have to be inferred from things they say in passing or interpreted from their actions, largely because life in school does not give many opportunities for verbalised reflection and overt discussion of that kind.)

It is possible to represent those *emotional* attitudes towards performance data along an axis ranging from 'cold' to 'hot' – by this I mean to suggest the degree of enthusiasm for the potential contribution of data to teaching and learning.

**Figure 9.1**

'cold' \_\_\_\_\_ 'hot'

For several interviewees, the use of test/examination data to arrive at an assessment of pupils' ability was experienced primarily as an externally-imposed requirement. This attitude was often justified by reference to those aspects of performance data analysis, assessment regimes, etc., which are less than ideal. And this in turn was sometimes linked with an implicit or explicit appeal to 'professional judgement' in preference to data – *'I know my kids'*; *'my job is to teach, not analyse data'*; *'I'm interested in science teaching, not the school's GCSE results'*; *teachers aren't looking over their shoulders at performance data [said favourably]*. Whilst a 'cold' attitude towards data was in some cases linked to an acknowledged ignorance (often ostensibly of other colleagues: *'QUASE is beyond the comprehension of most staff here'*) or a lack of confidence, in others this attitude was at least partly dependent on a prior personal/professional system of alternative values and in particular on the view that central government was imposing an inappropriate model on education (see, for example, Fielding, 1999, for a scholarly exposition of this view). Some of the interviewees who seemed 'cold' towards the use of performance data punctuated their conversation with lively anecdotes about pupils as people and gave the impression that they saw teaching as an essentially individualistic activity revolving around relationships. There were

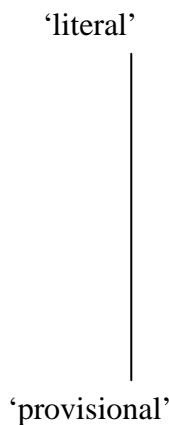
other, more tactical, reasons for 'coldness' towards data, for example, an anxiety about the potential identification of 'failing' teachers.

So far as QUASE data in particular was concerned, an issue for almost all the heads of department was that they had, quite understandably, a focus on current pupils and their future performance; for a few weeks in the autumn term, they might spend time thinking about last year's Year 11. The problem was that the focus of value added analyses like QUASE is necessarily on past years' performance; several heads of department tended to dismiss the data for that reason as being 'historical': *'it hasn't changed what we do'*.

At the other end of this spectrum were those interviewees who enthused about the use of data and were able to talk at length about the systems and procedures they had devised and/or the analyses they had carried out. Sometimes this activity seemed to offer a further elaboration of 'headline' performance data analyses, 'playing the numbers game'; sometimes, by contrast, it was a deeper investigation of teaching/learning issues and problems, for example, how to identify and tackle Year 11 boys' difficulties in the written component of a modern foreign language examination. However, sometimes it looked like the collection and recording of large amounts of data for no particularly clear reason. One or two interviewees gave the impression of being isolated enthusiasts in their school, hoping by dint of argument and reams of data analyses to convert their counterparts in other departments to what they saw as a more rational system for understanding and evaluating pupils', and their own/their colleagues', performance. [My fieldnotes for one interview record my rather unprofessional exasperation with the elaborate system being explained to me: *'this man is a mixture of the perfectly reasonable and the completely insane'*.] One or two others saw their role rather as taking the lead with colleagues within, and sometimes beyond, their own departments to develop the use of data analyses at individual pupil, group and departmental level to review curriculum design and delivery. It seems likely that the effectiveness of the flow, between Key Stages and subjects within the school as a whole, of information about pupils' achievements (touched on in Chapter Seven), influenced the degree to which individuals could, or felt they must, take on such a role.

Even so, the ‘cold’ and ‘hot’ axis by itself did not seem sufficient to differentiate the range and complexity of attitudes and approaches which had emerged in the fieldwork. So I tried putting in another axis to represent a range of *intellectual* stances towards the use of performance or assessment data, with ‘literal’ use at one extreme and ‘provisional’ at the other – by this I mean to suggest the degree of reliance on data as a manifestation of pupils’ ability in a given subject/skill.

**Figure 9.2**

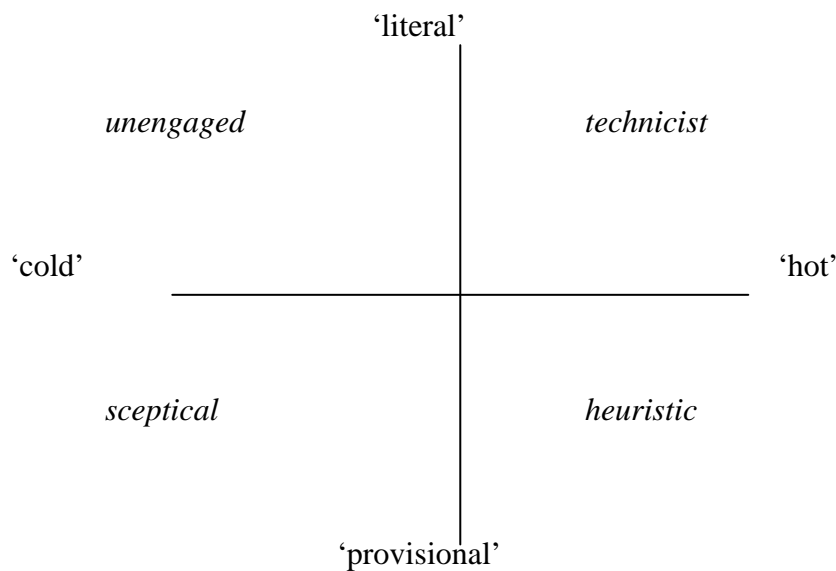


When I asked about how people dealt with any anomalies between results arising from different datasets/analyses, it was clear that for some interviewees this was explicable in terms of the possible inaccuracy of one or more of the datasets; whereas for others it was something to be expected from the nature of data. Moreover, some interviewees appeared to regard performance/assessment data as possessing an intrinsic truth about departmental performance and/or pupils’ ability. One head of department said what he most needed was *‘the statistics which tell me how I am doing on a weekly basis’*. One or two heads of department believed that pupils’ future performance could be safely predicted if only the system could assemble and analyse the right kind of data now. At the other end of this particular spectrum, a few interviewees regarded performance data as necessary but not sufficient, that is, useful and insightful only if professional judgement is brought to bear on it. A head of department in one school said, *‘the data is a catalyst for understanding the problem’*. Another head of department argued: *‘there is no way you can bring together the SATs results, the CAT (cognitive abilities test) scores, the grade based on homework and your own professional judgement about a pupil into a single target. The point is that the discrepancy between all these pieces of information can tell you something really*

*important – about your expectations of the pupil, for example, or her expectations of herself.'*

By putting one axis on top of the other, I created four quadrants which can be used to describe different combinations of intellectual and emotional responses to the use of data; I have labelled these combinations *unengaged*, *technicist*, *sceptical* and *heuristic* respectively:

**Figure 9.3**



Constructing a quadrant diagram makes at least some allowance for the fact that the attitudes and responses of heads of department were not easily ranged along only emotional or only intellectual axes. These quadrants can be further elaborated as follows:

- The **'unengaged' position** was characterised by an apparent resistance to taking the initiative in making use of data, often because the data was perceived as being 'out there' and not intrinsically relevant to pedagogical needs. The member of staff made use only of selected pieces of data either for idiosyncratic reasons or because required to do so by senior management. A pretty unambiguous example was a head of department who said: *'None of this [value added] is any use to me. After thirty years in the profession, I know what my pupils are capable of, I don't need bloody number-crunchers churning out reams of meaningless figures.'* This approach tended to place the teacher's professional judgement and the data in opposition to each other. A different kind of example was a head of science who thought it was not his job to collect and analyse data; as someone with a background in the private industrial sector, he saw himself as an operational manager and considered it was senior management's job,

not his, to set the required targets. For some other interviewees, their resistance decreased over time; their initial rejection of the data therefore appeared to stem more from their relative lack of knowledge and/or confidence about data analysis and its interpretation than from a set of principles or underlying assumptions.

- The **‘technicist’ approach** was one in which a great deal of enthusiastic reliance was placed on performance data as the key to monitoring and evaluating pupils’ performance. Such data was seen as problematic with regard only to its accuracy; its meaning and interpretation were largely taken at face value. This approach was often characterised by a dedication to gathering and recording large amounts of data at the individual pupil level (although sometimes without a clear purpose for this); it was also often allied to an expressed wish to be able to find a formula with which to ‘predict’ pupils’ future performance. An example of this approach was the head of department who advocated the use of a simple spelling test for all pupils in Year 7 which could be used by all departments as predictive of pupils’ capacity to obtain higher grade GCSEs. A departmental head in another school wanted to be given information which would firmly predict pupils’ GCSE grades from their Key Stage 2 and 3 results, and found the notion of ‘chances graphs’ (which allow for the inexact fit of any regression line to individual data) *‘just an irritation’*. This approach tended to put the spotlight on pupils and what they are (or are not) considered capable of, rather than on the curriculum or the pedagogy.
- The **‘sceptical’ approach** was actually marked by a resistance to the literal use of data rather than to a rejection of data *per se*. The difference between the ‘unengaged’ and the ‘sceptical’ approaches was in the relative sophistication with which proponents viewed data. An extreme example was a head of English who said at the beginning of the interview, *‘I don’t do sums’*. But it became clear that this was not the real point; in disagreeing with the senior management in the school, he had taken care to think through his arguments carefully, for example: *‘Statistics can tell you what you have done, not what you will do – unless all the conditions stay the same [implying, which never happens]’*. The point was that he had a strongly held set of values about teaching which were opposed to ‘the current orthodoxy’. He claimed that the professionalism of teachers was being eroded by the misapplication of figures, and argued that using prior attainment data to set target grades for individual pupils was *‘immoral and without foundation’*, because any individual child could always surprise their teacher by doing better than anyone anticipated. He said that his relative immunity from being required to set the prescribed targets was due only to the good performance of the department so far; although he believed this record depended on the quality of subject teaching and relationships with pupils much more than on performance monitoring. However, when pressed he acknowledged that he himself used performance data discreetly to monitor the performance of the teachers in his team. The headteacher’s view, misguidedly but perhaps not surprisingly, was: *‘there’s nothing much going on in the English Department, is there?’*

- The **‘heuristic’ approach** was one which accepted and valued performance, and especially value added, data; the data was used for raising questions rather than making judgements, however, and crucially these were strategic questions about how well the curriculum and pedagogy were meeting pupils’ needs. An example was the head of a mathematics department who reported scrutinising data from a number of different sources including QUASE. It became clear to him and his departmental team that, although the OFSTED inspection report had suggested the department was under-achieving with the top sets at GCSE, the scatterplot data told a different story. It indicated that the third sets had been under-achieving and that this was a consistent pattern over three or so years. Putting their views and impressions together with the data, the staff concluded that the current scheme of work was too dense and demanding for pupils in Set 3 and that in consequence these pupils had come to believe they were no good at mathematics. The head of department and his colleagues re-wrote the whole scheme of work to make the curriculum more lean and transparent. This approach acknowledges that data cannot ever be predictive, in the strict sense, of pupils’ future performance; but nor does it have to be perfect, i.e. totally valid and reliable, in order to be useful. People with this approach could handle the data confidently, so that it made sense of what they wanted to achieve, but also with respect for its limitations. One head of English and her deputy studied performance and assessment data regularly to make sure not only that they had assigned pupils to the appropriate bands but also that these organisational arrangements were not turning into self-fulfilling prophecies for pupils in the lower sets. They were concerned that an over-reliance on test data might mean that some pupils’ potential was never picked up. The phrase they used was, *‘we are always ready to fan the spark when it shows.’* Interestingly, this approach often seemed to be allied with a communicated sense of confidence in the department’s work and also a view of the subject as something worthwhile to be shared with pupils. Even so, these interviewees recognised that the developmental and strategic use of data involved serious challenges for the rest of the departmental staff in terms of time and understanding, and they saw themselves as needing to lead by example and *‘give direction’*; or even just to accept that it would take some time for colleagues to be able to use data in this way.

I believe this classification is useful for analytical purposes, because it suggests that the head of department’s own beliefs and values about teaching and learning are highly implicated in his or her narratives of departmental leadership and the uses made of performance data within and by the department.

However, as an important rider, I would have to say that the evidence presented in Chapter Three above strongly suggests that the ‘heuristic’ approach is the most rational and defensible. One of the interesting points to emerge from this small sample is that no particular subject discipline seemed

to confer a prerogative on the 'heuristic' use of data (in contrast to what Williamson *et al.*, 1992, found): one might have assumed, for example, that heads of mathematics – and, possibly, science – departments would be more likely than heads of other departments to welcome and utilise value added measurements. But this was not the case: interviewees in each of the core subject departments were distributed fairly randomly over the four quadrants. This leads one to postulate that an area worthy of further development is the connection between the kind of material presented here and the work on departmental cultures being undertaken by, for example, Busher and Harris (1999).

In fact, the description and analysis of departmental functions and cultures, as distinct from those of senior management, in schools has only recently begun to be the focus of academic (and policy) concern. Busher and Harris, for example, raise the question of whether heads of department are primarily 'subject leaders', providing professional leadership, subject skills and knowledge, and exemplifying reflective practice, or 'middle managers', responsible for implementing senior management policy, line-managing staff and mediating organisational pressures upwards and downwards. Presumably a key issue is the extent to which heads of department feel in control of either of these roles. In this context, performance data might operate as a tool for reflective collegial practice or might be perceived as a weapon wielded by senior management; the relationship between departmental and senior management roles seems to me to be another area worthy of further examination.

## **9.6 Teachers' Narratives**

Because of the low level of dissemination of QUASE data within the four schools, there was little point in interviewing staff below middle management; the only non-management teacher interviewed was a member of the mathematics department whose review of the lower set curriculum was discussed in section 9.4 above. This was because I thought it important to test whether the head of department's account accorded with that of at least one of his staff members. The evidence from this level within school hierarchy is therefore sparse in the extreme. What is interesting from the research point of view was the impression that, although the member of staff agreed with the diagnosis and the decision arrived at, she spoke very much as someone who

was following a lead (or, indeed, her leader) – whereas the head of department had talked more in terms of collegiality and joint decisions. She also did not seem to exhibit the same ‘heuristic’ approach to the data as he; in fact, her explanation of the ways in which the department used data was more akin to the ‘technicist’ approach. This is an intriguing finding, which signals the need for more investigation in this area.

This teacher expressed the concern that pupils these days were not prepared to put in the work to get good grades, and that there was even a feeling amongst pupils and their parents that all the responsibility for achievement should be laid at the door of the teachers; she said that some parents were asking her: ‘are you on target?’ It is arguable that teachers are being put under pressure through the use of performance data not just by government and the media but by individual parents as well.

### **9.7 Governors’ Narratives: A Watching Brief?**

The discussion in this section is based on accounts given in interview by the chair of governors in School F (new in post), and the chair of governors and chair of the curriculum committee in School G. Thus again the evidence is quite sparse. The terminology used by these governors was a kind of common-sense language unlike that used by staff, for example: (re target-setting) ‘*[The headteacher] sets the standard and talks it through with the staff*’; (re value added measures) ‘*We try to find out what the difference is between being “up” and “down” from where we should be*’; (re data analyses) ‘*If we can recognise that we are adding value that’s useful*’. This may have something to do with the view that ‘*in our jobs, analysis of performance is much more cut-and-dried*’. There was some feeling that ‘*there is too much information in schools*’ and also that governors and parents likewise ‘*have an immediate suspicion of statistics*’.

Nonetheless, all three governors were aware of the responsibilities they had for overseeing performance. In one case it was clear that the headteacher was very much the controller of the information about achievements and targets her governors received: the interviewee said the headteacher set ambitious targets which have always been met, so ‘*I think we’d be above average [on the Autumn Package tables]*.’ In the other case, it seemed that the governors and senior management had found a way of working quite closely together,

although again the headteacher summarised the performance information for governors. Here, it was recognised that '*we can do a lot better than we are doing*'.

It may be asking a great deal of governors to understand and interpret value added performance data of the type provided by QUASE alongside all the other performance indicator information which is sent to schools by various agencies, and it seems appropriate that headteachers should try to provide some way through all the data. The QUASE service provides, on request, a governor information sheet which summarises the school's results on one side of paper – one of the schools had commissioned this, but the interviewee did not mention it. The governors of the other school did not seem aware of this option.

Perhaps the best governors can expect to do is to keep a watching brief on performance data and ask the simple but challenging questions.

## **9.8 Summary and Conclusions**

So far as the evidence from this study is concerned, the exploration was constrained by the nature of the data, which prevented an investigation of such issues as the construction of teachers' professional identities and the micro-politics of schools that might have been particularly salient. Even so, the interview data collected for this study can be interpreted as a variety of 'narrative', in which different attitudes, emotional and intellectual, towards the use of (value added) data could be discerned. The interpretation of the evidence in this way suggests that the meanings of (value added) data are socially constructed and emerge from the *interaction* between:

- the actual 'numbers on the page';
- the politicised significance which performance data as a species of information possessed at the time, particularly for headteachers and senior managers;
- what individual staff, particularly at head of department level, were bringing – often implicitly – to the discussion in terms of their own skills in and knowledge of performance data, their values and attitudes towards schooling and teaching, and their expectations and assessments of their pupils. This in turn poses the question: was value added data sometimes acting as a *locus* for conversations staff would like to have had, or anxieties they wanted to express, about their professional work and its value in the organisation?

The exploration in this chapter, in so far as it was able to pursue the discourse constructed by school staff around value added, seems highly suggestive of several important and interesting areas for further data collection and analysis.

**CHAPTER TEN**  
**VALUE ADDED AND SCHOOL IMPROVEMENT: THE**  
**DEVELOPMENTAL USE OF QUASE ANALYSES**

**10.1 Introduction: Using QUASE Analyses Developmentally**

The purpose of this chapter is to take examples of QUASE analyses and to discuss their use in school settings for developmental purposes. This is done partly in order to give a better sense of the actual data (beyond the sample analyses provided in the Appendices) with which staff have had to work, but more importantly to locate the discussion about value added in a more explicitly evaluative context. That is to say, the current chapter moves away from the largely exploratory approach adopted in Chapters Seven, Eight and Nine in order to argue in favour of approaches and activities which appear to have potential for bringing about school development.

Two rather different examples have been chosen to indicate the sorts of conditions under which QUASE data can be used with intent to bring about worthwhile outcomes. The first example is taken from a case study school already described in this thesis; the second example is taken from one of the schools with which I worked as part of my professional responsibilities. For reasons discussed previously (see Chapter Six), I had attempted to keep my professional role at NFER separate from the implementation of the fieldwork dimension of this research study. However, it happened that the NFER was commissioned by a local education authority to provide QUASE performance feedback to school staff in face-to-face meetings in seven secondary schools at a time when the fieldwork for the research study was in progress (see Chapter Six, section 6.8). The commissioned programme provided an alternative source of information about the uses of QUASE data and illustrated well the kind of political context in which the dissemination of such data is often situated. I have therefore chosen to make use of both kinds of evidence – from ‘pure’ research and from developmental work – in this chapter. Comparisons and contrasts between the research study and the commissioned developmental programme in terms of their contexts and outcomes were thus more or less inescapable: the implications of this for the overall conduct of the study are discussed in Chapter Eleven.

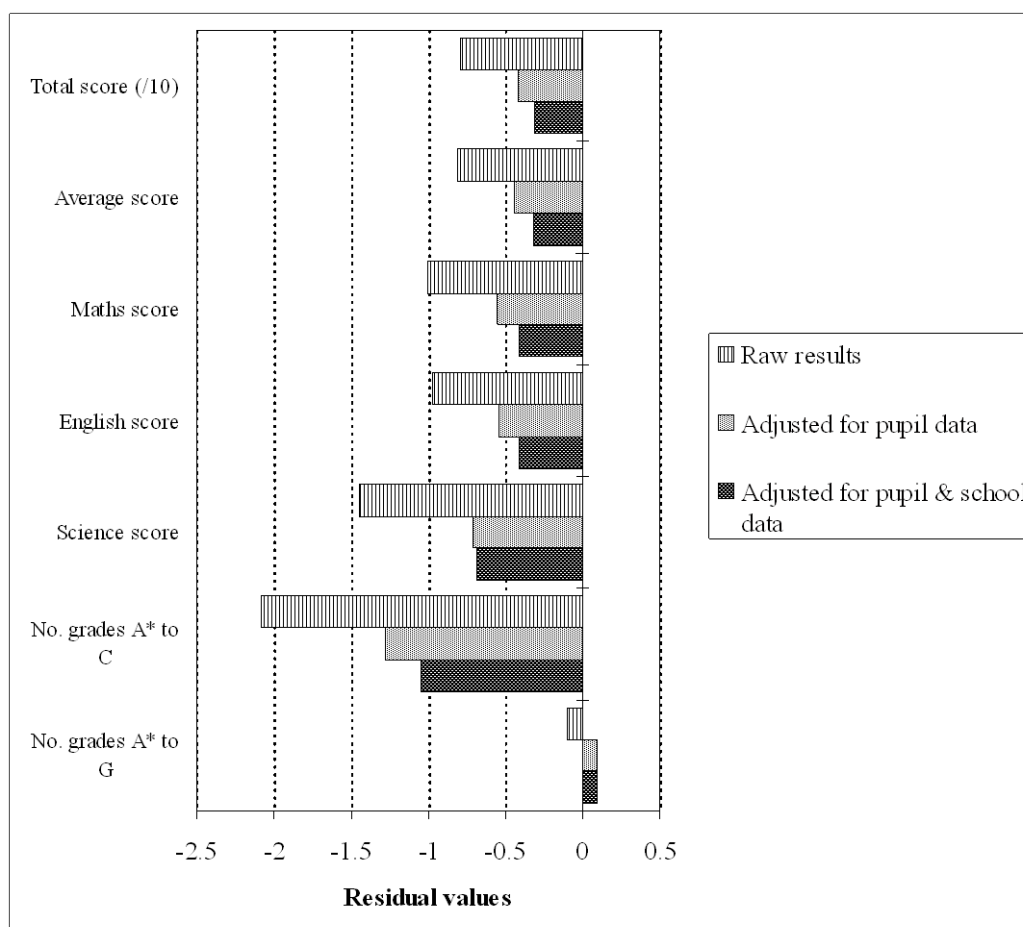
The current chapter ends by drawing some provisional conclusions about how the use of QUASE (and, by extension, other systems and services concerned with performance data analyses) can be further developed for the purposes of school development and improvement.

## **10.2 A Note on Presentation of Data**

QUASE, in common with many other ‘value added’ outputs, represents the results of analyses (using multilevel modelling) in the form of a residual or series of residuals, i.e. the difference between the average score and the individual school’s (or department’s) score. Such residual analyses are produced in the form of bar charts, as in **Figures 10.1** and **10.2** below. These kinds of diagram should always contain, or be supplemented by, information to show whether the residuals are statistically significant or not. In School Q’s case (**Figure 10.1**), for example – discussed in Section 10.4 below – the unadjusted residuals on total score, average score, mathematics score, English score and number of A\*–Cs were all significantly below average; the residuals adjusted for pupil data on English score and number of A\*–Cs remained significantly below what would have been expected; and the residual adjusted for pupil and school data on number of A\*–Cs was again significantly below what would have been expected. In this case, the diagram does not show this information and to that extent is misleading. The information is given on an accompanying table which provides upper and lower limits for each residual (not reproduced here, but see Appendices A and B). By contrast, in School X’s case (**Figure 10.2**) none of the subject areas performed, statistically speaking, better or worse than would have been expected on the basis of background variables. This is not the obvious conclusion to be drawn from the diagram, however.

Each of the diagrams in this chapter is reproduced exactly from anonymised originals.

**Figure 10.1 QUASE Performance Indicators (Residuals) for School Q**



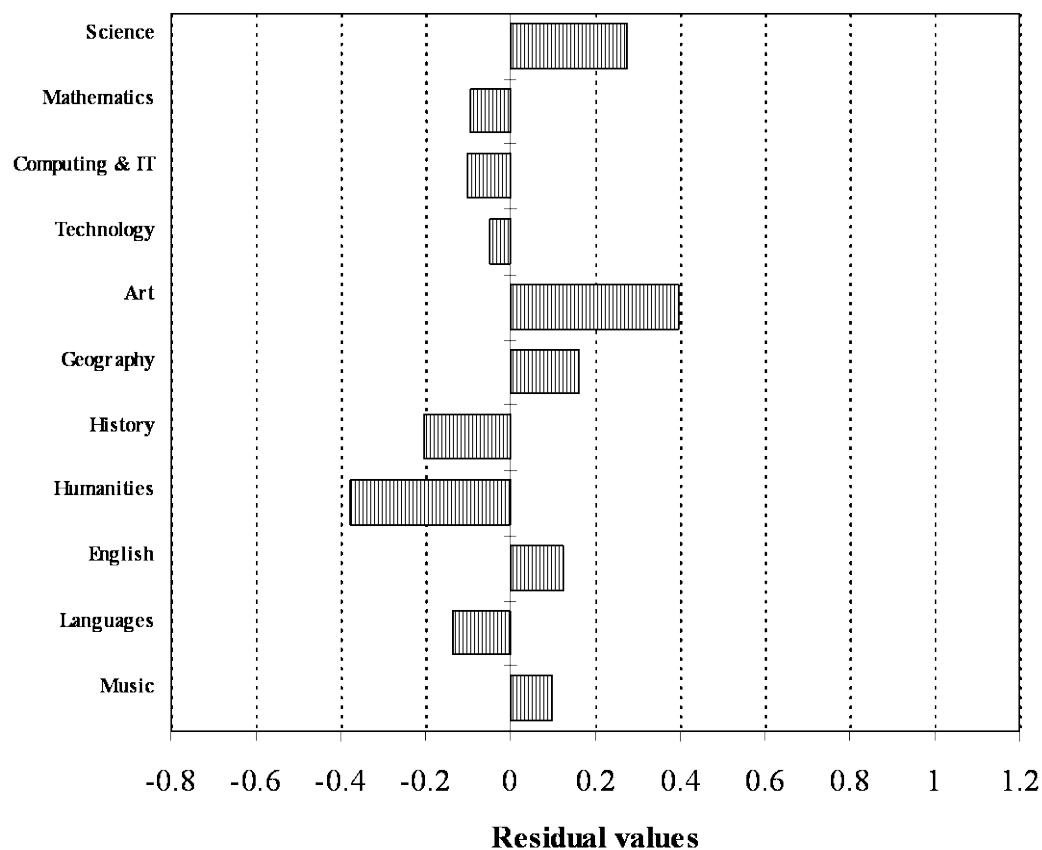
The adjusted values in the above table represent the differences between the actual performance indicators for the school and those 'expected', taking into account the background variables.

*Pupil variables used in the adjustment*    *School variables used in the adjustment*

Boys/girls	Size of Year 11
Prior attainment	Catchment area
Ethnic background	% eligible for free school meals
Age	

The background variables used are those which have been found empirically to correlate (positively or negatively) with performance, and which are consistently defined for all schools.

**Figure 10.2 QUASE Subject Area Results (Residuals) for School X**



The values plotted above represent the differences between the average score achieved in each subject area and the 'expected' value, taking pupils' total GCSE score and other pupil variables into account.

*Pupil variables used in the adjustment*

- Boys/girls
- Total GCSE score
- Ethnic background
- Age

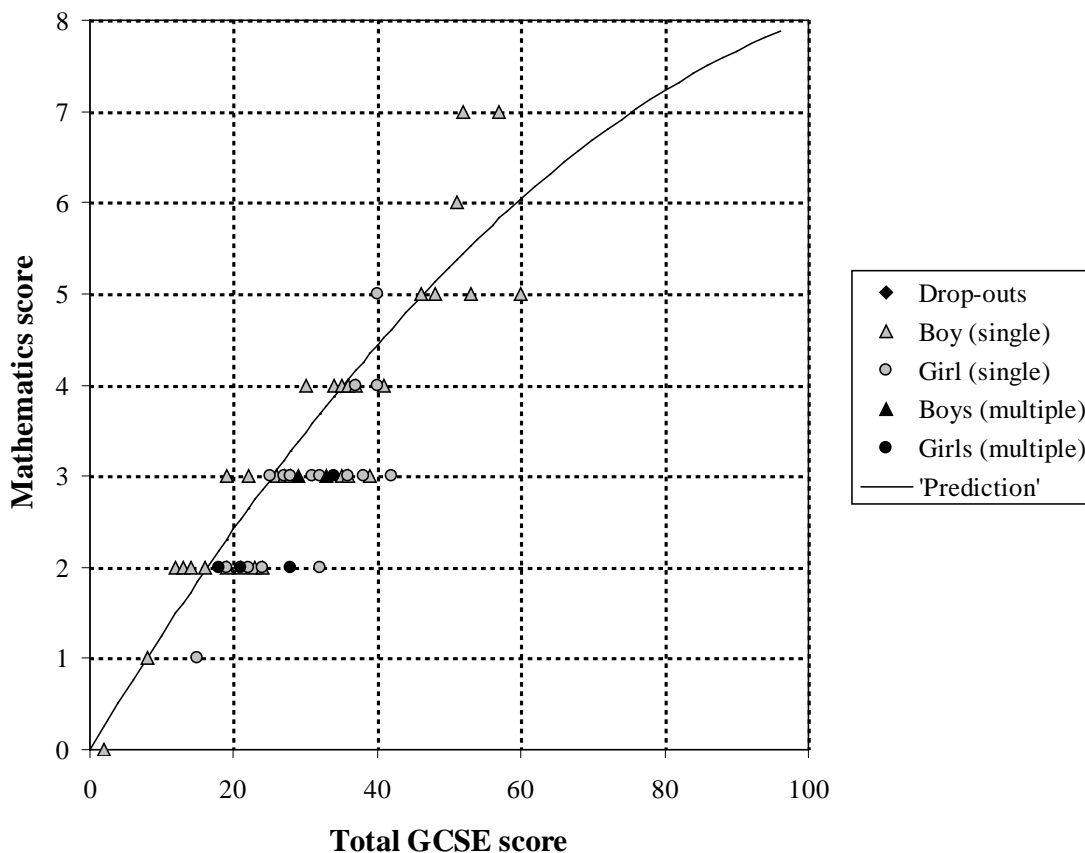
Subjects whose values are significantly above or below zero are indicated by an \* on the left hand side.

These diagrams provide a graphic representation of how well a school or department is performing relative to some expected norm and are intended to assist school and LEA personnel in making assessments of 'relative effectiveness'.

For subject managers and teachers, however, information which is more disaggregated may be more useful because it provides a clearer indication of how different pupil groups within the school/department have performed, for example, boys/girls, higher/lower attainers, whites/minority ethnic groups.

A way of representing pupil performance which shows how individuals have performed is in scatterplot diagrams. NFER's QUASE service (in common with products and services from several other organisations) produces scatterplots to represent the distribution of individual pupils' performance in relation to some notional 'predicted' or 'expected' performance, usually based on regression analysis: see **Figures 10.3, 10.4 and 10.5** below. Whilst these are not as statistically rigorous as value added scores based on multilevel modelling, they may be much more revealing as diagnostic instruments. One can see at a glance, for example, whether the results are closely bunched or widely scattered; how many and how anomalous the 'outliers' are; where, in relation to pupils' earlier performance and/or gender, the pattern of under-achievement, if any, is most prominent. Staff should be able to see any discernible patterns in the data as well as to note any individual outliers (to whom they may well be able to put names). Such information, as the examples which follow suggest, can be helpful in identifying under-achievement and suggesting strategies for improvement within a school or department.

**Figure 10.3 Pupils' results in mathematics in School F relative to total GCSE score**



This figure shows your pupils' performance in mathematics relative to their overall attainment at GCSE. The score for each pupil is calculated by allocating 8 points for A\*, 7 for A and so forth, calculating the average over all relevant subjects taken and summing over all subjects to obtain the total GCSE score.

The 'prediction' line is based on all the pupils in QUASE schools attempting this subject. The equation of the 'prediction' line is a best-fit curve to all the QUASE data, such that it predicts a score of zero when the total score is zero and never rises above a value of 8 (= A\*).

### 10.3 Example 1: Using the Data to Ask Questions about Teaching and Learning

The example discussed in this section is drawn from one of the research case study schools (School F) and has already been reported on in Chapters Seven and Nine above. The head of the mathematics department in School F held the post of data manager for the school and had assumed the role as ‘champion’ of value added analyses. He said in interview that he had convened a team meeting to see what could be derived from the departmental data from a number of sources over time. The departmental performance data, including that of QUASE – see **Figure 10.3** above – indicated that there was a consistent pattern of under-achievement in the lower sets, in particular the third set. The departmental team took time to consider the possible reasons for this. (The school’s OFSTED inspection had identified a quite different problem, that of not stretching the brightest pupils sufficiently!) Putting their views and impressions together with the data, the team concluded that the work was too dense and demanding for pupils in Set 3 and that in consequence these pupils had come to believe they were no good at maths and, not surprisingly, did not like the subject much. Having arrived at the school in Year 7 enjoying mathematics, pupils were reaching the end of Year 10 thoroughly demoralised.

The case study evidence showed that some departments were, with assistance, able and willing to carry out this kind of diagnosis. What is different in this particular case is the response of the head of department to the diagnosis. His response was strategic rather than tactical, focused on changing the curriculum rather than on a deficit model of pupils’ capacity: thus, instead of trying to improve pupils’ morale or giving extra support outside lessons, for example, he and a couple of colleagues spent a term re-writing the entire scheme of work to make the curriculum more lean and transparent for these pupils. As he acknowledged, this was a major undertaking whose investment in terms of staff time alone needed evidence of returns. At the time the research was conducted, there had not been time for a cohort to have taken GCSE based on this new curriculum, so no such evidence was available. The challenge, as the head of department recognised, would be for the department to make time properly to evaluate the changes made.

## **10.4 Example Two: Developing a Dialogue around QUASE Data**

As was said above, this second example was taken from my professional work at NFER, as part of a developmental project with an LEA. Section 10.4.1 gives the background to this particular example; section 10.4.2 describes what was done in developing the use of QUASE data and with what outcomes.

### **10.4.1 Working with an LEA to use QUASE data**

The LEA in question was one of the latest unitary authorities to be established and the performance of its schools had come under much closer scrutiny by the LEA team than when the schools were part of a large shire county. In recent OFSTED inspections, two of the schools, including School Q discussed below, had been designated as requiring special measures.

When the LEA decided to commission QUASE data analyses for all its secondary schools as part of a ‘value added and school improvement’ initiative, the depth of attention being paid to performance data was relatively new (only one of the schools had previously been involved in QUASE). The purpose of the commissioned work was accordingly to provide follow-up support for schools in interpreting and using the QUASE Profile (which also included pupil and parent attitude surveys). An initial seminar was hosted by the LEA, at which I explained the principles of QUASE to the schools’ senior management teams and LEA advisers, and gave an overview of what the data indicated about the schools’ performance. I then undertook a programme of school-based meetings planned around a detailed discussion of the data for each school and the departments within it, and the implications for pupils’ achievement. The meetings were explicitly set up to assist school staff, and the LEA personnel in attendance, to use QUASE to plan further activities and, if necessary, changes to practice. A member of the LEA, usually the school liaison adviser, attended the meeting with senior management in each school and made a record of action points; either the LEA adviser and/or a member of the senior management team attended, and wrote notes on, the meetings with heads of department.

I have chosen to discuss my visit to School Q in this chapter because it offers a clear illustration of the need to establish a receptive context and culture amongst staff for the discussion of value added data.

#### **10.4.2 Establishing a dialogue**

School Q – some of whose GCSE-related results are represented in **Figures 10.1** above and **10.4** and **10.5** below – is situated in the middle of a large estate of mixed social housing and owner occupancy on the edge of a medium-sized town (100,000 population). Although School Q is a so-called mixed ability school, the town also has four selective grammar schools which take up to 25 per cent of the higher academic ability children in the whole age cohort. School Q's most recent full OFSTED inspection placed it in special measures and the headteacher, appointed to improve the school's performance, had been in post for just under two years at the time of the QUASE feedback sessions. The school had been featured in rather hostile articles on the front page of both local newspapers and was facing a not unfamiliar downward spiral of difficulties of the kind that can culminate in closure: low examination results, buildings in need of renovation, socio-economic problems in pupils' families, negative publicity, falling roll, decreased funding. Expectations of pupil performance – on the part of some staff, parents and pupils themselves – were acknowledged, at least by the headteacher and the LEA adviser, to be a key issue in this school. (Indeed, this was acknowledged to be a key issue in all the non-selective secondary schools in the borough.) I visited the school twice to give feedback sessions on QUASE data, first to the headteacher and subsequently to all staff.

The background to these sessions was as follows:

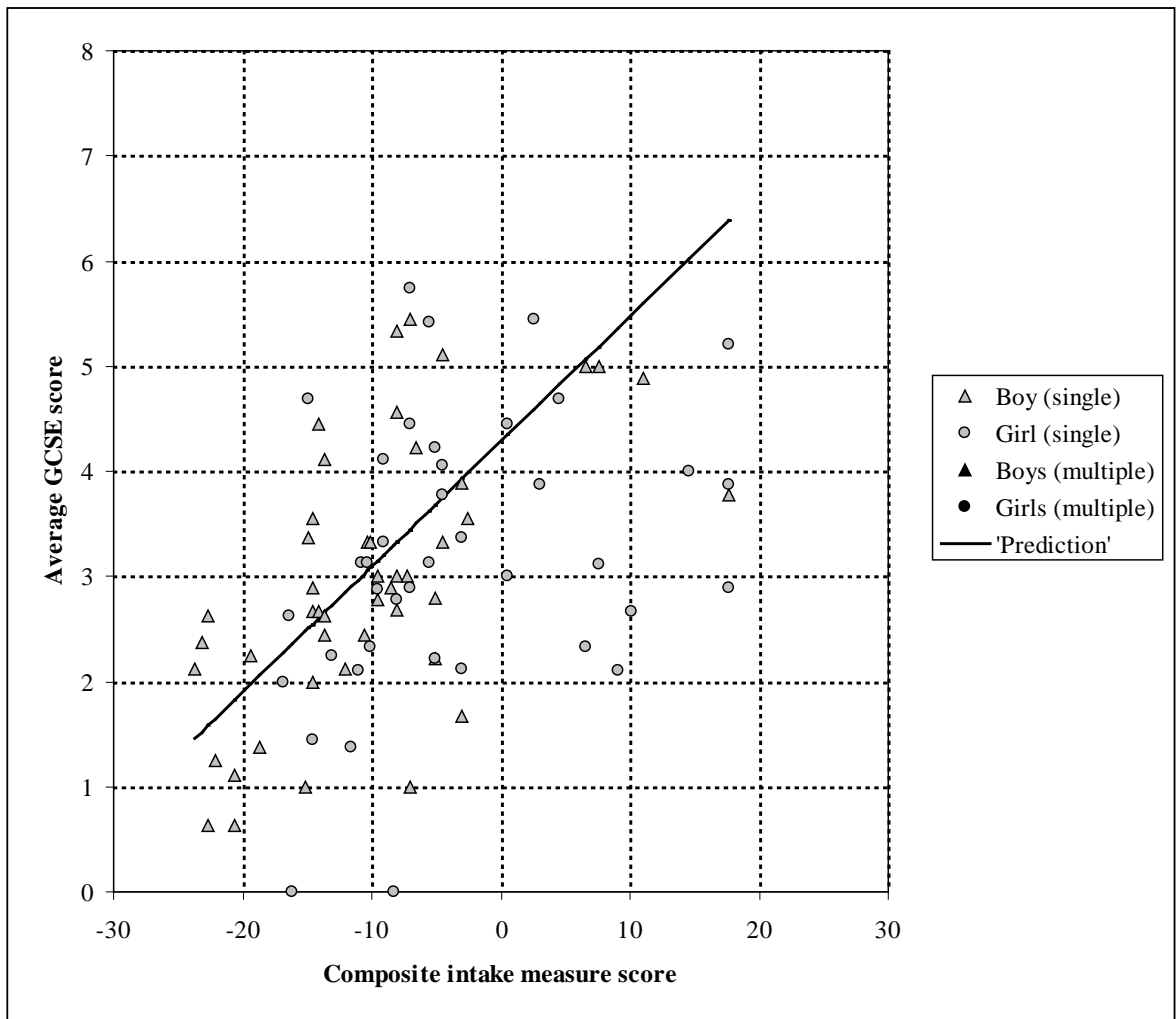
- As has been said, the LEA commissioned the analyses and feedback on behalf of all secondary schools in the borough for the 1998 GCSE results.
- Owing to some delays in the school sending the requisite data to NFER, however, the headteacher had not received the school's data and accompanying report on the 1998 GCSE results until the spring term of 1999; and it was not until the summer term that he asked to discuss the analyses with me. After that meeting, he said that he found the data a great deal more illuminating and useful than he expected from a cursory read, and wanted his staff to have a similar opportunity to discuss their own departmental performance. However, it took another term for this to be arranged. In effect, therefore, staff were being invited to discuss data which was a year out of date, the GCSE results for 1999 being the new focus of interest.

- The headteacher used the second of a pair of in-service training days on teaching and learning for the QUASE feedback sessions for staff; the previous day had been devoted to discussions of how to recognise effective teaching and effective learning in practice. The QUASE feedback sessions took the form of one-hour meetings for all members of each faculty group in turn in the school library (timetabled as designated faculty sessions throughout the whole day). Teaching assistants were included in the faculty groups.
- Although staff were personally courteous, it was clear from their tone of voice and/or body language that some of them were initially resistant to attending the QUASE feedback meetings: many would have preferred to spend the designated faculty time planning or marking, but were not given a choice.

Since this was the first time I had met the members of staff and therefore had no way of knowing what their (possibly quite diverse) expectations and levels of competence were, I adopted a strategy in which I:

- made it clear that these analyses belonged to the school and the LEA, and were not for publication nor for OFSTED inspection purposes;
- explained what the school's overall results showed (see Figures 10.1 and 10.4 above) before talking their own departmental data through (see Figure 10.5 below for one example), so that each department had the whole-school context;
- pointed out what was positive in the results before going on to make suggestions about what I thought I could see in the scatterplot distributions;
- kept checking out these observations with them, in the form of questions like, 'is that what you see as well or can you make out a different pattern?'; 'what do you think might lie behind this that the statistics do not show?'
- continuously tried to pick up verbal and non-verbal cues from people who might not understand the data and to give simple clear explanations; it seemed particularly important to include the teaching assistants in this process, since they tended not to talk much without being invited.

**Figure 10.4 Pupils' average GCSE scores results in School Q relative to intake measures**



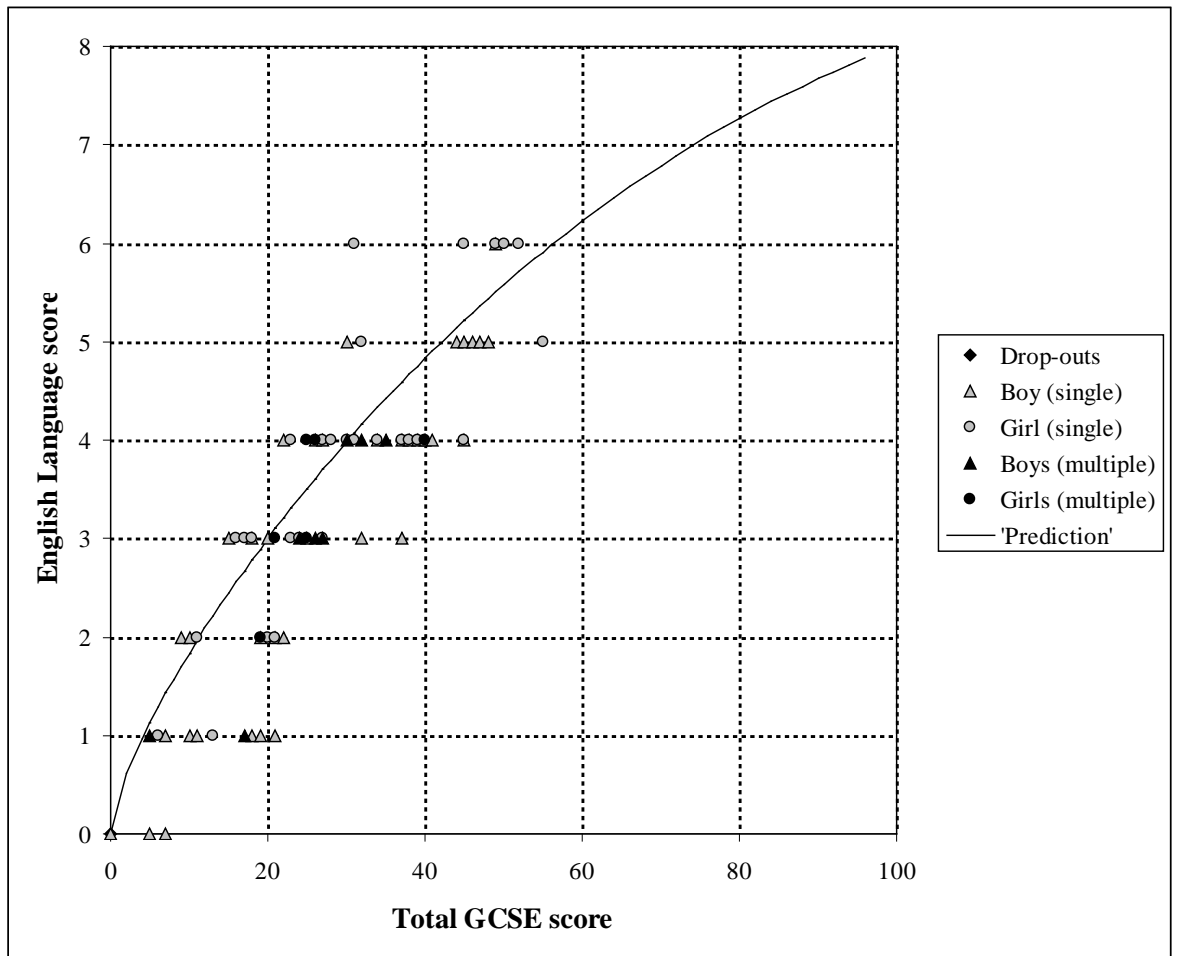
This figure shows your pupils' overall performance at GCSE relative to their prior attainment. The composite intake measure score is composed of a weighted sum of all available and suitable prior attainment measures for each pupil, such that a value of zero implies 'average' prior attainment.

The average GCSE score for each pupil is calculated by allocating 8 points for A\*, 7 for A and so forth, for each subject and averaging over all subjects taken.

The 'prediction' line shown is based on all the pupils in QUASE with prior attainment data and has the equation:

$$\text{Average GCSE score} = 4.285 + 0.119 \times \text{Prior attainment}$$

**Figure 10.5 Pupils' results in English in School Q relative to total GCSE score**



This figure shows your pupils' performance in English relative to their overall attainment at GCSE. The score for each pupil is calculated by allocating 8 points for A\*, 7 for A and so forth, calculating the average over all relevant subjects taken and summing over all subjects to obtain the total GCSE score.

The 'prediction' line is based on all the pupils in QUASE schools attempting this subject. The equation of the 'prediction' line is a best-fit curve to all the QUASE data, such that it predicts a score of zero when the total score is zero and never rises above a value of 8 (= A\*).

The outcomes from these faculty discussions can be summarised as follows:

- With some prompting from me on how to read the graphs, the faculty groups acknowledged that the school's lower-than-average performance could not be entirely explained by the school's context. This in itself was a breakthrough, since a common response to criticisms of the school's performance had up until then been to plead the school's catchment and status. But there were clear indications in the QUASE data of under-achievement by pupils who had come into the school with above average attainment scores – the 'most able' pupils in the school's terms. This was thought to have arisen because the school's policy had been to concentrate on the needier pupils, of whom there was a sizeable proportion. The more able pupils might well be individuals who had only just failed to reach the threshold for a 'pass' in the 12-plus tests, and who consequently needed not just support but challenge as well. The striking thing was that the majority of these pupils were girls. The faculty groups reached a consensus that it was a whole-school responsibility to identify these pupils during their first year at the school and to put better support/extension provision in place. These conclusions arose from the discussion of the whole-school data, particularly as presented in Figures 10.1 and 10.3 above, but also including the pupil attitude survey data.
- A theme to emerge in the discussions, which the numerical data could not show, was that the current GCSE options system and the way pupils were being advised about subjects needed reviewing. It was felt by several members of staff that one of the reasons for under-performance by some pupils at GCSE could be the relative rigidity of the options system which resulted in pupils having to take subjects in which they had expressed little or no interest.
- Some groups of staff said at the time that they felt that the feedback discussions had helped them to understand and engage with the data – some admitted that they had found the discussions unexpectedly illuminating. Other similar reactions were reported back to me by the headteacher and LEA liaison adviser. Moreover, staff were able to appreciate how, in the context of the two days' in-service training, this kind of quantitative information could underpin their more qualitative discussions about how to make teaching and learning more effective. This kind of comment was particularly gratifying when it came from members of staff who started off the meeting by expressing scepticism and/or lack of confidence in dealing with statistics. In such cases, it seems highly likely that this change of mind would have been difficult without the input from a facilitator (and perhaps someone who was not connected with the school and was perceived as having no particular 'axe to grind').
- The feedback from the LEA was also positive: there were concerns that the school was finding it hard to move on from a position of systemic under-achievement and that identifying the right combination of levers for change was particularly difficult. The QUASE feedback sessions were seen as a successful way of establishing dialogue, based on evidence, about expectations and achievement in the school.

## 10.5 Good Practice in Using Value Added for School Improvement

I argued in Chapter Five that *‘given the right culture, value added can help to pose better and more focused questions about the way a school or LEA has performed with its pupils and to stimulate more informed discussion amongst school staff’*. The two examples of QUASE data ‘in action’ presented above – supplemented by indications emerging in Chapters Seven, Eight and Nine – give some useful indications, I think, of what that culture looks like. I also believe the evidence allows us to make some provisional suggestions about good practice in using value added data for school development, i.e. for stimulating better and more informed dialogue about performance, achievement and expectations.

First of all, there are some points to be made about the presentation of the data:

- Staff appeared to appreciate the fact that the analyses showed what was **specific to the department/faculty** whilst putting this in the context of the **whole school’s performance**.
- Presentation of data which showed patterns of **individual pupil performance** seemed to engage teachers more readily than aggregate data.

Secondly, there are several things to be said about the context and processes within the schools:

- **The part played by the headteacher seems key:**
  - The two schools discussed in this chapter were dissimilar in terms of their stage of development – one was fairly secure while the other had recently come out of special measures – but in each school<sup>22</sup>, the headteacher understood the role which could be played by value added data in helping to **raise questions about teaching and learning** and levels of attainment. This was also true in the other schools in the case study where value added data was being used for school development.
  - Both headteachers had themselves previously **received input** from the QUASE team; there was evidence that this had been a positive factor in the other two case study schools where it had been commissioned (see Chapter Eight).
  - Consideration of the analyses by heads of department was **not optional**: the headteachers in both schools required staff to take time to discuss the data and to make a response to it.

---

<sup>22</sup> For the role played by the headteacher in School F, see Chapter Seven.

- So far as it was possible to tell, senior managers believed in, and were trying to create, a **culture of dissemination and sharing**. Although this was not necessarily strongly established throughout each school, and was inevitably vulnerable to pressures of time and (in one case) defensiveness brought about by adverse publicity about the school, it was nonetheless possible to see the signs of such a culture in practice. In particular, setting up **group discussions/team meetings** seemed to play a useful role in (i) establishing a common view of what the data showed and (ii) elaborating the significance of the data for teaching and learning.
- The potential importance of a **champion (internal) or a facilitator (external)** was already highlighted in Chapter Seven. The examples of School F and School Q in this chapter show in greater detail how the input from such a person can help the accessibility, interpretation and acceptability of the data, particularly in terms of relating past performance to present and future actions. Such input also helped to ensure that the data was presented and used as an **aid to reflection and development**. The other major function of the data – its use as a management instrument for monitoring performance – was kept distinct and, as far as possible, separate. When the data was used to stimulate discussion, staff contributed **other kinds and sources of information** to broaden, deepen or modify the conclusions that could be drawn from the ‘numbers on the page’.
- Such discussions led in each case to **specific actions to be taken** by the school or department being identified and agreed; this helped to impart a sense of purpose to the discussions, and clearly gave staff a feeling that they could find ways to move forward – sometimes from quite a negative starting-point.

## 10.6 Conclusions: Improving the Use of Value Added Data

If these factors are transferable (and my professional experience outside the research study suggests they are), then it is arguable that QUASE data analyses – and by extension, other kinds of performance data analyses – could be managed both more firmly and more sensitively than is often the case, with the result that their understanding and use could bring about more positive change and development. The final section in this chapter suggests, from an avowedly evaluative stance, what the main lessons are for making better use of value added data for school development.

**Providers of data – including government agencies – need to understand how it feels to be ‘on the receiving end’.** As has been said, School Q, in difficulties for some years and recently identified as needing ‘special measures’, had been pilloried in the local press. Staff were therefore already primed to be defensive, demoralised, angry and/or dismissive about further scrutiny of their pupils’ performance. Indeed, the QUASE results might have

seemed particularly unwelcome since they did not offer much by way of consolation or justification. But, although the data apparently held some hard messages, it would not have been at all helpful to start from an assumption that, simply for that reason, the staff were lacking in professionalism, let alone wholly or equally to blame for pupils' under-performance. On the contrary, it was evident, from talking with members of staff, that many of them felt protective towards their pupils and wanted to know how to help them overcome what they saw as multiple barriers to these young people making educational progress. The main need in such a context is to frame the data analysis in such a way that it can be used as a diagnostic tool rather than as a reinforcement of negative feedback which is already all too familiar. The example of School Q suggests that staff are more likely to be open to thinking about making changes if there is information that, rather than merely making judgements, helps them to diagnose what has been going on.

**Policy-makers need to understand the 'concrete messiness' of schools.** As Gray (2000) says (p. ii): *'school improvement is simply too complex a process to mandate from the centre'*. Perhaps the process of school development and improvement is more realistically thought of as a series of somewhat unpredictable social interactions – to understand which the discipline of social anthropology would be highly appropriate – than as a set of statutory procedures. Without an appreciation of how professional, institutional and personal pre-occupations may drive or inhibit individuals' or groups' actions, it is impossible to provide credible guidance on how schools can become 'learning organisations'. In the case of the mathematics department in School F cited above, my visit to the school on an ordinary working day made it quite easy for me to understand how – despite their intentions – the normal daily tasks of teaching and trouble-shooting had up to now stood in the way of the staff undertaking an evaluation of the changes they had made in response to the QUASE data. So, as well as providing time and space for staff to reflect on the available evidence and its implications, managers need to decide what further strategic resource commitments will be needed to assist the school to become a 'learning organisation' (rather than merely compliant with imposed initiatives).

**School senior managers and LEA advisers need to consider how to foster and make best use of an internal ‘champion’ or an external facilitator.**

The examples given above indicate that there is a role for expert input and support. As I hope I have demonstrated in this thesis, value added data is by its nature complex and makes demands on people’s understanding. Given the pressures on staff time and attention, it may be insufficient to rely on explanatory text, let alone merely sets of tables and graphs (both of which were, of course, supplied by QUASE). Face-to-face facilitation by someone well-versed in the data can first of all provide the requisite space in which attention can be single-mindedly focused on the analyses. Facilitation sessions can also take account of, and adapt to, the particular school circumstances on the day, any apparent anomalies in the data and the wide variety of individual staff concerns which may arise. Moreover, a facilitator – particularly an external one – can challenge a discussion which seems to be veering too far in the direction of self-congratulation or self-justification (see Dudley, 1999b). Finally, as Gray *et al.* (1999) say, there are likely to be pockets of ‘good practice’ even in the least successful of schools – a facilitator can help both to identify where these are and to convey to other staff the perhaps uncomfortable notion there may be something worth learning from their colleagues.

**School senior managers and LEA advisers need to think about how to establish ‘ownership’ by teachers of the data and its implications.**

The evidence suggests that if a climate of ‘ownership’ of the data can be established, through encouraging frank dialogue and asking open-ended questions, staff will respond positively and be able to find in the data *‘a place to put our feet’* (in the words of one headteacher) – this can happen even where it is known that the data is also being used as a monitoring instrument by senior management and/or the LEA. The sense of ownership needs to be extended to a discussion of what actions might need to be taken as a result of understanding what the data seems to show. Sometimes this may be a question of deciding to focus on how particular subjects are taught to particular pupil groupings (for example, mathematics for the third set); sometimes it may be a question of whole-school action (for example, reviewing how the subject options scheme works). The important thing in both cases was that the decision came from the staff rather than being imposed on them. A more general statement of this idea can be found in Harris and Hopkins (2000, p. 13):

In many school systems there is despair at the signal failure of both 'top-down' and 'bottom-up' reforms to significantly enhance the learning of students... All the more recent studies of centralised policy initiatives confirm the established finding [that] 'local implementation dominates outcomes'... Yet... a meta-analysis of school improvement strategies clearly supports the conclusion that most schools without some form of external support have no idea at all at how best to direct resources towards enhancing student achievement. So if neither 'top-down' nor 'bottom-up' works, what combination of 'pressure and support' is required to support school improvement?... Simply it is ... about... building capacity from within. Effective schools throughout the world have created internal contexts within which powerful learning and teaching occurs – they are schools that have norms of continuous improvement.'

**School senior managers, and LEA advisers and other professionals playing a support role need to acknowledge the importance of school culture.** The evidence also suggests that – at least in the short term – data does not change school culture but, rather, the school's culture shapes how data will be received and presented. The kind of culture a school has is often shaped by a few key individuals: Gray (2000) puts the issue like this:

Personalities are likely to loom large; learning to transcend them is important... (p. iv).

He goes on to suggest:

By understanding more about the processes they have been through and the small steps they have succeeded in making (rather than the larger ones which have defeated them), a greater impetus for change can be created... (p. iv).

Another way of countering difficulties may be to go quite slow at first: Gray also argues that:

persuading [teachers] too early to 'lock into' issues to do with their own performance may be too challenging and ultimately dysfunctional. They may well need time to work their way round to such issues. Furthermore, early attempts by schools to learn may create uncertainty and even (at least initially) an element of dissatisfaction; as a matter of course schools should anticipate some instability (p. iv).

The example of School Q above indicates that the headteacher, who was relatively new in post, may not have been wrong to delay the dissemination of QUASE analyses by almost a year. Understanding how to diagnose, ‘tap into’ and work on a school’s culture may be the single most important capacity for professionals wishing to bring about change to acquire.

**Both schools and central government need to recognise the supporting role of the LEA.** Certainly, in the example of School Q above, the LEA initiated the QUASE analysis as part of the borough school improvement initiative, commissioned the NFER to provide face-to-face feedback to the headteacher and then to faculty groups, and provided INSET to help to integrate the value added analyses with target-setting and self-evaluation activities. The school’s liaison adviser followed up with the school staff the activities planned as a result of the QUASE feedback. The LEA was acting not only as a source of funding for this work but also as a ‘critical friend’ or witness to the school’s development and a monitor of change.

**Providers of value added data might usefully review the issue of ‘timeliness’ of data.** It is assumed by most agencies that data needs to be analysed and distributed as soon as possible after the test/examination results have been published. Certainly, several schools in the fieldwork study had asked if it were possible for QUASE data to arrive in the first half of the autumn term. Staff expected to hold discussions about GCSE performance quite early in the autumn term, so as to put in place any arrangements deemed necessary for current Year 11 pupils as soon as possible. However, in the case of School Q above, the QUASE data was already a year out of date by the time the headteacher decided to disseminate it. This may paradoxically have made the analysis more acceptable: the results could be treated as historical while the principles were accepted and internalised. Moreover, the fact that the results were so obviously ‘out-of-date’ seemed to make it easier for staff to think about the data more strategically, in terms of what needed to be done in Years 7 and 9, as well as in Years 10 and 11.

All the above lessons are fairly practical, being suggested by a combination of the case study fieldwork and my professional developmental work. The following, final, chapter of the thesis discusses what the study as a whole – through the literature review and the conceptualising of the issues, as well as through the two kinds of empirical fieldwork – has added to the understanding of the national value added agenda and its potential role in school improvement.

## ***CHAPTER ELEVEN***

### **SUMMARY AND CONCLUSIONS: SOME WAYS FORWARD FOR VALUE ADDED AND SCHOOL IMPROVEMENT**

#### **11.1 Introduction: Concluding the Study**

This chapter concludes the thesis by presenting a summary of the key contributions the study has made to the field; these comprise, first, a comprehensive historical contextualisation and conceptual clarification of the principle of educational ‘value added’ and, secondly, a critical analysis of empirical evidence about value added in use. But because that analysis in turn revealed how much more there might be to understand about ‘the psychology and sociology of numbers’ than was encompassed by the present study, the chapter goes on to suggest other relevant perspectives and approaches which might be useful in pursuing the issues beyond this particular thesis. In doing so, the chapter addresses the tensions between research and development, and between exploration and evaluation, which the aims and conduct of the study exemplified.

Notwithstanding these strictures, the chapter ends by arguing that the study has some new and valuable things to say about the use of value added measures for school improvement, from the perspectives of both policy and practice.

#### **11.2 Summary of Key Findings from the Study**

The critical review of the literature on value added in education, undertaken in Chapters Two, Three, Four and Five above, attempted to impose some conceptual clarification on what is meant, in educational terms, by ‘value added’. These chapters charted the historical development in the uses of the term up to the time when a national value added system was established in England and Wales. The literature review raised a number of issues about the basis for understanding and measuring value added, and the implications this has for policy and practice; thus these chapters themselves constitute new knowledge as well as an overview of the field.

Chapter Two showed that during the 1980s and 1990s the combination of the development of sophisticated analytical models by school effectiveness researchers and the focus on value-for-money in further/higher education changed the attitude of educational managers towards performance data. But it also showed that different people can mean rather different things by the term ‘value added’; indeed, that there are ‘functional ambiguities’ in the term – that is to say, ambiguities which are integral to how the term is made to function. Notwithstanding this conceptual complexity, the chapter summarised the key points on which the value added agenda is premised, as follows:

- educational performance makes a difference not only to the individual’s life/career opportunities but also to the health of the national economy;
- robust measures of standards and quality, and of the extent of the contribution of education to the economy, are useful and desirable; it is necessary to go beyond the assertions of educationists about the efficacy of educational provision;
- value added has come to be a meaningful term in an educational context and can be deployed in the assessment of standards and quality;
- individual institutions make a difference to students’ educational performance; it is possible, and from many points of view desirable, to make comparative assessments of institutional performance; some institutions can be shown to be measurably more effective than others in terms of specified student outcomes;
- it is possible to identify and account for key factors beyond an institution’s immediate control which are implicated in its students’ performance;
- it is possible to calculate ‘predicted’ or expected levels of performance based on these factors, and then to assess individual institutions’ results against these.

Chapter Three reviewed a selection of material – upwards of fifty articles, papers and/or reports – in the academic and research literature, having first established cut-off dates and other criteria. This review revealed how the activities and issues encompassed by the term value added need to be ‘read back’ into previous research pre-occupations. The chapter also demonstrated how the focus of value added usage has shifted from being a contested methodology concerned with making fairer and more valid comparisons between schools to being a management tool concerned with raising pupils’ attainment. Finally, it argued that there were paradoxes associated with using value added measures for accountability purposes. In effect, value added measurements serve to reinforce the provisional and probabilistic nature of all

quantitative performance data: as Thomas and Goldstein (1995) remark, *'research emphatically demonstrates that the measurement of progress or value added... is neither simple nor straightforward'*.

Chapter Four examined the policy-related literature in an attempt to elucidate how the government's education policy had moved from repudiating to supporting value added; and what particular meanings and model of 'value added' it was espousing. It was noted that the developers of the national value added system (Fitz-Gibbon and Tymms, chiefly) had made animadversions in their final report (Fitz-Gibbon, 1997) about the misuses to which value added data might be put. Nonetheless, for policy-makers and educationists at large, value added does now seem to be, in Fitz-Gibbon's (1996) words *'the statistic of choice'* for assessing performance and progress. One unintended outcome may well be that value added has done nothing to undermine what one might call 'the tyranny of the measurable'.

Together, these chapters inform, situate and justify the fieldwork component of the thesis: they show how far there was a need to elucidate value added both as methodology and as ideology before any empirical work could be credibly undertaken into how schools might make use of value added data. Chapter Five shifted the focus to these more empirical considerations and examined the admittedly fairly meagre evidence from the literature purporting to describe how schools react to this kind of performance information (Harris *et al.*, 1997; West and Moore, 1998; Wikeley, 1995, 1998; Williamson and Fitz-Gibbon, 1990; Williamson *et al.*, 1992; Yang *et al.*, 1999). The issues – and the gaps in knowledge and understanding – which were thereby highlighted gave grounds for believing that a useful study could be undertaken in order to:

- give a good descriptive account of how value added data was used in different kinds of schools;
- highlight the conceptual, ethical and managerial challenges that ensued from the introduction of such data;
- make some recommendations for using value added data well, based on the notion that *'given the right culture, value added can help to pose better and more focused questions about the way a school or LEA has performed with its pupils and to stimulate more informed discussion amongst school staff'*.

The setting up of the study, in order to include schools of different types and at different stages of improvement, was described in Chapter Six. Ten schools were originally chosen for case study work, and nine participated. Chapters Seven, Eight and Nine analysed the findings of the fieldwork by providing descriptive, explanatory and speculative accounts respectively of what was seen and heard. It is arguable that this iterative approach itself constitutes a methodological outcome from the study, created in an attempt to problematise and re-conceptualise what challenged the author's initial understanding.

The findings from those three chapters showed, amongst other things, that:

- Compared with what had been anticipated, the use of value added data for school development was relatively limited in the schools visited at the time of the study. Nonetheless, there were signs in some schools that the use of data was 'taking off' and beginning to have a positive impact on monitoring and self-evaluation.
- There were noticeable differences between different departments in the same school in the level and type of use of value added data.

Five factors were identified which seemed, in combination, to be related to schools' utilisation of value added data for school improvement, as follows:

- active support for value added by senior management;
- 'championing' or mediation of value added analyses by a senior member of staff;
- length of time of the school's participation in – and presumably therefore familiarity with – the QUASE service;
- previous exposure of at least some staff in the school to guidance on the interpretation of QUASE analyses;
- the degree of incentive or impetus for action provided by the school's 'relative effectiveness'.

Moreover, when the fieldwork data was analysed in terms of what could reasonably be ascertained about staff attitudes towards value added data, it seemed certain that teachers' own beliefs and values about teaching played a crucial role in how they received, interpreted and put the data to use (or not). This evidence of how data is perceived by school staff is certainly compatible with Wikeley's (1998) and Dudley's (1999a) studies, which each found that staff attitudes were highly implicated in how data was received and conceptualised, and that a range of different attitudes was detectable. Consequently, it tends to throw doubt on the more simplistic account given in

West and Moore (1998), which did not really problematise these issues. Furthermore, if the findings from the present study were to hold good for different circumstances, times and locations, they might help to elucidate further the findings of Williamson and Fitz-Gibbon (1990) about the restricted use of ALIS data.

Chapter Ten rounded off the empirical dimension by looking in detail at two particular examples of schools where value added data was being used with a view to bringing about positive changes. From these examples, some ideas were developed for improving the use of value added data for school development. The chapter drew heavily on experiences and expertise which the author had acquired in her professional capacity – and this raises some important issues for the conduct of the research which are examined in Section 11.4.4 below.

### **11.3 Disquisition on the Findings**

Some of these findings and outcomes deserve further elaboration in order to bring out their implications. The main issues under review in this section are given in bold type below, followed by a brief explanation/discussion.

**‘Value added’ measures of performance are now a major feature in the educational landscape.** The context for value added has changed dramatically since the 1980s, when ‘value added’ was thought of, if at all, by many educationists as a quasi-technical idea which had strayed into education from economics; and was regarded by influential politicians with suspicion. One of the major forces for change were the published performance tables on individual schools, which were perceived as unjust and inequitable in concept and consequence. School managers and teachers joined with scholars to intensify the debate about how to measure the performance of pupils in the nation’s schools in a way which could shed light on progress made as well as standards reached. The change of political heart which instigated the national value added system that became operational in autumn 1998 can be seen as the consequence of a decade of sustained argument by this coalition of academics and practitioners. What started out as an academic debate on how to analyse ‘school effects and effectiveness’ through statistical modelling has become taken-for-granted knowledge (by some), now integrally linked with the national political agenda for educational quality.

**The idea of ‘value added’ contains within it definitional and conceptual problems.** Of the various implicit or explicit definitions of value added encountered, this study favours the definition given by McPherson (1992): *‘[value added is the] calculation of the [relative] contribution schools make to pupils’ progress’*. Making such calculations entails finding a way to analyse performance which takes account of factors which have been empirically found to be associated with performance but over which schools have little or no control. This notion in itself has been the subject of much debate, concerned chiefly with whether it is desirable to ‘make allowances for’ some groups of pupils, as the agenda has been described – although in fact the statistical procedure of ‘allowing for’ or ‘adjusting for’ can be thoroughly justified when the purpose is to create fairer methods of institutional comparison. The underlying aim of value added continues to be strongly contested: the school effectiveness agenda – to which the value added ‘project’ must be seen as belonging – has been criticised by some scholars (e.g. Elliott, 1996; Slee *et al.*, 1998) for being educationally and socially regressive and/or insufficiently grounded in (good) theory.

**Nonetheless, the continued attractiveness of the idea of ‘value added’ has meant that methods for calculating it have been robustly trialled, developed and refined.** ‘Value added’ was eminently an idea whose time had come in response to the crude ‘league tables’ of school performance. The attempts – which in retrospect seem little short of heroic – to generate a statistically accurate yet feasible system for producing information which can be used, on the one hand, to assess the ‘relative effectiveness’ or ‘value added’ of different schools and, on the other, to provide diagnostic assistance for school managers and staff have culminated in some consensus around the right kind of model to use. Most experts in the field (following Gray *et al.*, 1986) are now agreed that it is necessary to take account of the following factors:

- outcome measures which reflect *all* levels of pupil performance;
- a measure of each pupil’s prior attainment;
- information about each pupil’s background (gender, ethnic group, socio-economic status).

Additionally, the calculation should:

- be based on data collected at the individual pupil (not aggregate) level;
- use multilevel modelling as the statistical technique.

**Even so, there remain a series of technical problems to which the research is now turning its attention.** Chapter Three identified further questions which value added approaches have begun to explore, including:

- What can be said about social as distinct from academic effectiveness (as measured by attitudinal data and/or attendance)? Are the independent variables correlated with social effectiveness the same as for academic effectiveness? What evidence is there of stability/consistency/improvement in terms of attitudinal factors in individual schools?
- How far is academic effectiveness correlated with social/motivational effectiveness within the same institution?
- How might the correlates of *changes* in school effectiveness (e.g. towards improvement) be identified?
- Given that research suggests that the greatest variance is within-school rather than between-school, how might the relative effectiveness of teaching groups/classes, etc. be explored?
- Is it possible to investigate *'how the variation [in effectiveness by... ability group, by social class, by ethnicity and by subject department] is maximised or minimised within different types of "effective" or "ineffective" schools?'* (Reynolds, 1995.)
- Institutional 'effectiveness' might, on the other hand, be thought to be correlated with the level and type of LEA or other external support. How could this be explored using multilevel approaches?
- Going on from here, what other important factors – notably i. the effect of previous institutional impact on pupils' performance; ii. the relational or micropolitical aspects of schools; and iii. 'partnership' models of education which utilise agencies to work in or alongside schools – should now be brought into the analysis of school effectiveness? How could this be done validly and manageably?

Such questions constitute an ongoing and challenging agenda for the research community.

**Because of the technical complexity involved, the focus of much work in the value added field has been on methodological accuracy.** Given the complex nature of the data and the analytical models, work has concentrated on the 'provider' end of the process, from conceptualising the statistical model appropriately to collecting the right data in the right form (see Saunders and Thomas, 1998, for a summary of such issues). What use, if any, was made by headteachers and staff of the analyses remained an relatively unexplored area, although one about which assumptions were frequently made. Such is the official faith in the benefits of performance, especially value added, data that

most of the current guidance from the DfEE to schools on using data for target-setting, for example, continues to be based on prescription rather than in-depth description.

Yet the examination undertaken as part of the present study, in Chapter Five above, of the various kinds of empirical evidence concerning schools' use of value added data showed that a comprehensive and conclusive account of how schools – or rather staff at various levels within schools – use value added data had still to be constructed. That evidence – which now includes the new evidence generated by the author's own fieldwork – suggests some complicating factors may be at work.

**Value added data carries a particular symbolism for many school staff at present.** Performance data generally may be impartial in intention, but is rarely perceived as neutral in effect. Value added data in particular is complex, relatively new and often 'high stakes'; it can therefore represent an overt threat or at least a felt disturbance to some kinds of school, departmental and/or group culture; in some cases, the data's significance can be transformed by the institutional context. As many interviewees in the study made clear, value added data must be seen as part of a policy agenda, about scrutinising and making accountable school, departmental and teacher performance. It is probably not possible for any performance data to be perceived and used in a way which wholly ignores – or remains unshaped by – this context. And from a teacher's point of view the policy context is felt to be a rapidly moving one. When the fieldwork for the study started in summer 1998, the main new item on schools' agenda was target-setting. Since then, schools have needed to react to the 'Autumn Package', PANDAs and the Green Paper on teachers' pay. There is also now (in 2000) a strong drive towards developing centralised instruments for school self-evaluation (for example, by OFSTED and the Audit Commission). Scrutiny of results for accountability and diagnostic purposes, at LEA, institutional and sub-institutional levels, is therefore likely to be strengthened at least in the short term.

**In any case, data does not speak for itself.** Data depends on a series of prior decisions about *what* has been measured and *how* it has been measured, decisions which necessarily involve value judgements as well as complex technical ones. These are not always transparent or understood. Partly because of this, some staff are unduly sceptical about performance data to the point of wanting to dismiss it; others are perhaps over-optimistic about what it can tell them. In those cases in the study where value added data was utilised well – i.e. as part of ‘intelligence-gathering’ for school development – it was often ‘championed’ by someone who understood and was interested in the data and was prepared to mediate the technical aspects for colleagues; moreover, the data was used heuristically and in conjunction with other information, both numerical and qualitative, to initiate discussion.

**In short, the meanings of data are socially constructed.** The evidence from the present study suggested that the meaning of value added data emerged from the *interaction* between:

- the actual ‘numbers on the page’;
- the politicised significance which performance data as a species of information currently possesses;
- what individual staff, particularly at head of department level, bring – often implicitly – to their perceptions of value added in terms of their knowledge of data analysis, their values and attitudes towards schooling and teaching, and their expectations of their pupils.

This view of how data is used in turn seems to pose the question: does value added data sometimes act as a *locus* for conversations staff would like to have, or anxieties they want to express, about their professional work and its value in the organisation? If so, then who controls (access to) value added data and for what perceived purposes the data is being used are central rather than secondary questions. How well can a ‘champion’ or sponsor of value added data manage and mediate the political and ethical as well as the technical aspects for colleagues? For value added analyses evidently constitute, at least at the moment, an intervention which needs to be managed. In some schools in the study, senior managers were not sure how to make best use of the analyses, and let departmental heads ‘*do their own thing*’; this seemed to lead to inertia in some departments and/or an uncoordinated strategy for how data could be used for raising pupils’ levels of attainment. Other headteachers, especially those who had been using the data over several years, were more

confident about what they wanted to achieve by the strategic use of data. Even so, there were still conspicuous differences between different departments in the same school. This suggested that it might be just as important to explore and understand why schools, or departments within schools, are *not* using value added data.

**Several of the studies** (Wikeley, 1995, 1998; Williamson and Fitz-Gibbon, 1990; Williamson *et al.*, 1992; Yang *et al.*, 1999) – **as well as the present one – suggest or allow us to infer that external input in the form of in-service training on value added data and its use is necessary.** This input is likely to be necessary for two different kinds of reason. First, the data itself needs explanation and mediation in a way which takes account of different schools' and staff's levels of understanding. One important aspect for staff to comprehend is that, however robust the data analysis is, the results are still provisional from a statistical point of view, and are therefore better used to raise questions than to make judgements. As a report on a conference on the use of value added measures of achievement in primary schools claimed: *'Clearly a great deal of in-service work would be needed to help all headteachers to develop insights as to the value and limitations of research findings'* (Education Journal, 1997, p. 14). Secondly, this study has suggested that an institution's incentive and capacity to use the information cannot be taken for granted. This study found that the extent to which schools made use of the data was fairly much in accordance with what might have been expected from their 'relative effectiveness', i.e. according to whether they had an *extrinsic* incentive to use it or not. Moreover, the use (or lack of use) to which value added data was put seemed to follow the existing contours of the style of leadership at both senior and middle management levels. It is not at all evident that the introduction of data necessarily and of itself brings about a culture supportive of systematic self-review and planning for improvement; it seems rather the other way round, that the culture may determine what use is made of the data. From a policy perspective, this raises the question of whether further external encouragement and support from, for example, the LEA may be needed. It also reinforces the idea that training and support which takes account of individual school contexts may be necessary in order for the potential of the data to be better understood and utilised.

Taken as a whole these nine points make a serious contribution, I believe, to an area of educational practice which continues to be a key policy focus, and of great importance to school and LEA managers and teachers. The discussion I have embarked on constitutes one kind of response to the challenge thrown down in 1993 by Patten, then Secretary of State for Education, that the value added agenda been obfuscating rather than illuminating. This thesis suggests that the complexities that have been introduced to the measurement of educational quality *have* been at least partially justified, in that they have contributed – albeit in a patchy way and at an uneven pace – to a deeper professional understanding and interpretation of performance data. The challenges posed to school management and culture by the introduction of value added which the study has identified are ones it is important to understand and to help schools address. The fact that it is possible to find schools which are in the process of creating the right conditions for using value added analyses developmentally is surely an optimistic sign.

(Even so, I think I would still have to say that some of the concerns raised by the Secretary of State in 1993 remain unresolved. Might school governors and parents, for example – many of whom are deeply concerned about the quality of education but are unlikely to have the time or inclination to tackle scholarly papers – entertain the suspicion that academics have been engaged in an exercise no more meaningful or helpful than counting angels on a pinhead? Whose professional and career interests have been well served by the development and introduction of value added? There are no easy answers to these questions, but those of us engaged in the value added ‘project’ need to be very mindful of them.)

I outlined earlier in this section a research agenda for further development in value added statistical approaches. In terms of our understanding of how value added analyses are or can be utilised, what is there still left to do? I have already suggested that the empirical evidence – including that furnished by the current study – has not quite delivered a ‘comprehensive and conclusive’ view of how value added analyses are being used in schools. The following section explains why, and what other kinds of evidence and/or understandings might need to be acquired.

#### 11.4 Reprise of the Main Problems Encountered in the Study: How to ‘Complete the Description’?

The main issues attendant on the study’s initial findings were as follows:

- the lack of use/understanding of QUASE analyses in some schools and departments (even although the schools in question might have been thought to have a strong commitment to the use of QUASE);
- differential use of QUASE by departments within the same school;
- the emergence in some, but not all, schools of a ‘champion’ of QUASE data;
- the difficulty of dissociating QUASE from the discourse of performance data more generally;
- the possibility that micro-political and cultural issues within institutions were shaping forces in understanding and using performance data.

From these issues it is possible to identify four areas for framing further knowledge, as follows:

- the desirability of articulating in greater detail the various kinds of discourse teachers construct around quantitative data, by giving an account of QUASE data analyses *in action*, particularly focusing on any difficulties staff have in accessing, accepting and interpreting the results;
- the desirability of understanding the micro-politics of each school from different perspectives (and perhaps moving towards a conceptualisation of how teachers construct their professional identities);
- the desirability of identifying the processes by which ‘the champion’ acquires or takes on that role;
- the possibility – or necessity – of my engaging from the outset in a more active role with school managers and teachers, to assist in the framing of a context conducive to the utilisation of data.

With hindsight, therefore, there were many additional leads I would ideally have wanted to pursue about the schools and their staff in order to compose a more comprehensive and cogent explanation; but, as I have said, the case study data was not of a kind which I could use to follow these issues up. The following sub-sections give a brief account of the issues and possible alternative approaches which might be relevant in each of these areas.

#### **11.4.1 Articulating the discourse around quantitative data: the participant observer**

An area which the study began to explore, but had not been designed to investigate in depth, was how teachers construct and engage in a discourse around pupil-related quantitative data. As I argue in detail in Section 11.4.4, in the particular form of inquiry that I undertook, my role as adviser or ‘agent’ appeared somewhat more conducive to understanding aspects of this than was my research role, in that it enabled me to observe at first hand, and to address, some of the defences and resistances staff constructed.

It is perhaps not coincidental, therefore, that some particularly pertinent and perceptive empirical work in this area has been carried out and written up by someone who is an LEA adviser (Dudley, 1997, 1999a, 1999b). The work – which I came across only towards the end of my own study – was based on an LEA initiative (the Essex Primary School Improvement programme) which *inter alia* generated pupil attitude survey data. Dudley’s particular contribution, in the context of the issues with which I am concerned here, was to record and classify teachers’ responses as they engaged with the data. The survey elicited pupils’ responses to their school experiences and their perceptions of themselves as learners, and the aggregated results were presented and discussed by an adviser – Dudley himself – to their teachers. The process was thus analogous to the construction of QUASE Profiles and follow-up sessions. Dudley observed the teachers’ responses *in situ* as they reacted to his feedback of pupil perception data, and he devised a four-way classification with which to try to categorise and explain these responses. He reports that teachers responded to the data as either good news or bad news; the data could produce:

- an action-orientated positive response to improve the issue behind the data (this type of response was equally likely to be provoked by either good news or bad news data);
- a passive ‘filing away’ of the issue behind the data;
- a passive rejection of the issue behind the data;
- an active denial or rejection of the issue behind the data. (Dudley, 1999a, p. 97.)

Crucially, he then says that the factors which promoted an action-oriented response were found in *the way the data were presented and the feedback managed*. He notes that important issues within the data could easily be lost and that the discussion could be side-tracked or brought to a halt by staff focusing on variables over which the school had little control, such as home background.

Factors associated with positive, action-orientated responses were:

- the availability of comparative data from schools felt not to be dissimilar;
- preparation for reading the data;
- prompts which focused discussion of pupil data on teaching, learning and issues which the school could influence and which developed speculation skills in teachers;
- some prior groundwork done by the discussion leader in identifying possible improvement strategies to feed into the discussion so that people did not feel ‘cornered’ by the data;
- the introduction of ground rules into the discussion to ensure that apparent good news or reactions to perceived external influences such as home background were sufficiently challenged;
- a [school] climate where speculation and reflection were promoted among staff;
- a feeling that the process generating the data was valid and could be trusted.

(Amalgamation of Dudley, 1999a, p. 97 and 1999b, p. 117.)

Dudley reports that, as part of a process actively managed by the schools and the LEA, staff reached the point where they did not feel the need to continue to use the code of confidentiality which had originally bound them. He concludes:

These matters combine to suggest that for a formal model... of school improvement to be sustained, a strong cultural model needs to be developing alongside – enabling trust, reflection, challenge and risk-taking to occur among schools, within the LEA and between the two. (Dudley, 1999a, p. 98.)

Although Dudley's work was with primary rather than secondary school teachers and revolved around pupil perception rather than performance data, it seems to me that his analysis is insightful and the findings could be highly applicable to the use by schools of value added performance data. The general force of his argument seems to me ungainsayable, namely that:

taking account of the emotional issues in effective use of data is vital if it is to be used with sustained effect... [T]he management of data presentation and feedback is crucial if it is to be taken on by schools and used for improvement purposes. (1999a, p.97)

The critical differences between the design and methods of Dudley's work and those of my research study lie in the following:

- the EPSI data was fed back to the teachers in actively-managed sessions (in my research study, I assumed that the QUASE data had been disseminated without knowing precisely whether and how this had been done);
- Dudley was able to capture, *via* tape-recordings of the feedback sessions, teachers' responses to the data as it was given to them and as they began to interrogate and use it (rather than, as in my research study as I designed it, interviewing staff some time after they had received the data and in a setting which was not one of observable use).

This suggests that a focus on teachers' actual, 'real-time' engagement with QUASE data might have been more conducive to exploring how teachers construct a discourse around data. However, this might be quite hard to set up as a research study, since it would mean knowing in advance when such discussions were going to take place – and it may be assumed that these are not confined to formal, planned in-service sessions which can be notified ahead of time to a researcher. One solution to this would be if the researcher played a more active advisory role – for which there are other advantages, as discussed in Sections 11.4.4 and 11.4.5 below.

#### **11.4.2 Understanding the micro-politics of the school: insights from the literature on school culture and organisation**

As I wrote in Chapters Nine and Ten, I had an underlying notion that concepts developed in the literature which investigates how schools function as organisations might be useful in understanding some of the findings of this study. In the course of the fieldwork, I had been struck by how much of what is expressed in formal conversation with an outsider might be thought of as some kind of performance, i.e. an acting-out in semi-public of micropolitical tensions and stresses about status, resources and relationships which are manifested in less straightforward or explicit ways in the day-to-day interactions within the school.

Several articles from the literature on how schools function as organisations were accordingly consulted – including organisational and systems theory, conceptualisations of educational leadership, typologies of school culture, the management of change literature, and micropolitical approaches, such as small group theory, to the analyses of teachers' behaviour – and these seemed to have a lot to offer for the subject of this study. (In addition, some articles on the theme of the construction of teachers' professional identity were identified through a search of key-words: Beijaard (1995), Convery (1999), Maclure (1993). When reviewed in the light of Dudley's more focused work, however, these did not provide much additional insight into the particular issues in this study.)

Even though these cultural and organisational concepts could not be brought directly to bear on the fieldwork (for reasons given in previous chapters and essentially having to do with the restricted nature of the data collected), it is worthwhile summarising them as pertinent to the further study of the area.

Before doing so, a small piece of clarification ought to be attempted: some of the literature on school culture, or some aspects of some writers' arguments, state or assume that micro-politics equates to the tactical exploitation, manipulation or blocking of power, and in this context it often seems to slip into being used as a value-laden term (positive or negative). I wish to use the term here in a more neutral and descriptive way, as an analytical tool to understand the fissiparous nature of organisations and different people's perspectives at different levels of the hierarchy or with different kinds of authority or influence within them. (See Hoyle (1999) for a good discussion of this issue.)

The ideas which seem potentially most useful for contributing something to the understanding of the narratives presented in Chapter Nine above can be summarised thus:

- the exercise of, and challenge to, formal and informal power characterises the way organisations, including schools, work;
- the role played by teachers' beliefs and perceptions, whether consciously and/or collectively articulated or not, are crucial to an understanding of schools function and evolve – and, by extension, of how schools can improve;
- schools, and departments within them, may differ markedly in their organisational culture, partly because of differences in the beliefs, perceptions and values of the staff;
- because teachers' professional identity is tied up with their beliefs about the purpose and value of educational policy, micropolitical activity may be further stimulated by the policy context.

A brief discussion of the work from which these ideas are drawn follows.

- The most recent work in the school effectiveness and school improvement fields acknowledges the crucial role of psychological and intra-organisational factors. Sammons *et al.* (1998 p.33) note that '*school effectiveness researchers have paid too little attention to practitioners' perspectives and explanations of effectiveness*'; their landmark study (Sammons *et al.*, 1997) explores the extent to which practitioners' beliefs and perceptions about effectiveness are related to research findings about the key characteristics of effective schools. The importance of taking account of the '*habits of everyday teachers*' is remarked on by Gray *et al.* (1999 p.151, quoting Fullan, 1993) in another study of the links between educational outcomes and in-school processes. Other examples of work linking organisational culture and/or micropolitics with school improvement include Stoll, 1998; Busher and Harris, 1999; Hoyle, 1999, p.219.
- Such observations immediately expose the fact that teachers in the same school may have differing or even conflicting beliefs and perceptions. Sammons *et al.* (1998 p.33, citing Elliott, 1996 and White, 1997) say that '*staff can differ markedly in their views about what constitute appropriate goals for students and for schools*'. Bishop and Mulford (1999 p.179) that '*[teachers' efforts to advance their interests and beliefs] often result from, and produce further, **conflicting standpoints** about what and who matters in schools*' (my emphasis). It is not just matters of belief which may be at issue; Sammons *et al.* (1998 p.40) say '*...whole staff discussion of such issues [perceptions of strengths and weaknesses within the school] may well reveal multiple perspectives and could be painful.*' Furthermore, the fact that conflicts of values and beliefs do often arise amongst teachers in the same institution does not mean that those conflicts are always overt or

even recognisable as such: it is possible that they may become apparent only by chance or because of a particular kind of intervention.

- West goes further than Bishop and Mulford, and says (West, 1999 p.189, following Ball, 1987) that school is '*a place where differing ideological perspectives and differing ambitions and expectations are inevitable...*' One may take from this that not only is any school, as a norm, differently experienced and described by different people within it depending on their beliefs, their backgrounds and their perceived self-interests but also that any discussion of what happens in schools should be grounded in this understanding. This has not been universally acknowledged; for example, there are many forms of guidance and advice which pose the establishment of 'a collective vision and fully shared aims' as a prerequisite of institutional effectiveness.
- In contrast to the notion of a collective conscious vision is the idea of organisational culture. This is said (by Schein, 1985, quoted in Stoll, 1998) to be '*the deeper level of basic assumptions and beliefs that are shared by members of an organization, that **operate unconsciously**, and that define in a basic "taken-for-granted" fashion an organization's view of itself and its environment*' (my emphasis). If this is true, then it implies that, on the one hand, organisational culture helps to shape what individuals think and do about their job as formally constituted, their informal role and their assumed responsibilities, and similarly those of others, without their being aware of it. As Stoll (1998, citing Deal and Kennedy, 1983) says, culture becomes simply '*the way we do things around here*'. On the other hand, organisational culture is also to some extent the product of the enacted beliefs and behaviour of individuals with power and/or influence; the longer they occupy those positions, the more embedded the culture becomes, and the harder it may be for others inside the organisation to find effective ways of 'naming' the culture for themselves and others, still less of challenging it or changing its default. Indeed, the first of the two quotations given in this paragraph illustrates the point more tellingly than its author perhaps intended: the organisation has been personified to the point where it is said to have a view of itself!
- Talking specifically about subject departments, Busher and Harris (1999) argue that '*departmental cultures... represent the enacted views, values and beliefs of teachers and support staff about what it means to teach students in particular subject areas within particular institutional contexts. Such cultures will be represented through a variety of cultural artefacts... [including]: rituals; ceremonies; language and other agreed educational practices which people undertake.*' They claim that cultural symbols and rituals '*[help] participants construct their understandings of their work and their relationships with other people [and are] a potent means by which people in groups create and sustain a sense of identity.*' West (1999) argues that the pathological consequences of such psychological closeness – in terms not only of institutional divisiveness but also of the deteriorating quality of group judgements – outweigh the benefits to the group.

- For teachers, a large part of their identity must surely be a *professional* identity, however: one which is concerned with the praxis of pedagogy and the education of young people, as Sammons *et al.* (1998 p.33), quoted above, intimate. This is too large an issue to discuss here, but I have elsewhere (Saunders, 1998) drawn on evidence to suggest that teaching is perceived by many teachers as an activity with an inherently ethical significance. This might in turn suggest that teachers' responses to what they are asked or required to do will depend partly on how well the innovation fits with their sense of what is professionally acceptable as well as on how it impinges on their own job/career or on their group cohesion, even although this may not be wholly or explicitly articulated. As Gray *et al.* (1999, p.151) say, in their detailed study of improvement in 12 schools: '*Attitudes [to change] are ... intimately bound up with notions of the professional self [and] many [teachers] needed to be convinced that further change was necessary – their instinct was to resist*'. The constraining impact on policy implementation of '*[teachers'] efforts to advance their interests and beliefs*' is the explicit topic explored by Bishop and Mulford (1999 p.179).
- Introducing the idea of resistance means addressing directly the question of power; most commentators on micropolitics say that power – its acquisition and maintenance or challenges to it – is a large part of what is at stake in organisations. West (1999 p.193) points out that a number of studies of organisational functioning have found that '*informal influence is most frequently used to inhibit and frustrate, while formal authority is linked to initiation or development*'. This must be an especially cogent insight about power relationships in schools during periods when, as now, new educational initiatives seem to come thick and fast and the pressure for 'development' seems to be relentless. It also suggests that micropolitical activity does not revolve around intra-organisational issues only but is also stimulated by the external policy context. This has especial relevance to the perception of value added data on school performance – see below, final bullet point.
- Policy may even be responsible for generating micropolitical activity, as West argues (*ibid.* p.190): '*schools in England and Wales have never offered more scope for micropolitical influence than they do now – within the self-managing school.*' He notes '*the... recent phenomenon of informal (or semi-formal) groups within the school beginning to compete actively with each other in the new context provided by local management of schools.*' (p.193.)
- An observation to make about this whole area of investigation is that the use of terms like 'unconscious beliefs', 'symbols and rituals', 'sense of identity' and 'instinct' (found in some of the above quotations) signals that something other than a discourse about purely rational actions and activities is felt necessary. I believe these ideas can be usefully linked, for present purposes, to an idea put forward by Wise (1977), writing about schools as 'loosely coupled organisations'. He says: '*[The rational model] postulates that schools operate by setting goals, implementing programmes to achieve these goals and evaluating the extent to which the goals are attained. The goal-oriented process is assumed to be effectuated through a*

*bureaucratic distribution of formal authority and work responsibility. It is further assumed that the attainment of goals provides sufficient incentives to drive the system... [A]nalysis of interviews with teachers [of how they see school reality]... bears little resemblance to that imagined by those who believe in the rational model.'* All this implies a certain irreducible non-rationality about the way schools operate.

- Having said all this, it is important to bear in mind the possibility that '*micropolitical life in schools is infinitely variable and context-specific*' (Bishop and Mulford, 1999 p.179). This means that the researcher must be very cautious about imputing rationales and consequences found in one particular situation in one particular school to any other school or situation, and about generalising very far from local phenomena. At the same time, one can imagine that the researcher-*qua*-outsider may sometimes be the only person to access certain kinds or combinations of information: privy to several individuals' confidences on the one hand and able to make out the dim shape of the organisation's cultural unconscious on the other. This is evidently suspected by some interviewees – especially headteachers asking for a 'de-brief' at the end of the researcher's visit.
- Finally, in the present study, one must also consider the additional element specific to it: that is, value added data and what it symbolises. As argued in Section 11.3 above, there are grounds for believing that numerical or quantitative data, especially that which is complex, relatively new and seen to be 'high stakes', represents a threat to some kinds of school, departmental and/or group culture. This notion is given some indirect support by Maw (1999, cited in Chapter Five above), who shows that the press (including the *Times Educational Supplement*) coverage of 'value added' – which it may be surmised is more in the forefront of teachers' consciousness than academic papers – has largely been in the context of publishing the controversial and contested tables of school performance. Worse, the press coverage of key issues concerning the measurement and presentation of value added has not given a clear and coherent message: '*there appears to be some confusion in the press*' (p.8) is Maw's mild way of describing the collection of contradictory or ambiguous statements she quotes. Thus, far from functioning as an instrument of and for self-evaluation, value added data could actually serve in the current climate to contribute to '*personal alienation and disempowerment*' (Bishop and Mulford, 1999, p.179). Such possibilities must surely be kept in mind.

This summary of points is not intended to represent anything like overview of the literature: rather, it constitutes an initial discussion which might assist other studies to examine the cultural aspects of schools' interpretation and use of data with greater cogency; and eventually to work towards a provisional typology for classifying cultures according to their use of data. General cultural typologies have already been formulated at both school and departmental level (for example, Hargreaves, 1994, 1995; Harris, 1997, 1998; Stoll and Fink, 1996).

### 11.4.3 Understanding the champion's role

One of the aspects of the fieldwork in the present research study on which an exploration of micro-politics might be able to shed further light was the discovery that in some schools there was a senior member of staff who took responsibility for championing value added analyses in some way; and furthermore that this was usually associated with a better understanding within the school of the data.

It would be interesting to find out, first, whether this function – which was not necessarily formally recognised and, even where it was, seems to have been as a result of an existing, informally adopted role – is a common one or coincidentally associated with schools in this particular study. Second, it would be worth discovering whether there were any common characteristics shared by the people who had adopted this function in terms of:

- level of seniority
- length of service in the school
- subject discipline
- previous training in statistical methods
- participation in higher degree study (MEd, EdD, PhD) or in school-based research project (for example, *via* a DfEE Best Practice Research Scholarship grant)
- standing with colleagues.

The case study data was not of the right kind to explore the first of these issues, and not based on a sufficiently large sample of schools to investigate the second – which may be why no obvious pattern emerged from the schools in relation to these categories. Having said this, it is worth noting that the one ‘champion’ of value added who seemed to be a rather lone voice in the school, as distinct from being able to exert influence with his peers, was a head of department, not a member of senior management.

Third, it would be useful to compare the impact on schools’ use of value added data of input and training from an external facilitator with the presumably more informal and ‘hands-on’ contribution to colleagues’ understanding made by the champion. In doing so, one would need to take account of the internal cohesiveness of the school, and therefore to pay attention to the micro-political aspects discussed in the previous section. In another case study school, it appeared that the ‘champion’ of value added

analyses was having difficulty being heard by colleagues because of the degree of fragmentation and diversity in the school management structures and processes.

#### **11.4.4 ‘Pure’ research or ‘research-and-development’?**

Exploration of each of these areas could add to an understanding of the uses of value added data in schools, and with hindsight it is possible to construct an alternative methodology – as Section 11.4.5 goes on to do. However, a final, and highly significant, issue to explore in this review of the methodology is the belated impact on the study of the work I undertook in my professional role and its capacity to provide useable evidence. In this extended sub-section, I want, first of all, to examine what might have been gained (as well as lost) from merging the role of researcher and adviser/developer at an early stage of the study. I then take the opportunity to explore certain tensions – between ‘pure’ research and ‘research-and-development’, and between exploration and evaluation – which have been hard to resolve. Finally, I say something about the role and significance of teachers’ professional identity and autonomy which I think constitutes a further challenge for this study.

In Chapter Six I gave an account of how and why I took the view at the design stage that failing to distinguish between my research and my development roles would not have been a legitimate way to approach the fieldwork; this was on the grounds that the empirical evidence might have been contaminated if I had provided guidance on QUASE data to participants. In reviewing the methodology, however, it is appropriate to look again at both kinds of work (as adviser or ‘agent’ and as researcher), to give an overview of the different kinds of encounter I experienced in the two roles and to discuss what they actually contributed by way of evidence. This was made both possible and necessary because I became involved in an intensive programme of development work for an LEA at the same time as I was conducting the research study – as described in Chapter Ten.

The two dimensions of work – the research study and the commissioned advisory/‘agency’ work – had some basic features in common: detailed information on performance at school and departmental level had been sent to participating schools and I had studied the data closely; I met in person and by prior arrangement with school senior managers (always including the

headteacher) and with heads of the mathematics, English and science departments; QUASE data analyses were the ostensible focus of conversation.

The respective objectives of these meetings were nonetheless quite different. As has been described in previous chapters, the research interviews were ‘framed’ by a written research information sheet and initial invitational letter, as well as by the letters setting out the purpose of each of the visits and interviews – although it is possible that these were not always circulated to the people concerned. Confidentiality was explicitly assured and interviews were usually, although not invariably, held in private and on the ‘territory’ of the interviewee (Chapter Seven gives the details). There was no developmental aim, for example, to change people’s understanding or practice: indeed, this was to be avoided, since I assumed that this would have compromised the fieldwork evidence. Some of the interviews did not address the issue of QUASE data for reasons which have already been explained (Chapter Seven).

So far as the LEA-commissioned programme of visits was concerned, its primary purpose was to follow up the provision of QUASE Profiles so as to help senior and middle managers in the schools interpret the data and generate action points for development. The meetings were therefore always centred around a detailed discussion of the data and its implications. Moreover, the school meetings were held not under conditions of confidentiality but rather of accountability (Chapter Ten gives details). The meetings were usually held on ‘official territory’, such as the headteacher’s or deputy headteacher’s office or a conference room.

The broad contrasts I encountered between the two kinds of work programmes in their tenor and outcomes were instructive and can be summarised as follows:

- The LEA-commissioned meetings were highly focused on the data and its implications, an agenda that was driven by the LEA personnel present as well as by the author. Participants knew, although sometimes only vaguely, that this was part of the LEA’s strategy for school improvement as contained in the Education Development Plan. There was also an awareness that attention was being focused on under-performance in the secondary schools. The context itself was thus one of ‘heightened accountability’ and the interviews themselves tangibly reinforced this sense, with witnesses being present and intended outcomes requiring to be discussed in front of managers. Yet it emerged from some of the staff interviews that not all staff had been briefed by their headteacher about the

exact purpose of the meetings and/or about the author's role. Although a letter setting out the broad aim of, and arrangements for, the meetings had been sent by the LEA to the headteacher, this had not usually been circulated within the school. So, even though I began each meeting by stating that its purpose was to take participants through the data, to explain what it meant and to discuss with them ways in which the analyses might be used to plan further activities and suggest possible changes to pedagogy, etc., this was not sufficient to allay the anxieties and suspicions of some participants about the ground-rules and/or consequences of the meeting. The conditions were such as to predispose them to hear what they expected to hear: the general context of heightened accountability and lack of clear briefing by the headteacher evidently meant that the default for these participants was to construe the status of the author as equivalent to visiting inspector, which they indicated in a variety of ways (deference, defensiveness or even hostility). Unsurprisingly, then, some participants seemed to feel themselves to be on the receiving end of an inquisition rather than a participant in a discussion. It may be surmised that this rather suited the intentions of some headteachers. Arguably, therefore, the combination of conspicuous accountability and the *de facto* ambivalence surrounding my status allowed my role to be constructed as that of an 'agent' with undefined but possibly extensive power and/or influence.

- Comparing these meetings with the research interviews, the content of the latter was more diffuse and comparatively unfocused on QUASE data. Some interviewees, particularly headteachers, managed the interview at least partly according to their own agenda (as described in Chapters Seven and Nine); this was made that much easier when, as in several cases, they did not have much to say about QUASE data and the time had somehow to be filled! Interviewees often took the opportunity to express their views on issues related to value added data, such as published performance data, and occasionally on any other policy about which they held strong opinions. There was no sense of the author being perceived as an agent or inspector, although it appeared that on one or two occasions she was assumed to have a 'hot-line' to the DfEE *via* which she could relay practitioners' messages to policy-makers.

The contrasts between these types of encounter raise an immediate question: could these differences be attributable to my behaving in a significantly different way in my role as researcher compared with that of 'agent'? The answer is obviously hard to assess with certainty, but I think that what counted was probably not so much any variations in my demeanour or speech as the perceptions people held of my role – and, as I have just indicated, I have reason to believe that my role as adviser/agent encouraged more staff to regard me (rightly, in some ways) as an influential presence. They were consequently impelled by the context and circumstances in which my visit took place to respond more actively to the challenges posed by the data.

So then the important issue for the study is whether the evidence I acquired as an ‘agent’ differed in kind and/or in principle from that which I acquired as researcher. For a while, as I embarked on the programme of work for the LEA, I attempted to maintain a distinction between the two sources of evidence – to the extent of almost trying to ignore what I was finding out from the developmental work. In the end, however, I realised I could not sustain this self-denying epistemological ruse; and indeed I then came to believe that the observations I was able to make in my role as adviser or ‘agent’ gave me more insight than did many of my interviews in the research study. To some extent, this was because I was present at the time when staff were engaging with the data and thus could use direct personal observation of what some of their difficulties were – but this is actually a question of timing and could in theory have been built into the research study (*pace* the difficulty I mentioned in Section 11.4.1 above). Much more significantly, my role as agent gave me a richer understanding on the one hand of how the political context could and did affect teachers’ acceptance of value added analyses; and on the other of what steps teachers could be encouraged to take, having overcome their reservations and defences, to bring about change and development. (I should make it clear that I am not positing a general contrast between research and development work, but only giving an interpretation of a particular set of circumstances arising from in the way I conducted the study.) So I felt I had no recourse but to make use in the thesis of the evidence I acquired as an ‘agent’, despite having rejected this approach at the beginning.

Nonetheless, I also conclude that I could not have based a valid study entirely on data gathered in my professional role. This is not because of the risk of ‘contamination’ of the evidence – which I now see as rather less important, since the input of some initial training and guidance seems to be a fairly critical factor in how far school staff can then actively use the data. The greater risk seems to me that of never being able to know how far people’s perceptions of my role as an adviser/‘agent’ might have inhibited or otherwise distorted their willingness to discuss the issues which mattered to them; and/or to present themselves as being more positive towards value added data than they actually felt. Some of what was said by teachers to me *qua* researcher would, I judge, not have been thought acceptable to say to me *qua* adviser. This raises the question, what is the perceived and/or appropriate level and kind of accountability (= ‘truth-telling’) of participants towards individual researchers and to the research-and-development agenda in general?

I do not feel that the study allows me to answer this question definitively, although the fact that I became aware of it was due to my being impelled to address this dual role. Researchers are not usually in a position to demand accountability from participants; but they can take precautions against being told only what participants assume they want to hear: and however involved researchers become in development work they should consider how best to protect some essential impartiality, how best to ensure at least some detachment from implementation. They should also consider how to engage teachers' trust and collaboration in the research so that a compliance model is not the default. I believe I learnt how to do some of this better than previously through comparing the two kinds of encounter.

This leads me on to a discussion of the fact that there has been an unresolved tension for me in the work between exploration and evaluation. I think it is fair to say that during the course of undertaking this doctorate, I travelled from a position of unwitting to considered ambivalence. On the one hand, I wished to explore what was happening in schools around value added data, given the range and seriousness of the issues and challenges as I conceived them (and as presented in Chapters Two to Five); in the context of such an inquiry, teachers' attitudes, values and beliefs – of whatever disposition – and the impact of those on their understanding and use of value added data seem a supremely important human element to understand, the key intervening variable as it were. On the other, I could not (and did not ultimately want to) divest myself of the urge to apply the insights available from value added data, however probabilistic and provisional, for the benefit of schools and pupils. In this context, not all attitudes adopted by teachers can be conceived as equally valuable or valid and 'the heuristic' attitude, as I call it, is the one I would have to construe as being most conducive to change and development.

However, if this is a problematic ambivalence, I do not think I am alone in displaying it: most of the leading critical commentators whose work has at one time served to deepen our understanding of the conceptual, methodological and technical problems of 'value added' – Fitz-Gibbon, Goldstein, Gray, for example, in their different ways – and even in some cases to challenge the possibility of its practical as distinct from theoretical utility, have not forborne on other occasions to undertake developmental work for schools, LEAs and/or government agencies.

Ultimately, my position remains an evaluative one: as I said in Chapter Six (Section 6.4.2), my working supposition was that, in its design and intent, the study came closest to a qualitative evaluation. As the study progressed, throughout my analysis of the fieldwork evidence and in the suggestions I have made for improving the use of value added analyses, I have been centrally concerned with the possibility of using value added in a reflexive way for school development.

Along the way, however – and this brings me to the third major issue in this sub-section – I have found myself compelled to try and understand questions of teachers’ professional values and beliefs, not only in terms of their impact on the use of value added analyses but in their own right. This is because professional identity and the exercise of professional autonomy are integrally related to the practice of pedagogy, defined as ‘*[the] informed interpretation[s] of learners, knowledge and environments [by teachers] in order to manipulate environments in ways which help learners make sense of the knowledge available to them. It is an intense, complex and discursive act...*’ (Edwards, 2000, pp 5-6) – such a conception of pedagogy is arguably central to one way of talking about quality in education. This is a debate which impinges on, but goes far beyond the scope of, the present study; however, I think the questions it raises constitute a kind of ‘undertow’ in my thoughts about the uses of value added in schools. For the sake of professional honesty, therefore, it is right to try to explain what I mean.

The evidence and arguments I have brought to bear support the notion that establishing a sense of ‘ownership’ amongst teachers is important to generating developmental usage of value added data. In addressing the issue of ‘ownership’, I think we need a better understanding of what kinds of knowledge teachers feel moved and motivated by, and of which they can eventually feel ownership. A great deal is being said currently about the importance of ‘relevant’ research/data/information/evidence for both policy and practice (see, for example, Hargreaves, 1996 – his lecture to the Teacher Training Agency was a key source in the contemporary resurgence of this debate). I said in Section 11.3 above that data does not speak for itself: in particular, the links between performance data and pedagogy are not self-evident but have to be constructed by the teachers themselves. It might not be going too far to say that this study raises questions about the self-evidence of ‘relevance’. Bird *et al.*’s (1998, p. 35) definition of relevance in educational

research includes the notions of ‘importance’ and ‘contribution to existing knowledge’. If we accept this definition, the present study suggests that, so far as schools and teachers are concerned, we would then have to go on to ask ‘important for what and for whom?’ and ‘knowledge of what and for whom?’ In other words, it cannot be assumed that the policy context of raising standards – in which value added data as a form of information is most readily situated – constitutes a necessary and sufficient definition of importance as far as teachers are concerned: indeed, they may feel dominated and demotivated by such policy initiatives if these are imposed rather than negotiated. What teachers may then evince is ‘compliance with’ rather than ‘ownership of’.

Nor can it be simply assumed that value added analyses of performance constitute the kind of knowledge which is felt by teachers to contribute something worthwhile to what they already know or believe about the connection between pupils, progress and pedagogy. The convictions which some teachers and other educationists hold about this connection (and from which a scepticism towards value added data might imaginably ensue) would lead them to reject the notion that educational progress can be legitimately captured by numerical performance data based on test scores. Such a view might readily be encapsulated in the words of Stenhouse (1975), which actually pre-date the current preoccupations with performance:

Education as induction into knowledge is successful to the extent that it makes the behavioural outcomes of the students unpredictable.

So one thought I have had is that why and how teachers engage with some kinds of knowledge and not others is bound up with:

- their assessment of the perceived costs in relation to potential benefits of engaging with complex and difficult subjects like value added, when there is already a large workload outside as well as within the classroom for teachers to contend with;
- their construction of professional identity, particularly as this concerns:
  - the theories (albeit sometimes implicit) they have developed about their own learning as teachers: Elliott (1991, pp. 1–12) alludes to this relationship;
  - their values-base: what they think it is most important for young people to be able to accomplish as a result of their education;
- their particular institutional culture and its micro-politics.

Gray seems to insinuate that the notion of the ‘relevance’ of externally-derived knowledge may need to be questioned in the context of how teachers learn: *‘There are hardly any short-cuts – when it comes to learning how to change few teachers seem to be prepared to let others do the thinking for them.’* (Gray, 2000, p. iv). In the context of the present study, these ideas perhaps suggest that a management model – value added as an innovation to be managed – does not do sufficient justice to the professional milieu of schools in terms either of teachers’ cognitive processes and their theories about those processes, or of the institutional and political contexts in which they individually and collectively construct their cognition.

This all tends to confirm my view – intimated in Section 11.4.2 above – that we need to understand a great deal more about both the cognitive and the social processes involved in teachers’ professional learning and the construction of professional pedagogical identity before being able to say exactly how ‘evidence’ can lead to ‘improvement’.

The tension between research and disinterested inquiry on the one hand and evaluation and development work on the other has surfaced at critical points during the conduct of this study: I hope this discussion has shown how this has occurred and also indicated why I have not been able easily to resolve the issues. It may be, however, that further careful study could illuminate those issues better than I have been able to do.

#### **11.4.5 An alternative project design?**

So then, to return to the methodological review: in summary, I believe that the empirical findings from this study are insightful and revealing in themselves; but that they are given considerable added cogency by the more strongly evaluative evidence I could access through my professional developmental work. The tensions thus exposed make it highly likely that the application of other conceptual frameworks and empirical methods could further enhance these insights, although from quite different perspectives.

Some possibilities are as follows: so far as building a ‘comprehensive and conclusive’ evaluative evidence-base is concerned, the suggestions put forward in Chapter Ten about good practice in using value added analyses could be tested and refined through a research-and-development study with more departments in a large number of schools. OFSTED reports, the Beacon

Schools initiative and the practical knowledge of the many LEA and HEI advisers who are developing the use of performance data in schools could be drawn upon to identify suitable schools for such a study. One of the strands which it would be important to incorporate and develop is the role of facilitator (external or internal). Teachers and subject leaders would be encouraged to develop their interpretation and utilisation of value added analyses under the guidance of an expert adviser/facilitator who would manage the discussions so as to correct any factual misapprehensions and to steer them towards agreed action points for teaching and learning. Such an approach is amply supported by Dudley's work (see Section 11.4.1 above). Staff would be invited by the facilitator to reflect together on their processes of understanding, the needs and issues which arose, the actions they agreed and the evaluation strategies they devised. The facilitator would take a lead in working with other colleagues in the same school and/or in other schools to share, test and develop these discoveries. The outcomes could be written up as a series of evaluative case studies.

However, this kind of developmental project does not exhaust the research possibilities. If the subsidiary, more neutral impetus for my study has been the need to inquire and understand as distinct from the overriding urge to change and improve, then a different, more ethnographically-oriented study might be indicated in order to pursue, for example, the kinds of ideas suggested by the literature on micro-politics and/or teachers' professional identity (as outlined in Section 11.4.2 above). A worthwhile study of this kind would require:

- an analytical framework which made the micro-political and cultural aspects of the above processes explicit and capable of being classified;
- a conceptual framework for understanding how teachers' 'professional identity' and their professional cognitive processes impinge on, and/or are shaped by, the introduction of quantitative measures of performance (in this case, value added analyses);
- a 'real-time' frame, that is, a capacity to observe and record the in-school processes and the individual and collective reactions of staff, starting from the point in time when the analyses arrived in school and including an observation of whether, when, how and for what ostensible and covert purposes the data analyses were introduced and disseminated to staff.

There could be possibilities for incorporating comparative elements into the design, in such a way that (i) schools where data analyses were a relatively unused instrument could be compared, in terms of the salient kinds of professional culture and events, with schools where performance data was already part of the dominant discourse; and/or (ii) schools which had different profiles of performance (as discussed in Chapter Eight) could be similarly compared.

The considerations involved in this hypothetical re-designing of the empirical investigation lead, finally, to the point at which it is possible to take an overview of the implications of the study: to assess, in short, what the study has to say about the use of value added measures of school performance for school improvement.

### **11.5 Conclusions: Ways Forward for Value Added and School Improvement?**

Chapter Five noted that there has been an assumption on the part of the DfEE, QCA and OFSTED – as exemplified in the production and dissemination of the ‘Autumn Package’ – that the promulgation of performance data would help to bring about school improvement. In this concluding section, I briefly summarise how far the evidence from the present study suggests this is likely to be true and then (equally briefly) reiterate the key points about how further progress might be made.

The first point to make about the present study in the broader national context is that the circumstances were apparently highly conducive to the schools in question making use of value added information, in that:

- the schools – or rather, the headteachers/governors – had commissioned and paid for the analyses (rather than having the information imposed on them);
- the information was of direct relevance for the individual school (as distinct from information about national data for which they would have to generate their own comparisons);
- the information was accessible in the sense that the report was sent to the school (as distinct from the school having to seek out information, for example, from the DfEE website);
- the information was important for the (policy-driven) focus on performance.

The study found that in some case study schools the data was being disseminated widely and used constructively. Importantly, this was more likely to happen if staff had opportunities for constructing collective meanings through discussion and dialogue about the data. This was borne out by evidence from the advisory and developmental work I was simultaneously undertaking.

Nonetheless, despite the favourable pre-conditions in the QUASE schools, a variety of challenges and difficulties seemed to be in the way of at least some staff in all the schools using value added data to its fullest extent and for developmental purposes. So assertions that value added data will or should lead to school improvement need to be tempered by some countervailing evidence. We can readily surmise, by looking at the fieldwork findings in one way, that some possible reasons for this state of affairs are that:

- lack of structures for dissemination of the information;
- lack of time for staff to read and interpret the information;
- lack of competence/confidence amongst staff to understand the information;
- lack of inclination by staff to spend time or acquire competence in using an instrument whose helpfulness to their work was not immediately obvious (*'it tells us nothing new'*);
- lack of belief in performance data as an aid to pedagogical excellence.

These are the kinds of conclusions which a typical evaluation might draw; and they do have the capacity to suggest some ways in which schools, LEAs and government could act to improve the introduction and use of value added analyses for school improvement.

But it could be said that such an interpretation constitutes something of a deficit model. What else might the study's outcomes be telling us? What factors and forces might be at work which mean that a more proactive and positive approach is required? I hope that the evidence and arguments I have presented give strong grounds for suggesting that 'the psychology and sociology' of using data need to be better understood and respected. The force of Dudley's (1999a, p. 97, quoted in Section 11.4.1 above) claim, that *'taking account of the emotional issues in effective use of data is vital if it is to be used with sustained effect'*, seems a very important lesson to be drawn from the evidence. This is hardly surprising when the backdrop to the issue is seen to

be the volatility of the political context for raising educational standards, the complexity and ingenuity of the value added enterprise, and the ‘high stakes’ attached to the use of data in making judgements about schools’ and teachers’ performance – issues to which I trust this thesis has done adequate justice.

However, Dudley’s subsequent argument (1999a, p. 97) that ‘*the management of data presentation and feedback is crucial if it is to be taken on by schools and used for improvement purposes*’ does not seem to me to go quite far enough: for it is not just that the introduction and use of data needs to be managed better. I would contend that value added data analyses may actually inhibit the cause of school improvement if, for example, they encourage teachers to focus only on measurable learning outcomes, or to think of pupils’ ability as ‘fixed’ or to construe their progress as ‘predictable’. I suggested that we may need a deeper understanding of, and respect for, the kinds of knowledge which teachers construct to enable them to relate performance data to their own pedagogical practice, and which they consequently feel professionally moved and motivated by. This in turn suggests that the current heavy emphasis on external accountability – in which value added analyses are playing an increasing role – ought to be counter-balanced by an approach which values and trusts informed professional self-evaluation. The study supports these claims conceptually and practically: Chapter Nine expounded an analytical framework of teachers’ approaches to quantitative data in which the ‘heuristic’ approach was identified as the one most likely to enable informed questions about teaching and pupils’ learning to be posed; Chapter Ten then gave examples of how a self-evaluative strategy can work in practice, when supported by a strong vision for improving teaching and learning, together with ‘hands-on’ facilitation for the mediation and interpretation of data. Such a strategy can be construed as an integral part of the process of ‘*capacity building from within*’. Policy-makers may not be able to legislate for this, but they can respect and enable it.

So finally, I wonder whether – in the light of the title of the thesis – ‘value added’ should now be conceived not as a science so much as a technology, that is to say, the practical application of a system of knowledge; the ‘art’ in question, of course, being education. The task then is to find the right ways of harnessing the technology, which require that we understand and work with, rather than against, the diversity of professional, social and cultural – that is, human – values with which it dynamically interacts.

## REFERENCES

- ANON. (1992). 'The value-added tasks' (Digest), *Education*, **180**, 5, 31 July.
- ATKINSON, P., DELAMONT, S. and HAMMERSLEY, M. (1993). 'Qualitative research traditions.' In: HAMMERSLEY, M. (Ed) *Educational Research: Current Issues*. London: Paul Chapman.
- AUDIT COMMISSION (1991). *Two Bs or Not...? Schools' and Colleges' A-Level Performance*. London: Audit Commission.
- BALL, S. (1987). *The Micro-politics of the School: Towards a Theory of School Organization*. London: Methuen.
- BARBER, M. and WHITE, J. (1997). *Perspectives on School Effectiveness and School Improvement* (Bedford Way Papers). London: University of London, Institute of Education.
- BEIJAARD, D. (1995). 'Teachers' prior experiences and actual perceptions of professional identity', *Teachers and Teaching: Theory and Practice*, **1**, 2, 281-94.
- BIRD, M., GOMM, R., HAMMERSLEY, M. and WOODS, P. (1998). *Educational Research in Action*. Open University Study Guide for the MA in Education (E835). Milton Keynes: The Open University.
- BIRNBAUM, I. (1993). 'Value added variations' (Digest), *Education*, **182**, 20, 12 November.
- BISHOP, P. and MULFORD, B. (1999). 'When will they ever learn?: another failure of centrally-imposed change', *School Leadership & Management*, **19**, 2, 179-87.
- BLUESTONE, B. (1994). 'The inequality express', *The American Prospect*, **20**, 81-93.
- BROWN, S., DUFFIELD, J. and RIDDELL, S. (1995). 'School effectiveness research: the policy makers' tool for school improvement?' *EERA Bulletin*, **1**, 1, 6-15.
- BUCK, D. (1993). 'Value addition: a reappraisal of school exam leagues tables', *Educational Psychology in Practice*, **9**, 2, 89-93.
- BUSHER, H. and HARRIS, A. (1999). 'Leadership of school subject areas: tensions and dimensions of managing in the middle', *School Leadership & Management*, **19**, 3, 305-17.
- BYATT, A.S. (1997). *Babel Tower*. London: Vintage.

CONVERY, A. (1999). 'Listening to teachers' stories: are we sitting too comfortably?' *Qualitative Studies in Education*, **12**, 2, 131-46.

COOPERS & LYBRAND (1988). *Local Management of Schools*. London: DES.

CORNALL, P. and LOFTHOUSE, M. (1993). *Value Added – "Taking Account of the Social Economic and Cultural Context of Schools"* (CSCS Discussion Paper 2). Northampton: Centre for the Study of Comprehensive Schools.

CUTTANCE, P. and GOLDSTEIN, H. (1988). 'A note on national assessment and school comparisons', *Journal of Education Policy*, **3**, 2, 197-202.

DEAL, T.E. and KENNEDY, A. (1983). 'Culture and school performance', *Educational Leadership*, **40**, 5, 140-1. Cited in: STOLL, L. (1998). 'School culture', *SIN Research Matters*, **9**, (whole issue).

DEARING, R. (1993). *The National Curriculum and its Assessment: Interim Report*. London: SCAA.

DUDLEY, P. (1997). 'How teachers respond to pupil data.' Paper presented at the British Educational Research Association Annual Conference, University of York, York, 10-14 September.

DUDLEY, P. (1999a). 'Primary schools and pupil "data".' In: SOUTHWORTH, G. and LINCOLN, P. (Eds) *Supporting Improving Primary Schools: the Role of Heads and LEAs in Raising Standards*. London: Falmer Press.

DUDLEY, P. (1999b). 'Using data to drive up standards: statistics or psychology?' In: CONNER, C. (Ed) *Assessment in Action in the Primary School*. London: Falmer Press.

EDUCATION (1992). 'The value-added tasks' (Digest), *Education*, **180**, 5, 31 July.

EDUCATION JOURNAL (1997). 'Using value added measures of achievement in primary schools', *Education Journal*, **17**, 14.

EDWARDS, A. (2000). 'Researching pedagogy: a sociocultural agenda'. Inaugural lecture delivered at the University of Birmingham, November.

ELKIN, S. (1995). *Adding Value and Improving Performance*. London: City Technology Colleges Trust.

ELLIOTT, J. (1991). *Action Research for Educational Change*. Milton Keynes: Open University Press.

ELLIOTT, J. (1996). 'School effectiveness research and its critics: alternative visions of schooling', *Cambridge Journal of Education*, **26**, 2, 199-223. Cited in: SAMMONS, P., THOMAS, S., MORTIMORE, P. and WALKER, A. (1998). 'Practitioners' views of effectiveness', *Improving Schools*, **1**, 1, 33-40. FALLON, M. (1991). 'Education Matters' (Extract), Radio 5, 17 November.

FIELDING, M. (1999). 'Target setting, policy pathology and student perspectives: learning to labour in new times', *Cambridge Journal of Education*, **29**, 2, 277-87.

FITZ-GIBBON, C. (1991a). "'A" Levels: corrective comparisons', *Managing Schools Today*, **1**, 2, 44-5.

FITZ-GIBBON, C.T. (1991b). *Evaluation of School Performance in Public Examinations: a Report for the Scottish Office Education Department*. Newcastle: University of Newcastle upon Tyne.

FITZ-GIBBON, C.T. (1992). 'School effects at A level: genesis of an information system?' In: REYNOLDS, D. and CUTTANCE, P. (Eds) *School Effectiveness: Research, Policy and Practice*. London: Cassell.

FITZ-GIBBON, C.T. (1994). 'Three routes to the wrong answer', *Times Educ. Suppl.*, **4093**, 9 December, 11.

FITZ-GIBBON, C.T. (1995). 'Measuring value added in schools.' In: IMAC RESEARCH *Education and Training Statistics Conference Papers, Statistics Users Council Annual Conference, 22 November 1995*. Esher: IMAC Research Ltd.

FITZ-GIBBON, C.T. (1996). 'The Value Added National Project', *Head Teachers Review*, Spring, 18-20.

FITZ-GIBBON, C.T. (1997). *The Value Added National Project: Final Report. Feasibility Studies for a National System of Value Added Indicators*. London: SCAA.

FITZ-GIBBON, C.T. and TYMMS, P. (1993). *Value Added: a Perspective on the Contribution from Examination Results* (CSCS Discussion Paper 1). Northampton: Centre for the Study of Comprehensive Schools.

FITZ-GIBBON, C.T. and TYMMS, P. (1995). *The Value Added National Project: General Report. Issues to be Considered in the Design of a National Value Added System*. London: SCAA.

FRANKFORT-NACHMIAS, C. and NACHMIAS, D. (1996). *Research Methods in the Social Sciences*. London: Arnold.

FREEMAN, R. and KATZ, L. (1994). 'Rising wage inequality: the United States vs. other advanced countries.' In: FREEMAN, B. *Working Under Different Rules*. New York, NY: Russell Sage Foundation.

- FREIBERG, H. J. (1999). (Ed) *School Climate*. London: Falmer Press.
- FULLAN, M. (1992). *Successful School Improvement*. Buckingham: Open University Press.
- FULLAN, M. (1993). *Change Forces: Probing the Depths of Educational Reform*. London: Falmer Press.
- FURTHER EDUCATION DEVELOPMENT AGENCY (1995). *Current Developments in Value Added*. London: FEDA.
- FURTHER EDUCATION UNIT (1993). *Value Added in Further Education*. London: FEU.
- GALLAGHER, A. (1991). 'Comparative value added as a performance indicator', *Higher Education Review*, **23**, 3, 19-29.
- GOLDSTEIN, H. (1987). *Multilevel Models in Educational and Social Research*. London: Charles Griffin.
- GOLDSTEIN, H. (1997a). 'From raw to half-baked', *Times Educ. Suppl.*, **4229**, 18 July, 15.
- GOLDSTEIN, H. (1997b). 'Methods in school effectiveness research', *School Effectiveness and School Improvement*, **8**, 4, 369-95.
- GOLDSTEIN, H. (1997c). 'Value added tables: the less-than-holy grail', *Managing Schools Today*, **6**, 6, 18-19.
- GOLDSTEIN, H. and MYERS, K. (1996). 'Freedom of information: towards a code of ethics for performance indicators', *Research Intelligence*, **57**, 12-16.
- GOLDSTEIN, H. and THOMAS, S. (1995). 'School effectiveness and "value-added" analysis', *Forum*, **37**, 2, 36-8.
- GRAY, J. (1979). 'The statistics of accountability: two approaches to improving information for LEA decision-making', *Education Policy Bulletin*, **7**, 1, 23-42.
- GRAY, J. (1981a). 'A competitive edge: examination results and the probable limits of secondary school effectiveness', *Educational Review*, **33**, 1, 25-35.
- GRAY, J. (1981b). 'School effectiveness research: key issues', *Educational Research*, **24**, 1, 49-54.
- GRAY, J. (1994a). 'Relative values', *Times Educ. Suppl.*, **4094**, 16 December, 15.

GRAY, J. (1994b). *Value Added Approaches in School Evaluation: the Experiences of Three LEAs in England – Lessons and Challenges*. Edinburgh: The Scottish Office Audit Unit.

GRAY, J. (1996). ‘The use of assessment to compare institutions.’ In: GOLDSTEIN, H. and LEWIS, T. (Eds) *Assessment: Problems, Developments and Statistical Issues*. Chichester: John Wiley.

GRAY, J. (2000). ‘How schools learn: common concerns and different responses’, *Research Papers in Education*, **15**, 3, i–v.

GRAY, J., HOPKINS, D., REYNOLDS, D., WILCOX, B., FARRELL, S. and JESSON, D. (1999). *Improving Schools: Performance and Potential*. Buckingham: Open University Press.

GRAY, J., JESSON, D., GOLDSTEIN, H., HEDGER, K. and RASBASH, J. (1995). ‘A multi-level analysis of school improvement: changes in schools’ performance over time’, *School Effectiveness and School Improvement*, **6**, 2, 97-114.

GRAY, J., JESSON, D. and JONES, B. (1984). ‘Predicting differences in examination results between Local Education Authorities: does school organisation matter?’ *Oxford Review of Education*, **10**, 1, 45-68.

GRAY, J., JESSON, D. and JONES, B. (1986). ‘The search for a fairer way of comparing schools’ examination results’, *Research Papers in Education*, **1**, 2, 91-122.

GRAY, J., JESSON, D. and SIME, N. (1990). ‘Estimating differences in the examination performances of secondary schools in six LEAs: a multi-level approach to school effectiveness’, *Oxford Review of Education*, **16**, 2, 137-58.

GRAY, J., REYNOLDS, D., FITZ-GIBBON, C. and JESSON, D. (Eds) (1996a). *Merging Traditions: the Future of Research on School Effectiveness and School Improvement*. London: Cassell.

GRAY, J., REYNOLDS, D. and JESSON, D. (1996b). ‘The challenges of school improvement.’ In: GRAY, J., REYNOLDS, D., FITZ-GIBBON, C. and JESSON, D. (Eds) *Merging Traditions. The Future of Research on School Effectiveness and School Improvement*. London: Cassell.

GREAT BRITAIN. DEPARTMENT FOR EDUCATION (1994). ‘Value-added by schools between ages 5-16.’ Letter from Secretary of State to the Chairman of SCAA, 21 November.

GREAT BRITAIN. DEPARTMENT FOR EDUCATION. (1995a). *GCSE to GCE A/AS Value Added: Briefing for Schools and Colleges*. London: DfE.

GREAT BRITAIN. DEPARTMENT FOR EDUCATION (1995b). *Value Added in Education: a Briefing Paper from the Department for Education*. London: DfE.

GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE and WELSH OFFICE (1988). *National Curriculum: Task Group on Assessment and Testing: a Digest for Schools*. London: DES.

GREAT BRITAIN. PARLIAMENT. HOUSE OF COMMONS (1997). *Excellence in Schools* (Cm.3681). London: The Stationery Office.

HARGREAVES, A. (1994). *Changing Teachers, Changing Times: Teachers' Work and Culture in the Postmodern Age*. London: Cassell.

HARGREAVES, D. (1995). 'School culture, school effectiveness and school improvement', *School Effectiveness and School Improvement*, **6**, 1, 23-46.

HARGREAVES, D. (1996). 'Teaching as a research-based profession: possibilities and prospects'. Annual Lecture delivered to Teaching Training Agency, April.

HARRIS, A. (1997). 'Differential departmental culture and strategies for change.' Paper presented at the British Educational Research Association Annual Conference, University of York, York, 11-13 September.

HARRIS, A. (1998). 'Improving ineffective departments in secondary schools: strategies for change and development', *Educational Management & Administration*, **26**, 3, 269-78.

HARRIS, A. and HOPKINS, D. (2000). 'Introduction to Special Feature: alternative perspectives on school improvement', *School Leadership and Management*, **20**, 1, 9-14.

HARRIS, A., JAMIESON, I and RUSS, J. (1997). 'A study of "effective" departments in secondary schools.' In: HARRIS, A., BENNETT, N. and PREEDY, M. (Eds) *Organizational Effectiveness and Improvement in Education*. Buckingham: Open University Press.

HARRIS, A. and RUSS, J. (1995). *Pathways to School Improvement: Identifying Principles and Practices*. Sheffield: ED.

HEWITT, D. and HUTCHISON, D. (1995). *Measuring Value Added in GNVQ Courses. Pilot Study for Kent TVEI*. Unpublished report.

HILL, P.W. and ROWE, K.J. (1996). 'Multilevel modelling in school effectiveness research', *School Effectiveness and School Improvement*, **7**, 1, 1-34.

HOPKINS, D., AINSCOW, M. and WEST, M. (1997a). 'Unravelling the complexities of school improvement: a case study of the Improving the Quality of Education for All (IQEA) project.' In: HARRIS, A., BENNETT, N. and PREEDY, M. (Eds) *Organizational Effectiveness and Improvement in Education*. Buckingham: Open University Press.

HOPKINS, D., HARRIS, A. and JACKSON, D. (1997b). 'Understanding the school's capacity for development: growth states and strategies', *School Leadership and Management*, 17, 3, 401-11.

HOYLE, E. (1999). 'The two faces of micropolitics', *School Leadership & Management*, 19, 2, 213-22.

HUTTON, W. (1996). 'Educated guesses are wide of mark', *Guardian*, 8 January, 15.

JESSON, D. (1992). 'Valuable addition' (Digest), *Education*, 180, 4, 24 July.

JESSON, D. (1996). *Value Added Measures of School GCSE Performance: an Investigation into the Role of Key Stage 3 Assessments in Schools. Interim Report* (DfEE Research Studies 14). London: HMSO.

KELLNER, P. (1997). 'Hit-and-miss affair', *Times Educ. Suppl.*, 4238, 19 September, 23.

KENNEDY, J. (1992). 'Adding value to raw scores', *Managing Schools Today*, 1, 4, 13-14.

LAUDER, H., JAMIESON, I. and WIKOLEY, F. (1998). 'Models of effective schools: limits and capabilities.' In: SLEE, R., WEINER, G. and TOMLINSON, S. (Eds) *Effective for Whom? Challenges to the School Effectiveness and School Improvement Movements*. London: Taylor and Francis.

MacBEATH, J., BOYD, B., RAND, J. and BELL, S. (1996). *Schools Speak for Themselves: Towards a Framework for Self-evaluation*. Glasgow: University of Strathclyde, Quality in Education Centre.

MacBEATH, J., MOOS, L. and RILEY, K. (1998). 'Time for a change.' In: MacBEATH, J. (Ed) *Effective School Leadership: Responding to Change*. Thousand Oaks: Sage Publications.

MacGILCHRIST, B., MYERS, K. and REED, J. (1997). *The Intelligent School*. London: Paul Chapman.

McGEEVOR, P., GILES, C., LITTLE, B., HEAD, P. and BRENNAN, J. (1990). *The Measurement of Value Added in Higher Education: a Joint PCFC/CNAA Project Report*. London: Council for National Academic Awards.

- McPHERSON, A. (1992). *Measuring Added Value in Schools* (NCE Briefing 1). London: National Commission on Education.
- MACLURE, M. (1993). 'Arguing for your self: identity as an organising principle in teachers' jobs and lives', *British Educational Research Journal*, **19**, 4, 311-22.
- MANDEL, E. (1962). *Marxist Economic Theory*. London: Merlin Press.
- MARSHALL, C. and ROSSMAN, G.B. (1995). *Designing Qualitative Research*. 3rd edn. London: Sage Publications.
- MAUGHAN, B. and RUTTER, M. (1986). 'Black pupils' progress in secondary schools: II. Examination attainments', *British Journal of Developmental Psychology*, **4**, 1, 19-29.
- MAUGHAN, B. and RUTTER, M. (1987). 'Pupils' progress in selective and non-selective schools', *School Organization*, **7**, 1, 50-68.
- MAW, J. (1999). 'League tables and the press – value added?' *The Curriculum Journal*, **10**, 1, 3-10.
- MESSICK, S. (1989). 'Meaning and values in test validation: the science and ethics of assessment', *Educational Researcher*, **18**, 2, 5-11.
- MORTIMORE, P. (1998). *The Road to Improvement: Reflections on School Effectiveness*. Lisse: Swets and Zeitlinger.
- MORTIMORE, P., SAMMONS, P. and THOMAS, S. (1994). 'School effectiveness and value added measures', *Assessment in Education*, **1**, 3, 315-32.
- MORTIMORE, P. and STONE, C. (1990). 'Measuring educational quality', *British Journal of Educational Studies*, **39**, 1, 69-82.
- MORTIMORE, P. and WHITTY, G. (1997). *Can School Improvement Overcome the Effects of Disadvantage?* London: University of London, Institute of Education.
- MYERS, K. (1995). (Ed) *School Improvement in Practice: the Schools Make a Difference Project*. London: Falmer Press.
- NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH (1996a). *Code of Practice*. Slough: NFER.
- NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH (1996b). *Code of Practice: Implementation Guidelines*. Slough: NFER.

- NICKELL, S. and BELL, B. (1995). 'The collapse in demand for the unskilled and unemployment across the OECD', *Oxford Review of Economic Policy*, **11**, 1, 40-62.
- NUTTALL, D. (1991). 'An instrument to be honed: tables do not reflect schools' true performance', *Times Educ. Suppl.*, **3924**, 13 September, 22.
- NUTTALL, D., GOLDSTEIN, H., PROSSER, R. and RASBASH, J. (1989). 'Differential school effectiveness', *International Journal of Educational Research*, **13**, 7, 769-76.
- NUTTALL, D.L., THOMAS, S. and GOLDSTEIN, H. (1992). *Report on Analysis of 1990 Examination Results*. London: Association of Metropolitan Authorities.
- PATERSON, L. (1991). 'Socio-economic status and educational attainment: a multi-dimensional and multi-level study', *Evaluation and Research in Education*, **5**, 3, 97-121.
- PATTON, M. Q. (1990). *Qualitative Evaluation and Research Methods*. London: Sage.
- POST-16 EDUCATION CENTRE (1996). *Value-added and Raising Achievement in the 14-19 Curriculum. Report from the National Conference 24 January 1996*. London: University of London, Institute of Education, Post-16 Education Centre.
- PRATT, J., TRAVERS, T. and BURGESS, T. (1978). *Costs and Control in Further Education*. Windsor: NFER Publishing Co.
- QUALIFICATIONS AND CURRICULUM AUTHORITY (1998). *Value Added Measures in School Performance Tables: Consultation Paper*. London: QCA.
- RAUDENBUSCH, S. and BRYK, A.S. (1986). 'A hierarchical model for studying school effects', *Sociology of Education*, **59**, 1, 1-17.
- REYNOLDS, D. (1995). 'The effective school: an inaugural lecture', *Evaluation and Research in Education*, **9**, 2, 57-73.
- REYNOLDS, D., BOLLEN, R., CREEMERS, B., HOPKINS, D., STOLL, L. and LAGERWEIJ, N. (Eds) (1996). *Making Good Schools: Linking School Effectiveness and School Improvement*. London: Routledge.
- REYNOLDS, D., HOPKINS, D. and STOLL, L. (1993). 'Linking school effectiveness knowledge and school improvement practice: towards a synergy', *School Effectiveness and School Improvement*, **4**, 1, 37-58.
- RUDDUCK, J. (1996). *Pupils' Experiences of Teaching and Learning*. Swindon: Economic and Social Research Council.

RUDDUCK, J., CHAPLAIN, R. and WALLACE, G. (1996). (Eds) *School Improvement: What Can Pupils Tell Us?* (Quality in Secondary Schools and Colleges Series). London: David Fulton.

RUTTER, M., MAUGHAN, B., MORTIMORE, P. and OUSTON, J. with SMITH, A. (1979). *Fifteen Thousand Hours: Secondary Schools and their Effects on Children*. London: Open Books.

SAMMONS, P. (1996). 'Complexities in the judgement of school effectiveness', *Educational Research and Evaluation*, **2**, 2, 113-49.

SAMMONS, P., THOMAS, S., MORTIMORE, P. and WALKER, A. (1998). 'Practitioners' views of effectiveness', *Improving Schools*, **1**, 1, 33-40.

SAMMONS, P., THOMAS, S. and MORTIMORE, P. with WALKER, A., HIND, A., SMEES, R., BAUSOR, J. and CAIRNS, R. (1997). *Forging Links: Effective Schools and Effective Departments*. London: Paul Chapman.

SAUNDERS, L. (1996). 'Using value-added measurement of performance for school improvement: what, how and why?' Presentation to European Conference on Educational Research, Seville, Spain, 28 September.

SAUNDERS, L. (1997). 'Value added principles, practice and ethical considerations.' In: HARRIS, A., BENNETT, N. and PREEDY, M. (Eds) *Organizational Effectiveness and Improvement in Education*. Buckingham: Open University Press.

SAUNDERS, L. (1998). 'Who or what is school 'self'-evaluation for?' Paper presented at the European Conference on Educational Research, University of Ljubljana, Slovenia, 20 September.

SAUNDERS, L. and SCHAGEN, I. (1994). 'Developing a system for school self-evaluation based on quantitative data.' Paper presented at the British Educational Research Association Annual Conference, University of Oxford, Oxford, 9 September.

SAUNDERS, L. and THOMAS, S. (1998). 'Into an uncertain world', *Times Educ. Suppl.*, **4280**, 10 July, 26.

SCHAGEN, I. (1991). 'Beyond league tables. How modern statistical methods can give a truer picture of the effects of schools', *Educational Research*, **33**, 3, 216-22.

SCHAGEN, I. (1996). *QUASE: Quantitative Analysis for Self-Evaluation. Technical Report 1996: Analysis of GCSE Cohorts 1993 to 1995*. Slough: NFER.

SCHAGEN, I. (1997). 'Value added taxes the statisticians', *Times Educ. Suppl.*, **4210**, 7 March, 14.

SCHEIN, E.H. (1985). *Organizational Culture and Leadership*. San Francisco, CA: Jossey-Bass. Cited in: STOLL, L. (1998). 'School culture', *SIN Research Matters*, 9.

SCHOOL CURRICULUM AND ASSESSMENT AUTHORITY (1994). *Value-added Performance Indicators for Schools*. London: SCAA.

SCHOOL CURRICULUM AND ASSESSMENT AUTHORITY (1997a). *Value Added Indicators for Schools. Consultative Paper: Primary*. London: SCAA.

SCHOOL CURRICULUM AND ASSESSMENT AUTHORITY (1997b). *Value Added Indicators for Schools. Consultative Paper: Secondary*. London: SCAA.

SILVERMAN, D. (1993). *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*. London: Sage Publications.

SLAVIN, B. et al. (1992). *Success for All: a Relentless Approach to Prevention and Early Intervention in Elementary Schools*. Arlington, Va.: Educational Research Service.

SLEE, R., WEINER, G. and TOMLINSON, S. (1998). (Eds) *Effective for Whom? Challenges to the School Effectiveness and School Improvement Movements*. London: Taylor and Francis.

SMITH, D. and TOMLINSON, S. (1989). *The School Effect: a Study of Multiracial Comprehensives*. London: Policy Studies Institute.

SMITH, P. (1995). 'On the unintended consequences of publishing performance data in the public sector', *Journal of Public Administration*, 18, 2/3, 277-310.

SOCIETY OF EDUCATION OFFICERS (1995). *Value Added and School Improvement*. Manchester: SEO.

SPOURS, K. and HODGSON, A. (1996). *Value-added & Raising Attainment: a Formative Approach. A Resource Pack for Practitioners*. Poole: BP Education Service.

SPOURS, K. and YOUNG, M. (1994). *Enhancing the Post-16 Curriculum: Value-added Perspectives* (Post-16 Education Centre Report No.10). London: University of London, Institute of Education.

STENHOUSE, L. (1975). *An Introduction to Curriculum Research and Development*. London: Heinemann.

STOLL, L. (1998). 'School culture', *School Improvement Network Research Matters*, 9, (whole issue).

STOLL, L. and FINK, D. (1992). 'Effecting school change: the Halton approach', *School Effectiveness and School Improvement*, **3**, 1, 19-41.

STOLL, L. and FINK, D. (1996). *Changing Our Schools: Linking School Effectiveness and School Improvement*. Buckingham: Open University Press.

STRONACH, I. and MACLURE, M. (1997). *Educational Research Undone: the Postmodern Embrace*. Buckingham: Open University Press.

SWEETMAN, J. (1997). *Curriculum Confidential 7: the Complete Guide to the National Curriculum*. Westley: Courseware Publications.

THOMAS, S. (1998). 'Value-added measures of school effectiveness in the United Kingdom', *Prospects*, **28**, 1, 91-108.

THOMAS, S. (1999). Optimal Multilevel Models of School Effectiveness: Comparative Analyses Across Regions (ESRC Full Report of Research Activities and Results). Unpublished report.

THOMAS, S. and GOLDSTEIN, H. (1995). 'Questionable value', *Education*, **185**, 11 March, 17.

THOMAS, S. and MORTIMORE, P. (1996). 'Comparison of value-added models for secondary-school effectiveness', *Research Papers in Education*, **11**, 1, 5-33.

THOMAS, S., SAMMONS, P., MORTIMORE, P. and SMEES, R. (1997). 'Stability and consistency in secondary schools' effects on students' GCSE outcomes over three years', *School Effectiveness and School Improvement*, **8**, 2, 169-97.

TYMMS, P. (1997a). *The Value Added National Project. Technical Report: Primary 3. Responses of Headteachers to Value-added and the Impact of Feedback*. London: SCAA.

TYMMS, P. (1997b). *The Value Added National Project. Technical Report: Primary 4. Value-added Key Stage 1 to Key Stage 2*. London: SCAA.

TYMMS, P.B. (1990). 'Can indicator systems improve the effectiveness of science and mathematics education? The case of the UK', *Evaluation and Research in Education*, **4**, 2, 61-73.

VINCENT, L. (1997). *The Value Added National Project: Analysis of 1994 Key Stage 3 Data Matched with 1996 GCSE Data. Secondary Technical Report No.3*. London: SCAA.

VISSCHER, A., KARSTEN, S., BOSKER, R., BRESSOUX, P., DEMAILLY, L., FITZ-GIBBON, C., GOLDSTEIN, H., MEURET, D., SCHAGEN, I., TYMMS, P., WEST, A. and WRAGG, E. (2000). Issues in Public School Performance Indicators. Unpublished report.

WEBSTER, W.J., MENDRO, R.L. and ALMAGUER, T.O. (1994). 'Effective indices: a "value added" approach to measuring school effect', *Studies in Educational Evaluation*, **20**, 113-45.

WEST, M. (1999). 'Micropolitics, leadership and all that...: the need to increase micropolitical awareness and skills of school leaders', *School Leadership & Management*, **19**, 2, 189-95.

WEST, M. and MOORE, S. (1998). 'Using added-value data to improve pupil attainment: a case study of practice in one secondary school.' Paper presented at the International Congress on School Effectiveness and Improvement 1998, University of Manchester, 4-7 January.

WHITE, J. (1997). 'Philosophical perspectives on school effectiveness research.' In: BARBER, M. and WHITE, J. *Perspectives on School Effectiveness and School Improvement* (Bedford Way Papers). London: University of London, Institute of Education. Cited in: SAMMONS, P., THOMAS, S., MORTIMORE, P. and WALKER, A. (1998). 'Practitioners' views of effectiveness', *Improving Schools*, **1**, 1, 33-40.

WIKELEY, F. (1995). 'Dissemination of effectiveness research and school improvement.' Paper presented at the European Conference on Educational Research, University of Bath, 14-17 September.

WIKELEY, F. (1998). 'Dissemination of research as a tool for school improvement?' *School Leadership & Management*, **18**, 1, 59-73.

WILIAM, D. (1992). 'Value-added attacks: technical issues in reporting National Curriculum assessments', *British Educational Research Journal*, **18**, 4, 329-41.

WILLIAMSON, J. and FITZ-GIBBON, C.T. (1990). 'The lack of impact of information: performance indicators for A Levels', *Educational Management & Administration*, **18**, 1, 37-45.

WILLIAMSON, J., TYMMS, P. and HADDOW, M. (1992). 'ALIS through the looking glass: changing perceptions of performance indicators', *Educational Management & Administration*, **20**, 3, 179-87.

WILLMS, J. D. (1985). 'The balance thesis: contextual effects of ability on pupils' O-grade examination results', *Oxford Review of Education*, **11**, 1, 33-41.

WISE, A.E. (1977). 'Why educational policies often fail: the hyperrationalization hypothesis', *Curriculum Studies*, **9**, 1, 43-57.

WOODHEAD, C. (1997). 'Inspecting schools: the key to raising educational standards.' Presentation of 'The Last Word Lecture' at the Royal Geographical Society, London, 21 January.

YANG, M., GOLDSTEIN, H., RATH, T. and HILL, N. (1999). 'The use of assessment data for school improvement purposes', *Oxford Review of Education*, **25**, 4, 469-83.

**APPENDIX A.**  
**QUASE Sample Profile (tables and figures)**

**APPENDIX B.**  
**QUASE Sample Profile (text)**

**APPENDIX C.**  
**Invitational letter to QUASE schools**

**APPENDIX D.**  
**Holding letter to respondents**

**APPENDIX E.**  
**Research Information Sheet**

**APPENDIX F.**  
**Confirmatory letter**

**APPENDIX G.**  
**Non-participant letter**

**APPENDIX H.**  
**Semi-structured interview schedule**

**APPENDIX I.**  
**Sample Information for Governors**