

The Leadership Paradox: Can school leaders transform student outcomes?

Bernard Barker*

Centre for Educational Leadership and Management, Leicester, UK

(Received 24 May 2006; accepted 7 September 2006)

This qualitative case study of an exceptional school in the south of England challenges the hypothesis that transformational leaders significantly impact on student outcomes. Interviews with staff and students, together with classroom observation, confirm that the head, appointed in 1995, has played an important role in transforming internal processes and in changing the context of the school. Although the observed and reported behaviour of leaders, teachers, and students matches expectations from the literature, the consequences for student achievement are unclear. Background variables seem to explain most of the apparent improvement in student outcomes. An effectiveness framework that assigns disproportionate value to examination results seems to have created a leadership paradox, where heads reported to be transformational produce only limited gains in performance. The study concludes that the government's determination to assume a strongly positive relationship between leaders and outcomes has compromised the principle of evidence-informed policy-making and that we need a different approach based on a broadly defined, qualitative conception of student success.

Introduction

This paper explores the fault line between two policy themes that have marked the UK government's approach to modernizing public services since 1997. The imperative to "transform" the public sector seems to have overridden an equally strong official commitment to "evidence-informed policy development and delivery" (Levitt & Solesbury, 2005, p. 3). Although transformational leadership is an important strand in education policy, few studies suggest that school leaders have more than a small, indirect impact on achievement (Bell, Bolam, & Cubillo, 2003; Leithwood & Jantzi, 2000). The result is a troubling gap between a widespread belief in leadership

*Centre for Educational Leadership and Management, 162-166 Upper New Walk, Leicester LE1 7QA, UK. Email: bcb3@le.ac.uk

and the absence of hard evidence of leaders producing substantial improvements in school and pupil outcomes (Bush, 2004).

Studies of leadership effects have tended to adopt a quantitative approach, with the analysis of large-scale datasets seeking to confirm a positive relationship between leadership and outcomes. In general, these have found that school leadership has a small but significant effect (Hallinger & Heck, 1998). Research programmes in Canada and Australia, for example, have found “significant indirect effects” on student engagement (Leithwood & Levin, 2005, p. 8). A more complex picture has emerged from a survey of 3,500 Year-10 students and 2,500 of their teachers and principals in South Australia and Tasmania. The data have led the authors to emphasize the interrelationship of a wide range of variables and to question the Hay–McBer model of transformational leadership adopted by the British government (Mulford, Silins, & Leithwood, 2004, p. 14). A very recent study based on 20,000 students enrolled in 250 American schools has found that “organic management,” including supportive leadership and staff collaboration, had no effect on “achievement growth” (Miller & Rowan, 2006, p. 242). Aware of this inconclusive evidence, the Department for Education and Skills (DfES) has commissioned research to further investigate the links between leadership and student outcomes (DfES, 2006).

This study adopts a different approach, and aims to challenge rather than confirm the theory that certain types of leadership necessarily produce improved student attainment. At first sight, The Shire School, a full range rural comprehensive in the south of England, described by inspectors as “very good and extremely effective” (Office for Standards in Education [OfSTED], 2000, p. 8), appears to confirm the hypothesis that strong leaders significantly impact on student outcomes. On closer inspection, however, there is evidence that, while outstanding leadership has enhanced the character and quality of educational provision, the consequences for student achievement are unclear and unproven. As at Hillside (Barker, 2005), inspectors and insiders reported a dramatic improvement in the school’s effectiveness that does not seem to have led to transformed organizational outcomes.

The evidence presented below raises important questions neglected by current policy, although a single case study cannot, of course, confirm or refute the general proposition that certain types of leadership tend to produce much better test and examination scores. Is it misleading for government agencies to emphasize “transformation” and improved academic results when even a much-praised leader seems to have had only a limited impact on outcomes? Does it make sense to evaluate leadership mainly in terms of an indicator that is conditioned by the “difficult to influence background characteristics of pupils” (Scheerens, 1989, p. 71)? Has the official emphasis on data and performance produced a distorted picture of how leaders contribute to improvement?

Transformational Leadership

Ministers and civil servants have argued consistently that:

By tackling our management and leadership deficit with real vigour, we will unlock the doors to increased productivity, maximise the benefits of innovation, gain advantage from

technological change and create the conditions for a radical transformation of public services. (DfES, 2002)

In education, the Prime Minister and his colleagues (Blair, 2001; Kelly, 2005; Morris, 2002) have continuously promoted the theme of transformational leadership, insisting that “Good heads can transform a school” (Department for Education and Employment [DfEE], 1997a, p. 46). Government agencies and supportive academics emphasize the scope for headteachers to change the context, vision, organization, and performance of their schools (Fullan, 2003). When the National College for School Leadership (NCSL) was launched with a budget of £60 million per year (Bolam, 2004, p. 263), the then Secretary of State announced that it would “play a key role in the Government’s strategy to transform our schools, drive up standards and ensure that every school is excellent or improving or both” (Blunkett, 2000, p. 1).

The NCSL (2001, 2003) emphasizes the added value produced through distributed leadership. Individual leaders, or groups of leaders empowered by their heads, are recommended to adopt styles and strategies to induce heightened motivation and change, especially in student outcomes (Bell et al., 2003; Fullan, 2003; Gold, Evans, Earley, Halpin, & Collarbone, 2003; NCSL, 2003). Effective leaders are expected to transform cruising, strolling, and struggling schools (Stoll & Fink, 1995).

Advocates of the leadership hypothesis are explicit about the typical actions associated with transformational models. These include:

mobilizing commitment to an explicit educational vision... coaching and mentoring designed to support individuals and increase leadership capacity generally; visible dispersal of leadership responsibility throughout the staff group whose members are trusted to initiate and complete tasks; and group decision-making that is highly participatory, open and democratic. (Gold et al., 2003, p. 128)

Earlier conceptions of transformational leadership have had to be extended because “there is no evidence to suggest that, on its own, it brings about anything but modest improved consequences for pupil outcomes” (Gold et al., 2003, p. 129). Leaders are encouraged to attend to instructional issues and to concentrate on the “behaviours of staff as they engage in activities directly affecting the quality of teaching and learning” (Gold et al., 2003, p. 129). Successful leaders “don’t impose goals but work with others to create a shared sense of purpose and direction,” and galvanise “effort around ambitious goals and by establishing conditions that support teachers.” They offer intellectual stimulation and individualized support, they develop the organization and strengthen the school’s collaborative culture, and they provide instructional guidance while empowering others to make significant decisions (Leithwood & Riehl, 2003, pp. 2–3). Their organizational impact is indirect rather than direct, as in earlier models, and is mediated through middle leaders, teachers, and the internal conditions they develop.

An evaluation of 20 high-performing specialist schools found that, rather than implementing a shopping list of improvements, successful multiskilled heads were encouraging a greater “interconnectedness.” Multiple inputs were integrated into a wider vision so that “superb student – teacher relations,” “a shared sense of vision,”

and a “genuine working together” generated an institutional buzz. Energetic, visionary leadership, a focus on the individual student, the active use of performance data, a broad and flexible curriculum, and enhanced status and resources contributed to a climate where teachers were prepared to go “the extra mile” (Judkins & Rudd, 2005, p. 4). An NCSL-sponsored enquiry into learning-centred leadership (Madison & Allison, 2004) confirmed the central role of the headteacher and leadership team, illustrated the extent to which school culture facilitates school improvement, and indicated the need to concentrate on raising achievement. Another study found that in all types of schools the effective use of data was valuable in challenging the expectations of staff and children and in stimulating questions and discussion that prompt improvement (Kirkup, Sizmur, Sturman, & Lewis, 2005).

Improving Results

This stress on leadership has been accompanied by a parallel emphasis on data and target setting (Gorard, Rees, & Selwyn, 2002; Gorard, Selwyn, & Rees, 2002). Every school is required to evaluate its performance and to set “challenging targets” (DfEE, 1997b), using data provided in the Performance and Assessment Report (PANDA) and in the Pupil Achievement Tracker (PAT), soon to be merged in a “new product” to be known as RAISEonline (Reporting and Analysis for Improvement through School Self-evaluation). This is intended to provide schools “with a tool for reviewing their performance data in greater depth as part of their self evaluation and target setting” (OfSTED, 2006).

A strong consensus has developed amongst policy-makers (e.g., Barber, 2000), inspectors (e.g., OfSTED, 1993), and academics (e.g., Fullan, 2003; Reynolds et al., 1996) that well-trained leaders have the ability to transform the organizational effectiveness of their schools and that effectiveness should be measured in terms of substantially better test and examination results (DfES, 2005b).

There is insufficient evidence, however, to inform or fully justify this policy. Early large-scale studies of school effects in general (Coleman et al., 1966; Jencks et al., 1972) concluded that achievement is mainly determined by family background, and that pupils’ subsequent careers are little affected by the quality of education they experience. Although these findings were challenged by Rutter, Maugham, Mortimore, and Ouston (1979) and Edmonds (1979), “later school effectiveness studies pointed to hardly greater differences between schools (in terms of explained variance) than Coleman and Jencks had done” (Scheerens, 1989, p. 70). Mortimore and Whitty (2000, p. 10) reported a “strong negative correlation between most measures of social disadvantage and school achievement,” while Thrupp (2001, p. 446) was disinclined to believe that school effectiveness research “can overcome the effects of social inequality.” Once social background and prior attainment are considered, the degree of between-school variance to be explained can be very small:

The extensive data sources which we have used gave consistently “low” estimates of the extent to which schools contributed differentially to pupils’ performance. In nine out of the eleven datasets the variance attributed to schools we estimated to be 8 per cent or

less... in six of the datasets the estimate was 5 per cent or less. (Gray, Jesson, & Sime, 1995, pp. 114–115)

There is even less data to suggest that leaders have a transformational impact on school effectiveness as measured by student outcomes. The EPPI systematic review (Bell et al., 2003), for example, found few empirical studies with unequivocal evidence of a direct leader impact on student outcomes and indicated that positive impact is usually indirect and mediated by teachers. On the basis of 41 studies conducted since 1980, Hallinger and Heck (1998) reported that principal effects are small and usually require exceptionally sophisticated research designs to detect.

Studies that associate desirable leadership characteristics with superior outcomes (e.g., Judkins & Rudd, 2005) can be insufficiently rigorous in evaluating claims that particular schools are more effective than others. The assertion, for example, that non-selective specialist schools outperform other nonselective schools (Jesson, 2001) is based on a value-added analysis that disregards the relatively high number of single sex, ex-grammar, and foundation status schools that have achieved designated specialist status and the additional resources that they have enjoyed (Gorard & Taylor, 2001).

The great majority of schools seem to be performing at levels that could be predicted from knowledge of their intakes. Schools found effective in one year continue to be effective subsequently (Gray et al., 1999). Across the system, socioeconomic status remains “the most powerful predictor of student success” (Leithwood & Jantzi, 2000, p. 422) and “in assessing relative performance,” student background factors are “paramount” (Gorard & Taylor, 2001). Clear links between improved student outcomes and “collaborative, democratic and distributed forms of leadership” have yet to be established (Harris, 2004, p. 13). The DfES (2004b, para. 22) has itself recognized that there is a “lack of consensus about the contribution of different elements to the linkages, indirect effects and mediating factors for improving school leadership.” The NCSL (Collarbone, 2001, p. 3) acknowledges the need for “objective and independent measures, which measure impact on school improvement.”

Despite these doubts and reservations, the DfES, the NCSL, and OfSTED continue to advocate “strong” leadership to transform our schools. Their sustained, orchestrated attention to leadership and performance has created the widespread conviction that a step change in results can be achieved, although the precise nature of the expected transformation is not defined (Dunford, 2002), and no agency has attempted to estimate the expected improvement. In the absence of conclusive evidence to support this emphasis on leadership and results, academics are beginning to argue for research that documents a wide range of school and student outcomes, not just those prioritized by government (Bush, 2004).

Methods

The persistent search for evidence of impact on outcomes is symptomatic of “pathological” social science (Park, 2000), where researchers “prefer using post hoc

statistical analysis to search for the phenomenon, rather than conduct any form of definitive test.” Despite numerous studies seeking the missing link between school leadership and results, “there appears to be no progress in the nature of the evidence over time” (Gorard, 2006, p. 87). Less effective heads, and other possible sources of discord, have received comparatively little attention (Barker, 2001). The problem stems from the inductive method criticized by Popper (1963, p. 54), who argued that theory cannot be “inferred from empirical evidence”. In his view, the true function of repeated observations and experiments in science is to test our hypotheses, not to confirm their predictions. Scientific theories should be tested by attempts to refute or falsify them.

The Shire School was selected, therefore, as a case where the characteristics described in the literature were most likely to be found, so that the theory that leaders can transform student results could be tested fairly. The inspection conclusion that there were “no major issues for action” (OfSTED, 2000, p. 8) was unusual, especially for a state comprehensive. The Shire was also listed as “outstanding” by Her Majesty’s Chief Inspector (HMCI) in 2002. The choice of a school that has experienced sustained, exceptionally strong leadership is a variant of extreme or deviant case sampling that “focuses on cases that are unusual or special. The findings of research on extreme cases can provide an understanding of more typical cases” (Gall, Borg, & Gall, 1996, p. 231).

Seventeen staff members were selected for interview in June 2005, such that age, gender, length of service, role (teaching or administration), and status were reflected within the constraints of a convenience sample inevitably conditioned by the availability of particular individuals (see Table 1 for a full list and for the codes used

Table 1. List of interviewees

| Interview | Role | Interview date | Years of service | Code in text |
|-----------|--------------------|----------------|------------------|--------------|
| 1 | Administration | 21 June 2005 | 14 | 1Ad |
| 2 | Administration | 21 June 2005 | 16 | 2AD |
| 3 | Assistant Head | 21 June 2005 | 14 | 3AH |
| 4 | Assistant Head | 21 June 2005 | 8 | 4AH |
| 5 | Assistant Head | 21 June 2005 | 8 | 5AH |
| 6 | Deputy Head | 22 June 2005 | 6 | 6DH |
| 7 | Head | 22 June 2005 | 10 | 7H |
| 8 | Head of Department | 21 June 2005 | 7 | 8HoD |
| 9 | Head of Department | 22 June 2005 | 13 | 9HoD |
| 10 | Head of Department | 22 June 2005 | 2 | 10HoD |
| 11 | Head of Year | 21 June 2005 | 12 | 11HoY |
| 12 | Head of Year | 22 June 2005 | 25 | 12HoY |
| 13 | NQT | 21 June 2005 | | 13NQT |
| 14 | Teacher | 21 June 2005 | 2 | 14T |
| 15 | Teacher | 21 June 2005 | 2 | 15T |
| 16 | Teacher | 22 June 2005 | 11 | 16T |
| 17 | Teacher | 22 June 2005 | 7 | 17T |

to ensure the anonymity of those concerned and for reference in the text below, for example, 7H refers to the headteacher). Interviews lasted approximately 30 min each, and were semistructured around issues emphasized in the literature (e.g., Gold et al., 2003).

Interview notes were word processed in a style as close as possible to a verbatim first person statement. Almost 15,000 words were recorded in total. Each interview was read and confirmed by the respondent as a fair account of the conversation. This method produces a version of an interview that is not as full or exhaustive as may be possible with transcription from tape or voice recordings (Lapadat, 2000) and may add to the risk of subjectivity and bias implicit in social investigations. There seems to be no escape from the dilemma that “transcription decisions both reflect the researcher’s theory and constrain theorizing,” and in this context even validity may be a social construction (Lapadat, 2000, pp. 208 & 210). As we do not have “independent, immediate and utterly reliable access to reality” (Hammersley, 1992, p. 69) and depend on the interpretations of witnesses for our understanding, there is little alternative to forming a judgement on the basis of the best evidence that can be gathered.

Interviews were compared and contrasted to triangulate respondents’ opinions, and the data were also collated and coded, using headings derived from the literature (e.g., distributed leadership, trusted to complete tasks, participatory group decision-making). During analysis, ideas about the conceptual significance of the data were recorded as memos (Day, Harris, Hadfield, Tolley, & Beresford, 2000; Punch, 2005). A feature of the study is the extent to which teachers, varying by age, gender, service, role, and status, were consistently enthusiastic in describing the role of the head and her impact on the quality of learning and teaching.

Classroom observation was undertaken in October 2005 to triangulate earlier interview comments about teacher–student relationships and to confirm the current relevance of judgements made by the OfSTED inspection team in 2000. Was there up-to-date evidence that: “Pupils’ and students’ very good learning owes a great deal to excellent relationships between them and with their teachers, whose management of lessons and learning are very good” (OfSTED, 2000, p. 21)?

A small opportunity sample of eight classes across the age and ability range was visited, so that half lessons were observed in English (3), mathematics (2), science (2), and geography (1). The number and choice of lessons was constrained by the timetable and other teacher commitments on the chosen day, as well as by the limited funds available for the study. Only one of the teachers interviewed in June was amongst those observed in October. Documentation, including schemes of work and lesson plans, was provided on arrival at each lesson, while the flow of classroom events and dialogue was detailed in contemporary handwritten notes. Fleeting conversations with students and teachers during and between lessons were followed by individual discussions at the conclusion of the afternoon session. Perceptions of the climate, learning, and behaviour observed during the day were checked with those concerned.

Although far removed from a systematic approach to classroom observation (Galton, Hargreaves, Comber, Wall, & Pell, 1999), ethnographic immersion in the life of the school (Wragg, 1999) encouraged confidence in the strongly positive

statements included in the OfSTED Report (OfSTED, 2000). My own experience of classrooms as a teacher, principal, inspector, PGCE tutor, and consultant enabled me to compare and contrast my observations with a wide range of other schools and settings. Such a small sample of classes has limited validity, however, and was primarily useful in providing evidence to confirm the continuing relevance of OfSTED judgements and the remarkably consistent view of learning and behaviour at The Shire School expressed by those interviewed and observed.

Other sources of information for the study include the OfSTED inspection report (OfSTED, 2000), online performance data (BBC, 2005; DfES, 2005a), the school's *Self-Evaluation Form* (The Shire School, 2005), a research paper presented by the head (Thomson, 2006), and the school brochure.

The Case Study: The Shire School

Under the leadership of the present head, Sara Thomson (pseudonym), since September 1995, The Shire School has emerged from a period of consolidation with a reputation as a successful 11 – 18 foundation comprehensive that serves a rural and generally prosperous community in southern England. The school has attracted people from a widely distributed variety of towns and villages. Students are now drawn from 26 primaries, and from three neighbouring counties. The school has been well placed to benefit from the many initiatives promoted by the government since 1997. Language College status (2002) has helped to maintain momentum, while Leading Edge status was achieved in 2004. This enhanced framework has facilitated numerous collaborative partnerships, especially around the leisure centre and the Language College (7H and 6DH). Eleven percent of compulsory age students were eligible for free school meals (FSM) in 2000, compared with the national mean of 16.5% (Gorard & Taylor, 2001; OfSTED, 2000).

Evidence from the case study is presented below to confirm the head's success in developing characteristics reported in the literature to be strongly associated with school effectiveness and improvement. To what extent have Sara Thomson and her colleagues improved organizational effectiveness and student outcomes by enhancing commitment and vision, dispersing leadership responsibility, and increasing leadership capacity?

Commitment and Vision

Students and teachers alike emphasized the head's personal role in creating the climate and conditions within which good teaching and learning can flourish. The head's ability to motivate her colleagues and inspire their commitment to improvement is widely recognized (15T). At her interview, she "promised to make this the best in the area" (7H) and 10 years later is still perceived to be "very driven and determined...like a dog with a bone...keeps at it all the time, she's adamant" (13NQT). Inspectors confirmed the strong leadership provided by the head and senior team (OfSTED, 2000).

Sara's family upbringing and Christian faith are the bedrock of her vision: "the traditional values associated with religion are what most parents want for their children . . . heads are the last bastions upholding those traditional values—you can be pushed away from them" (7H). Respondents saw the head as "the driving force behind all the improvements" (4AH) and were clear about their "ultimate goal . . . to maximize pupil progress" (3AH). One teacher summed up:

Sara is a remarkable person . . . she's so wonderful at the decisions, talking to people, knowing what is the best way to go, how to approach the staff—I've worked for a lot of people, she's so focused, a wonderful person to work for—makes the right decisions at the right times—the leadership pulls everyone together. (2AD)

A new teacher wondered whether women heads are more open to new ideas: "Men are perhaps more conscious about presenting themselves as a leader" (13NQT).

Sara herself believed that "being a woman has made me much more determined and tougher and more competitive" because you are "always having to prove yourself, people under-estimate you . . . when a little person arrives there are different assumptions" (7H). According to one teacher, the head's personal qualities had enabled her to be "successful in bringing in the same kind of people, who share that view, who fit into the whole scheme of things" (14T). These perceptions are fully representative of the interview data as a whole.

Dispersing Leadership Responsibility

When she was appointed in the summer of 1995, the head quickly assessed her colleagues' potential:

I did an analysis of their capacity for change—who wanted to change and progress, I worked out a see-saw model, intuitively assessing their reactions and responses. Who is moving in my direction? I worked closely with those who looked like wanting movement, for example the head of languages who said we would be ready to go for language college status in three years. (7H)

She was "continuously multi-tasking and knows what is going on" and had an eye for the "big picture," as well as "fine details" (5AH). She provided "clear leadership" and had the "balance right between giving people responsibility and allowing them to exercise it" (11HoY). She had a strong "people instinct" and was a "good personnel person" (5AH). She was "always encouraging, always praising and supportive" (13NQT). One colleague commented that she:

has a nose for a good appointment; has a gut instinct for the teacher or head of department who will fit in with the ethos of the school . . . she will know whether that is a Shire teacher or not. (5AH)

The head's strategy was to "grow" leaders within the school, challenging people early in their careers with extra opportunities and responsibility. All the members of the senior leadership team (SLT) had been promoted from within, having shown their

personal qualities and ability to contribute in other positions. The SLT was described as a “well-oiled machine” (13NQT), remarkable for its efficiency and helpfulness. One teacher with experience of five other schools commented that:

There is no blame, at a lot of schools it is always someone else to blame, people fester. Here there’s always someone there to take some of your workload, it comes right from the top, the head would sort it out if I were overloaded. (17T)

According to one assistant head, target setting also contributed to the effective distribution of authority and responsibility:

Leadership is widely dispersed across the school— huge responsibility is given. Target setting can give people more autonomy . . . highlights the excellent job that the vast number of departmental people are doing. (4AH)

The leadership team “make sure that we go through the middle managers—we trust the middle managers . . . give them responsibility for a chunk of the curriculum” (5AH). As a result, middle managers reported that they felt themselves to be leaders “empowered to do things” (5AH). Teachers in different departments described similar experiences of trust and collaboration:

The head seems to pick out the ones who are good; you know they are going to be reliable or trustworthy. (8HoD)

The head of department is in cooperation mode; support from us to her; but she disseminates a good work ethic . . . explains what we have to do. (14T)

She always expects good results from us but we all expect good results from ourselves anyway. (14T)

I get the feeling she trusts in the teachers’ judgement and independence, though we work as a team. Things get said and they get done and integrated in policy. (15T)

Departments were held accountable, however, and operated within a clear framework of policies and expectations. All subject staff followed the same guidelines on reporting, target setting, and other key areas.

This distribution of responsibility was also associated with participatory decision-making. Teachers reported that they felt “part of the process rather than someone who has to do as they are told” (8HoD). For ordinary members of staff, it was never a case of Sara and “the rest of us” because “we are there as a group . . . I never get the feeling that I’m being ordered or told” (13NQT).

All those interviewed confirmed this picture of widely distributed responsibility and their own, personal sense of involvement at the departmental level. When a new scheme of work was introduced, for example, subject staff felt “we were all involved” (14T). One department head was said to be “democratic, organised” so that everyone feels “you can contribute.” Middle managers were perceived to be “open to other people’s ideas,” although “everything is very structured” and “people have clear responsibilities” (13NQT).

Increasing Leadership Capacity

The head's role as a selector and developer of staff was widely perceived to be critical to the school's success. She had a very "conscious and deliberate" approach. She was staff development coordinator in her previous school, and "from the moment I meet people I'm assessing them for development and what they can do in the future" (7H). Teachers valued the opportunities that have unfolded as a result:

I know that I've been encouraged to take on extra roles. SLT are investing in me... I'm getting stuff I could take away to another school. (8HoD)

Valuing people is incredibly important—I've felt valued and I've been given the confidence, the training and the belief to do what's needed. (3AH)

The school has an extensive CPD programme, regarded by one interviewee as "a real strength and an improvement since the last OfSTED... there's lots of in-house training... training on discipline, on what makes a good lesson, inset days used well to target key areas" (5AH). Newly qualified staff reported the "fantastic induction programme" (14T) run by the deputy, praising in particular the open forum where they could be honest about their feelings and problems. Colleagues were pleased that the "teachers lead the inset" because they believed people were "far more likely to take on board the advice and adopt the practice" (5AH) when recommendations come from within. The administrative team was fully included in the programme and participated in regular training and development, particularly to equip individuals for new roles arising from the remodelling agenda.

Senior staff emphasized the importance of frameworks and policies that include monitoring, evaluation, and marking. An assistant head claimed that the target-setting process had been an important vehicle for coaching. A comprehensive mentoring programme had "made it clear to all explicitly what we are here for; moving forward and working hard" (4AH).

School Characteristics

The leadership behaviour described here is similar to that reported at other successful schools and is of the type government agencies believe produces the climate and conditions associated with outstanding performance. As the inspectors reported:

This is a very good and extremely effective school with many outstanding features. The headteacher and senior management team provide clear vision and strong leadership. These have a major impact on improving standards of attainment for pupils of all abilities. Improvement since the last inspection is very impressive. (OfSTED, 2000, p. 8)

The interview data are consistent across the whole sample in emphasizing that: a clear, shared vision has been established; a multiskilled head has created a greater interconnectedness; leadership has been widely distributed; decision-making empowers staff at all levels; opportunities are engineered to give people early responsibility;

ambitious goals are set and individuals receive intellectual stimulation and individualised support; there is an institutional buzz with a strong emphasis on individual needs; and coaching and mentoring activities make use of performance data. These findings are consistent with expectations derived from the literature (Gold et al., 2003; Judkins & Rudd, 2005; Leithwood & Riehl, 2003).

By these means, the head was believed to have impacted directly on factors that have constrained the school in the past. She was regarded as an “outstandingly fantastic PR person” (1AD), who had drawn in “brighter students from out of catchment” so that there was a “brighter end producing good academic results” (12HoY). The school has become “bigger, more popular, the results are very good; surrounding villages hear about it and that’s when you pull in the more middle class people” (16T). These exceptionally positive remarks, echoed to a greater or lesser extent by all those interviewed, are consistent with what has happened to the school over the last 10 years. Since 1994, the number on roll has increased from 594 to 925, the sixth form has grown from 71 to 164, the number entered for A/AS examinations has risen from 26 to 62, and the GCSE cohort has climbed from 94 to 133 (see Table 2). Since 1995, the average verbal reasoning quotient of the students has risen from 92 to over 100 (Thomson, 2006). An enhanced reputation has attracted more capable staff (11HoY), as well as an improved intake. A virtuous circle of success has developed (Thomson, 2006).

The reported quality of relationships between students and teachers shows that this type of leadership can also influence behaviour and expectations indirectly and so create a positive climate for learning. According to one teacher, “the ethos is one of calm friendliness” and “pastoral care is really superb.” Although the school had grown, and there was a remorseless drive for success, there was a “sense of closeness between teachers and students” that you might expect to be lost in a large school. Relationships were “fantastic in all sorts of ways at all sorts of levels” (17T). The aim

Table 2. The Shire School: performance data 1994–2005

| Date | NOR | 16+ | 15+ | Percent obtaining 5 A*–C | | | Percent A–G | Pts | A/AS entry |
|------|-----|-----|-----|--------------------------|---------|--------------|-------------|-----|------------|
| | | | | LEA | England | Shire School | | | |
| 2005 | 925 | 164 | 133 | 62 | 56 | 65 | 100 | 306 | 62 |
| 2004 | 885 | 140 | 139 | 61 | 54 | 65 | 95 | 292 | 57 |
| 2003 | 840 | 122 | 126 | 61 | 53 | 71 | 94 | 251 | 48 |
| 2002 | 778 | 98 | 128 | 61 | 52 | 69 | 96 | 252 | 44 |
| 2001 | 733 | 85 | 114 | 58 | 50 | 62 | 97 | 24 | 28 |
| 2000 | 690 | 88 | 98 | 57 | 49 | 56 | 100 | 21 | 47 |
| 1999 | 662 | 89 | 96 | 56 | 48 | 47 | 95 | 24 | 26 |
| 1998 | 644 | 85 | 93 | 55 | 46 | 49 | 84 | 22 | 46 |
| 1997 | 621 | 81 | 106 | 52 | 45 | 43 | 96 | 18 | 25 |
| 1996 | 614 | 78 | 112 | 50 | 44 | 46 | 85 | 20 | 40 |
| 1995 | 615 | 75 | 91 | 49 | 43 | 38 | 81 | 15 | 29 |
| 1994 | 594 | 71 | 94 | 46 | 43 | 45 | 84 | 11 | 26 |

was to “deliver for the individual rather than the average child . . . the school is there for their individual needs” (5AH). Data were “only useful in terms of helping individuals” and the “right students in the right subjects” were targeted (6DH).

Although the interview data here are consistent in supporting this account of the school, there are grounds for caution. Burns and Stalker (1995, p. 211) commented on the “extraordinary importance ascribed to the personal qualities” of managing directors and the tendency for leaders to symbolize or personify their concerns in the eyes of their subordinates. Malcolm (2004) was inclined almost to discount the interview as a source of information on the grounds that subjects can be too eager to please their questioners and to represent their work in positive terms. Even so, other sources of evidence, including the OfSTED Report (OfSTED, 2000), confirmed the very positive interpretation provided by the teachers interviewed.

Learning and Teaching

Inspection and observation reports provided further evidence that the school’s leadership had induced conditions that were positive for colleagues seeking to enhance the quality of learning and teaching. In December 2000, inspectors found that teachers had very high expectations for pupils of all abilities and achieved high standards of academic performance through very good teaching (OfSTED, 2000). The head of English, like other middle managers, was said to provide a very good example that was having an impact on standards. Skilful questioning encouraged pupils to extend and develop their answers (OfSTED, 2000).

Interviews with teachers suggested that the leadership team continues to promote improvements that impact on students’ learning. The “infrastructure of the school has hugely improved,” so that “every teacher has a designated classroom.” There were “no itinerant teachers who are all over the shop” (4AH). According to a head of department, a “fantastic building” had made a “big difference to the feel of the place” (8HoD). Electronic white boards, wireless laptops, and new software were said to have “radically changed the way we teach” (4AH) through their interactive, creative potential (17T). One head of department explained that:

I can show images . . . you can do so much with video clips . . . you can get on the Amnesty website, show prison, show a letter to write . . . my NQT has been doing brilliant work with it, designing starters and plenaries using wireless. (8HoD)

According to another teacher, “everyone at the school is committed—people are always coming up with ideas, people have a commitment, there’s a steady strand of ideas to how we’re to achieve the vision” (13NQT). Although OfSTED identified no issues for action, the school was working on a full timetable review because “unless we ask the questions we don’t know if there is a solution” (5AH).

Observation of lessons and short conversations with staff and students as they worked confirmed the links between the climate encouraged by the school’s leaders and the quality of life in the classroom and also indicated that little had changed since the OfSTED Report (OfSTED, 2000). Students were punctual, attentive, and well

behaved. With only one minor exception, classes waited quietly for teachers to begin lessons and responded eagerly to questions and tasks. Classroom exchanges created the atmosphere of a good-natured conversation with a purpose. Lessons were thoroughly well planned and documented, with a consistent emphasis on learning objectives and frequent reference to schemes of work. The goals of each activity were communicated clearly, using a variety of media (black boards, white boards, PowerPoint slides, and wall charts). The skills and knowledge to be acquired were explained in relation to previous learning and future needs. Practical activities and group tasks were managed so that changes of pace and direction engaged and motivated students of all ages and abilities. All the lessons seen combined accessibility with challenge (author's notebook, 6 October 2005, various entries).

One girl commented that she liked "all the teachers, they're all cool, only one or two grumpy ones," while a boy confided that "Shire kids are nice kids, there's no bullying." Another student felt that the teachers "are so good, I feel sorry for my friends who tell me about other schools . . . our maths is at a level to challenge us but we can still do it" (author's notebook entry, 6 October 2005).

An English lesson showed how the school climate encouraged inventive teaching. Students from Year 11, equipped with prepared poems and carefully planned learning activities, worked in groups with Year-7 children. The two sets of young people were given name badges, were introduced to one another, and quickly generated a steady buzz of explanation and discussion about a wide variety of poems. The teacher moved between tables with a clipboard, recording oral contributions but also pausing to explain the difference between a simile and a metaphor. She commented that the format had created "so many opportunities to score points . . . their oral grades have rocketed up" (author's notebook entry, 6 October 2005).

Although this observational evidence is impressionistic and drawn from a very small sample, the examples cited are consistent with the OfSTED Report (OfSTED, 2000, p. 20), when teaching and learning were found to be "good or better in 80% of the 172 lessons seen . . . no lessons were less than satisfactory," while behaviour was described as "outstanding" (p. 10) around the school and in class.

Measuring Effectiveness

This case study provides comprehensive evidence, therefore, of the type of leadership recommended by government agencies. Sara Thomson and her team were reported to have displayed, at a consistently high level, all the qualities described in the National Standards (DfES, 2004a). An important example of her strong commitment to government priorities was the emphasis on ensuring "a consistent and continuous school-wide focus on pupils' achievement, using data and benchmarks to monitor progress in every child's learning" (DfES, 2004a, p. 7).

The head and her colleagues were said to be passionate about raising achievement and to concentrate on factors that are believed to improve the quality of learning and teaching. OfSTED (2000), an agency seldom prone to enthusiasm, confirmed eyewitness reports that leadership at the school was particularly effective in raising

standards. With so much government emphasis on transformation and performance, there should be substantial evidence of improved results at a school that seemed to exemplify recommended leadership practice.

Improved effectiveness in producing results is less easily measured, however, than the simple world of the performance tables suggests. This is because the chosen measures (tests and examinations) are contested (Apple, 1989), unreliable (Rowe, 2000), and subject to frequent change (Goldstein, 2001; West & Pennell, 2000). Untrustworthy methods are used to evaluate a “nation on the move” (see the discussion in Gray et al., 1999, pp. 9–20), so that unstable criteria are applied to a moving target. As the percentage of candidates nationally achieving five or more A*–C grades has risen by 30% since 1994, “it is not possible to decide whether any change in test score is really due to a change in performance or a change in the difficulty of the test, or a mixture of the two” (Goldstein, 2001, pp. 434–435).

Tests and examinations that are not easily compared with one another (Goldstein, 2001; Gorard, Rees, & Selwyn, 2002; Hilton, 2001), inappropriate methods of analysis (Woodhouse & Goldstein, 1988), and misleading value-added indicators (Rowe, 2000) are used to quantify differences between schools. The DfES value-added tables, for example, disregard social background and are based on “over-simplifications of a more complex underlying structure” and “lead to biased estimates” when different types of school are compared (Goldstein, 2004). Comparisons that use average, unadjusted examination results are invalid (Goldstein & Thomas, 1996). Goldstein (2001, p. 435) concluded that the data and advice contained in the DfES “Autumn Package” are “at best confusing and at worst misleading.”

Despite these well-known doubts about the methods and data used to evaluate school performance, current measures do provide clear evidence that a steady improvement in effectiveness has taken place at The Shire during Sara Thomson’s headship. As Chart 1 shows, between 1994 and 1999 the percentage of students achieving five or more GCSE A*–C grades fluctuated around the national average and trailed the LEA average. Between 2000 and 2003, however, the GCSE headline figure climbed rapidly, to a point significantly above the national and LEA averages.

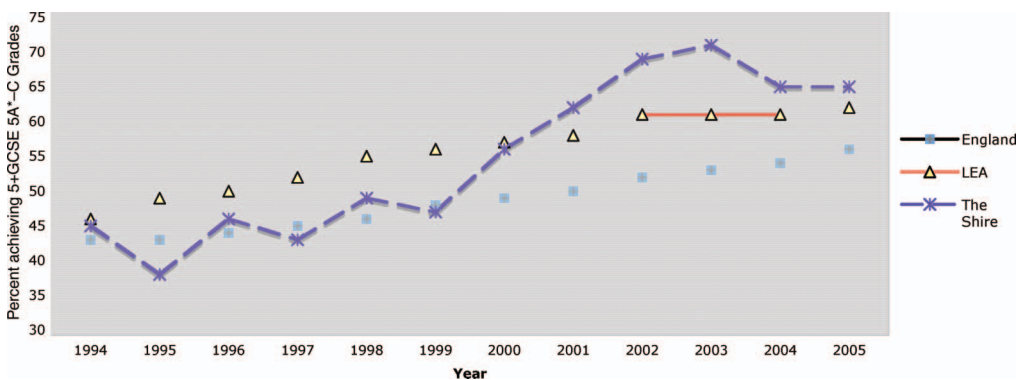


Chart 1. The Shire School of Local & National GCSE % 5A*–C.

The percentage of students obtaining five or more GCSE higher grades in 2004 and 2005, however, was less than the peak reached in 2003. As a result, the mean percentage of students attaining five or more higher grades through the whole period from 1994 to 2005 (55%) is similar to the LEA mean. On the new higher grade measure that includes English and mathematics (introduced in 2005), the school scored 54%, compared with the LEA average of 52.6%. The unadjusted five or more GCSE A*–C score has improved from 1% below the LEA average in 1994 to 3% above the LEA average in 2005 (Table 2).

A similar improvement has occurred at A-level, with the points score rising from 252 to 306 between 2002 and 2005. There is also evidence that the school's current effectiveness is well above the national average but only slightly above average for the LEA:

In 2004, despite an average verbal reasoning quotient of just 96.4 . . . 65 per cent of Year 11 pupils achieved 5+ A*–C, 94.2 per cent A*–G including maths and English and a capped points score of 40.7—all equal to or well above comparator group and county averages. (The Shire School, 2005, p. 10)

When the GCSE total points are compared with the corresponding A-level points, “the general trend is above or around the average expectation” (The Shire School, 2005, p. 12). The current A-level points score (306) is only slightly above the LEA average of 298.7 (BBC, 2005).

Value-added indicators also rank the school in an above-average position. The value added between Key Stages (KS) 2 and 4 in 2005 was 1009.9, compared with an LEA average of 999.5. Twenty-one schools in the LEA were shown as adding more value on this measure, while 25 added less. The value-added between KS 3 and 4 was 1005.4 (BBC, 2005).

Explaining Improved Effectiveness

As our knowledge of an individual school's effectiveness depends on untrustworthy measures and procedures, student performance is seen “through a glass, darkly” (1 Corinthians 13:12). Annual variations in test and examination results should be interpreted cautiously. The size and origin of the intake at 11+ and 16+, for example, has altered so much that simple comparisons between past and present performance are of doubtful validity. The total number of students entered for A-level (62 in 2005) is too small for fully satisfactory conclusions to be drawn.

A number of internal and external changes other than leadership seem to have contributed to the improved examination results and qualify our perception of the school's enhanced effectiveness. Additional resources and enhanced status, achieved through designation first as a foundation school, then as a Language College, seem to have helped The Shire exploit its accessibility for a widely dispersed but reasonably prosperous and mobile population. The school's relatively improved results may be plausibly associated with the better intake reported by those interviewed (11HoY, 16T, and 1AD) and the higher verbal reasoning average (up from 92 to 100) logged

by the head (Thomson, 2006). This is consistent with the emergence of a significantly below-average FSM eligibility by the time of the OfSTED inspection in 2000 (11%, compared with the national figure of 16.5%). Specialist Language Colleges (mean FSM eligibility of 10.2%), particularly those with foundation status, seem in general to have less disadvantaged intakes than other nonselective schools (Gorard & Taylor, 2001, p. 375). An admissions officer in another LEA was prompted to comment:

because if you're doing languages you're going to be bright and if you're bright it's going to be a good school, and if it's a good school you're going to go there. (quoted in Gorard & Taylor, 2001, p. 378)

Evidence of “superior organizational performance” (Mabey & Ramirez, 2004) is limited, therefore, because changes in the family background and ability of the students may explain a significant part of the school's improved effectiveness. In this light, the claim that there has been “a major impact on improving standards of attainment” (OfSTED, 2000, p. 8) is less than convincing.

Discussion and Conclusion

Once allowance is made for background variables, the available data seem to corroborate studies that have found a small, indirect impact on organizational effectiveness and outcomes (Bell et al., 2003; Hallinger & Heck, 1998). A significantly better intake seems, for example, to explain a good part of the increase of 4% (relative to the local authority mean) in the proportion of students achieving five or more GCSE higher grades between 1994 and 2005. Leadership-related changes (commitment to vision, ambitious goals, interconnectedness, dispersal of leadership responsibility, and a collaborative culture), however exceptional in this case, seem to have produced a relatively small gain in organizational effectiveness, as measured by academic outcomes. This finding that even a head like Sarah Thomson has succeeded in inducing no more than a limited improvement in measured effectiveness is consistent with previous studies and suggests that the government's transformational expectations are remote from what most schools can hope to achieve. The gap between policy goals and the evidence that they are realistic is wide indeed.

On many levels, however, The Shire School has been transformed. The number on roll has risen by 55%, the sixth form has increased by 130%, the percentage gaining five or more GCSE A*–C is 71% higher than on the head's appointment, and the premises and facilities are enormously improved. Language College status has helped the school to become an important centre for the local community. Interview respondents, supported by OfSTED, would agree that the head is a “level 5” leader who “builds greatness through a paradoxical blend of personal humility and professional will” (Collins, 2000, p. 20). She had changed the context (Fullan, 2003) so that better teachers, an improved student intake, and enhanced resources have combined to produce an “outstanding” school (OfSTED, 2000).

The choice of examination results as the main indicator of success helps explain the disconcerting paradox presented by outstanding leaders, who seem to have only a small impact on relative effectiveness. This measure is a severe test for any of the interacting school variables that may contribute to improvement, including leadership, because results are strongly conditioned by student background (Scheerens, 1989). When only a small margin is available to be influenced by a wide range of factors, it is not surprising that leadership effects are often hard to trace (Bush, 2004).

An unrelenting preoccupation with internal conditions that raise achievement also leads us to value only those leadership qualities that register on the convergent radar screen of the performance tables. The selected measures reward techniques that may sacrifice deeper learning for the sake of short-term gains. Preparation for reading tests, for example, has been reported to undermine “children’s ability to read sustained passages” (McNeil, 2000, p. 238), while writing can become no more than “daily practice in ‘the persuasive essay’” (McNeil, 2000, p. 239). When instructional leaders do encourage breadth and depth of learning, the expected improvement may not appear in the performance data (Frost & Durrant, 2002; Van Houtte, 2005). Promising initiatives may work in one area but fail in another (Muijs, Campbell, Kyriakidis, & Robinson, 2005).

The search for a “golden rule-book or recipe for effective leadership” (MacBeath & Myers, 1999, p. 67) that transforms examination results seems also to have produced a distorted picture of how leaders contribute to improvement. In this case, the head’s work to reposition The Shire within the local hierarchy of schools seems to have been an important factor in transforming its relative size and success and confirms a growing literature that associates student outcomes with intake variables (Levačić & Woods, 2002a, 2002b; Lupton, 2004). A more favourable student mix seems to have enhanced the school’s relative effectiveness (Thrupp, 1999). Such an improvement is likely to be missed by studies designed to confirm a positive relationship between particular types of leadership, internal conditions, and student outcomes.

Once allowance has been made for intake, leadership appears to become a marginal factor. The consistent finding that leadership has only a small impact on results challenges our common sense assumption that leaders are important and impedes our understanding of their role and influence. The government’s desire for results that exceed expectations has diverted research attention from the more usual scenario, where leaders and teachers contribute valuably to student outcomes without raising the grades above predictions based on prior attainment.

This study does help explain, therefore, why leadership effects are often found to be less than policy-makers expect, although the scope for generalizing from a single, exceptional school is obviously limited and there is a clear need for more research from a critical perspective. The government advocates leadership models that are essentially incomplete and unproven (Bush, 2004; Mulford et al., 2004) but demands, nevertheless, a transformation in schools and results that is hard to achieve with an effectiveness framework that assigns disproportionate value to narrowly defined and “difficult to influence” outcomes (Scheerens, 1989, p. 71). These policy decisions have created the conditions for an apparent leadership paradox,

like that at The Shire, where there is a marked disparity between the reported, transformational qualities of the leader and the less than remarkable improvement in examination results.

This paradox is symptomatic of an uncomfortable fissure between the drive to transform public service organizations and a parallel commitment to “evidence-informed policy development” (Levitt & Solesbury, 2005, p. 3). The government’s choice of test and examination scores as the only valid measure of student outcomes, and a research emphasis on large-scale quantitative studies to confirm that leadership improves results, have conditioned the available evidence of improvement to the point where it seems as if the transformational agenda may be unjustified. This conclusion is at odds, however, with the prevalent assumption that leadership is a critical organizational variable and with positive indications from a large number of qualitative studies, including this one (Barker, 2005; Judkins & Rudd, 2005). The government’s determination to drive forward with a strategy based on an assumed relationship between leaders and outcomes has compromised the principle of evidence-informed policy-making and has undermined the prospect of successful reform.

A different approach is required if the potential role of leadership in bringing about improvement is to be fully understood and realized. Policy-makers should begin to “distinguish between measured performance in tests and other kinds of knowledge, understanding, skills and development” (Frost & Durrant, 2002, p. 145). Leithwood and Levin (2005, p. 4) believed we should also “measure a more comprehensive set of leadership practices than has been included in most research to date” if we are to understand the varieties of school leadership that contribute most effectively to student progress. We need a “collaborative research programme” (Bush, 2004) to explore how leaders contribute to a more broadly defined, qualitative conception of student development and outcomes.

Acknowledgement

I should like to thank Professor Clive Dimmock for his unfailing support and guidance in the shaping of this article.

References

- Apple, M. (1989). Critical introduction: ideology and the state in educational policy. In R. Dale (Ed.), *The state and education policy*. Milton Keynes, UK: Open University Press.
- Barber, M. (2000). *High expectations and standards for all, no matter what: Creating a world class education service in England*. Keynote speech given at the Smith Richardson Foundation, Washington, DC, 31 May 2000. Retrieved February 4, 2002, from http://www.ncsl.org.uk/index.cfm?pageid=ev_auth_barber
- Barker, B. (2001). Do leaders matter? *Educational Review*, 53(1), 65–76.
- Barker, B. (2005). *Transforming schools: Illusion or reality*. Stoke, UK: Trentham Books.
- BBC. (2005). *Performance tables*. Retrieved March 8, 2006, from http://www.news.bbc.co.uk/1/shared/bsp/hi/education/04/school_tables/secondary_schools/html/916_va_lea.stm

- Bell, L., Bolam, R., & Cubillo, L. (2003). *A systematic review of the impact of school headteachers and principals on student outcomes*. London: University of London Institute of Education.
- Blair, T. (2001, February 14). Our schools can only get better if they are distinct. *The Times*, p. 22.
- Blunkett, D. (2000). *National college for school leadership* (Letter to Richard Greenhalgh, with list of tasks and responsibilities). London: DfEE.
- Bolam, R. (2004). Reflections on the NCSL from a historical perspective. *Educational Management Administration & Leadership*, 32(3), 251–267.
- Burns, T., & Stalker, G. (1995). *The management of innovation*. Oxford, UK: Oxford University Press.
- Bush, T. (2004). *The performance of leaders: Aims, impact and development*. Keynote paper presented at the British Educational Leadership, Management and Administration Society's Annual Conference, Stone, UK, October. Retrieved January 23, 2006, from <http://www.shu.ac.uk/bemas/bush2004.html>
- Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, DC: US Government Printing Office.
- Collarbone, P. (2001). *Leadership programme for serving headteachers: A review*. Nottingham: NCSL.
- Collins, J. (2000). *Good to great*. New York: HarperCollins.
- Day, C., Harris, A., Hadfield, M., Tolley, H., & Beresford, J. (2000). *Leading schools in times of change*. Buckingham, UK: Open University Press.
- Department for Education and Employment (DfEE). (1997a). *Excellence in schools*. London: The Stationery Office.
- Department for Education and Employment (DfEE). (1997b). *Setting targets for pupil achievement*. London: DfEE.
- Department for Education and Skills (DfES). (2002). *Government response to the report of the Council for Excellence in Management and Leadership*. Nottingham, UK: DfES Publications.
- Department for Education and Skills (DfES). (2004a). *National standards for headteachers*. London: DfES. Retrieved April 15, 2006, from <http://publications.teachernet.gov.uk/eOrderingDownload/NS4HFfinalpdf.pdf>
- Department for Education and Skills (DfES). (2004b). *School leadership: End to end review of school leadership policy and delivery*. London: DfES.
- Department for Education and Skills (DfES). (2005a). *Secondary school performance tables*. Retrieved March 16, 2006, from http://www.dfes.gov.uk/performance/tables/schools_05.shtml
- Department for Education and Skills (DfES). (2005b). *Trends in education and skills*. Retrieved April 17, 2006, from <http://www.dfes.gov.uk/trends/index.cfm?fuseaction=home.showIndicator&cid=5&iid=32>
- Department for Education and Skills (DfES). (2006). *Programme of research, 2003/080 Research into the impact of school leadership on pupil outcomes, Tendered, Expected start date 2.01.06*. Retrieved January 28, 2006, from <http://www.dfes.gov.uk/research/programmeofresearch/index.cfm?type=>
- Dunford, J. (2002). Transforming secondary schools? *Headlines*, 40, 6–8.
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37, 15–27.
- Frost, D., & Durrant, J. (2002). Teachers as leaders: exploring the impact of teacher-led development work. *School Leadership & Management*, 22(2), 143–161.
- Fullan, M. (2003). *The moral imperative of school leadership*. Thousand Oaks, CA: Corwin Press.
- Gall, M., Borg, W., & Gall, J. (1996). *Educational research: An introduction* (6th ed.). New York: Longman.
- Galton, M., Hargreaves, L., Comber, C., Wall, D., & Pell, A. (1999). *Inside the primary classroom: 20 years on*. London: Routledge.
- Gold, A., Evans, J., Earley, P., Halpin, D., & Collarbone, P. (2003). Principled principals? Values-driven leadership: Evidence from ten case studies of “outstanding” school leaders. *Educational Management & Administration*, 31(2), 127–137.
- Goldstein, H. (2001). Using pupil performance data for judging schools and teachers: Scope and limitations. *British Educational Research Journal*, 27(4), 433–443.

- Goldstein, H. (2004). *A commentary on the 2004 league tables for England*. Retrieved March 23, 2006, from http://www.mlwin.com/hgpersonal/league_tables_England_2004_commentary.htm.
- Goldstein, H., & Thomas, S. (1996). Using examination results as indicators of school and college performance. *Journal of the Royal Statistical Society*, 159(1), 149–163.
- Gorard, S. (2006). Is there a school mix effect? *Educational Review*, 58(1), 87–94.
- Gorard, S., Rees, G., & Selwyn, N. (2002). The “conveyor belt effect”: A reassessment of the impact of national targets for lifelong learning. *Oxford Review of Education*, 28(1), 75–89.
- Gorard, S., Selwyn, N., & Rees, G. (2002). “Privileging the visible”: A critique of the national learning targets. *British Educational Research Journal*, 28(3), 309–325.
- Gorard, S., & Taylor, C. (2001). The composition of specialist schools in England: track record and future prospect. *School Leadership & Management*, 21(4), 365–381.
- Gray, J., Hopkins, D., Reynolds, D., Wilcox, B., Farrell, S., & Jesson, D. (1999). *Improving schools: Performance and potential*. Buckingham, UK: Open University Press.
- Gray, J., Jesson, D., & Sime, N. (1995). Estimating differences in examination performances of secondary schools in six LEAs: A multi-level approach to school effectiveness. In J. Gray & B. Wilcox (Eds.), *Good school, bad school: Evaluating performance and encouraging improvement*. Buckingham, UK: Open University Press.
- Hallinger, P., & Heck, R. (1998). Exploring the principal’s contribution to school effectiveness: 1980–1995. *School Effectiveness and School Improvement*, 9(2), 157–191.
- Hammersley, M. (1992). *What’s wrong with ethnography?* London: Routledge.
- Harris, A. (2004). Distributed leadership and school improvement—leading or misleading. *Educational Management Administration & Leadership*, 32(1), 11–22.
- Hilton, M. (2001). Are the Key Stage Two reading tests becoming easier each year? *Reading*, 35(1), 4–11.
- Jencks, C., Smith, M., Ackland, H., Bane, M., Cohen, D., Grintis, H., Heynes, B., & Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Jesson, D. (2001). *Educational outcomes and value-added analysis of specialist schools for the year 2000*. London: Technology Colleges Trust.
- Judkins, M., & Rudd, P. (2005). *Evaluation of high-performing specialist schools*. Paper presented at the British Educational Research Association Annual Conference, University of Glamorgan, Pontypridd, 15–17 September. Slough, UK: NFER.
- Kelly, R. (2005). *Interview*, 23 February, BBC Radio 4. Reported in *Education Guardian*, February 23. Retrieved April 15, 2006, from <http://www.education.guardian.co.uk/1419education/story/0,15147,1423404,00.html>
- Kirkup, C., Sizmur, J., Sturman, L., & Lewis, K. (2005). *Schools’ use of data in teaching and learning* (DfES Research Report 671). London: DfES.
- Lapadat, J. (2000). Problematizing transcription: Purpose, paradigm and quality. *International Journal of Social Research Methodology*, 3(3), 203–219.
- Leithwood, K., & Jantzi, D. (2000). Principal and teacher leadership effects: A replication. *School Leadership and Management*, 20(4), 415–434.
- Leithwood, K., & Levin, B. (2005). *Assessing school leader and leadership programme effects on pupil learning* (DfES Research Report 662). London: DfES.
- Leithwood, K., & Riehl, C. (2003). *What we know about successful school leadership*. Philadelphia, PA: Temple University. Retrieved January 23, 2006, from <http://www.cepa.gse.rutgers.edu/whatweknow.pdf>
- Levačić, R., & Woods, P. (2002a). Raising school performance in the league tables (part 1): Disentangling the effects of social disadvantage. *British Educational Research Journal*, 28(2), 207–226.
- Levačić, R., & Woods, P. (2002b). Raising school performance in the league tables (part 2): Barriers to responsiveness in three disadvantaged schools. *British Educational Research Journal*, 28(2), 227–247.

- Levitt, R., & Solesbury, W. (2005). *Evidence-informed policy: what difference do outsiders in Whitehall make?* (Working Paper 23). London: ESRC UK Centre for Evidence Based Policy and Practice. Retrieved April 17, 2006, from <http://www.evidencenetwork.org/Documents/wp23.pdf>
- Lupton, R. (2004). Understanding local contexts for schooling and their implications for school processes and quality. *BERA Research Intelligence*, November, 89.
- Mabey, C., & Ramirez, M. (2004). *Developing managers: A European perspective*. London: Chartered Management Institute. Retrieved April 14, 2006, from www.managers.org.uk/euromanagementdevelopment.
- MacBeath, J., & Myers, K. (1999). *Effective school leaders*. Harlow, UK: Pearson Education.
- Madison, L., & Allison, C. (2004). *Learning-centred leadership project: A study of middle leadership in 21 secondary schools*. Nottingham: NCSL. Retrieved January 28, 2006, from www.standards.dfes.gov.uk/ntrp/lib/pdf/Patrick.pdf.
- Malcolm, J. (2004). *The journalist and the murderer*. London: Granta.
- McNeil, L. (2000). *Contradictions of school reform*. New York: Routledge.
- Miller, R., & Rowan, B. (2006). Effects of organic management on student achievement. *American Educational Research Journal*, 43(2), 219–253.
- Morris, E. (2002). *The London challenge*. Speech by Estelle Morris, Secretary of State for Education. London: DfES.
- Mortimore, P., & Whitty, G. (2000). *Can school improvement overcome the effects of disadvantage?* London: University of London Institute of Education.
- Muijs, D., Campbell, J., Kyriakidis, L., & Robinson, W. (2005). Making the case for differentiated teacher effectiveness: An overview of research in four key areas. *School Effectiveness and School Improvement*, 16(1), 51–70.
- Mulford, W., Silins, H., & Leithwood, K. (2004). *Educational leadership for organisational learning and improved student outcomes*. Dordrecht, The Netherlands: Kluwer Academic.
- National College for School Leadership (NCSL). (2001). *Leadership development framework*. Nottingham: NCSL.
- National College for School Leadership (NCSL). (2003). *Leadership programme for serving headteachers: facilitator guide*. Nottingham: NCSL.
- Office for Standards in Education (OfSTED). (1993). *Handbook for the inspection of schools* (May 1994 Amendment). London: HMSO.
- Office for Standards in Education (OfSTED). (2000). *Inspection report, the Shire School*. London: HMSO.
- Office for Standards in Education (OfSTED). (2006). *Improved data usage and analysis: Merger of PANDA and PAT*. London: HMSO. Retrieved April 15, 2006, from <http://www.ofsted.gov.uk/schools/dataandinformationsystems.cfm>.
- Park, R. (2000). *Voodoo science: The road from foolishness to fraud*. Oxford, UK: Oxford University Press.
- Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. New York: Basic Books.
- Punch, K. (2005). *Introduction to social research: Quantitative and qualitative approaches*. London: Sage.
- Reynolds, D., Bollen, R., Creemers, B., Hopkins, D., Stoll, L., & Lagerweij, N. (1996). *Making good schools: Linking school effectiveness and school improvement*. London: Routledge.
- Rowe, K. (2000). Assessment, league tables and school effectiveness: Consider the issues and “let’s get real!”. *Journal of Educational Enquiry*, 1(1), 73–98.
- Rutter, M., Maugham, B., Mortimore, P., & Ouston, J. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. London: Open Books.
- Scheerens, J. (1989). *Effective schooling: Research, theory and practice*. London: Cassell.
- Stoll, L., & Fink, D. (1995). *Changing our schools: Linking school effectiveness and school improvement*. Buckingham, UK: Open University Press.

- The Shire School. (2005). *Self-evaluation form*. The Shire School.
- Thomson, S. (pseudonym) (2006, February). *Improving the Shire school*. Paper presented at a research seminar held at the Centre for Educational Leadership and Management, School of Education, University of Leicester, UK. Copy of anonymised paper is available from the author.
- Thrupp, M. (1999). *Schools making a difference: Let's be realistic*. Buckingham, UK: Open University Press.
- Thrupp, M. (2001). Recent school effectiveness counter-critiques: Problems and possibilities. *British Educational Research Journal*, 27(4), 443–457.
- Van Houtte, M. (2005). Climate or culture? A plea for conceptual clarity in school effectiveness research. *School Effectiveness and School Improvement*, 16(1), 71–89.
- West, A., & Pennell, H. (2000). Publishing school examination results in England: Incentives and consequences. *Educational Studies*, 26(4), 423–436.
- Woodhouse, G., & Goldstein, H. (1988). Educational performance indicators and LEA league tables. *Oxford Review of Education*, 14(3), 301–320.
- Wragg, E. (1999). *An introduction to classroom observation*. London: RoutledgeFalmer.

Copyright of School Effectiveness & School Improvement is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.