Northumbria Research Link

Will the lessons be learned? Reflections on Local Authority evaluations and the use of research evidence

Abstract

Sure Start programmes are complex, community-based initiatives – fore-runners of the Children's Centres Initiative - which have been evaluated nationally and locally. Using an in-depth, retrospective case study of an evaluation of one local programme, the authors raise key issues pertinent to both practice and evaluation in the field, highlighting conflicts and dilemmas both within evaluation generally and, specifically relating to the evaluation of this programme. We illustrate the difficulties placed upon local evaluators by the lack of clear structures within which to work, and provide useful lessons as we move forward into the development and evaluations of new services for children and families.

Keywords: Sure Start; Evaluation; complex community initiatives.

Introduction

University research has many facets: there are high-status projects, funded by research councils and government departments which take almost as much time and energy to bid for as to complete, commissioned work from partner organisations, charities and other funders where shared agendas can be generated or significant compromises brokered and there are smaller projects, which may add to our understanding of the field and our research expertise but which are also pragmatically undertaken to keep researchers in their jobs between the (hopefully) larger projects. The authors of this paper have been working in social science research for a combined thirty years, and on 50+ projects, mostly in educational research focused on issues of inclusion, social justice, family learning, the exploration of professional roles and multi-professional working. We have a profound interest in, and a commitment to, the innovative work carried out in the Local Authorities (LAs) in our region. We place great respect and value on work where professionals from a range of education, health and social care backgrounds constantly break new ground in finding ways to work with young children and their families, to provide services, opportunities and spaces where needs can be more precisely identified and met.

This paper reports on our experience as a Higher Education Institution (HEI), and those of the LA, of one such evaluation study – an evaluation of a local Sure Start programme. Our intention is to make explicit how professionals can operate in this difficult area. We firstly explore the nature of the evidence hoped for through both the local and the national evaluation of Sure Start, followed by a discussion on what evidence was actually practicable to collect. The complexities and tensions that such a

task reveals are highlighted, which will serve to illuminate the potential experiences of others who are embarking on similar evaluative work. Finally, given the complex nature of evaluating community initiatives, we explore the reality of how the information gleaned through the evaluation was used. We present a descriptive model of evaluative research, which highlights areas where formative information-flow between project workers/programme managers and evaluators is problematic. We argue that there are gaps between the expectations of the evaluations and what was feasible, and then in turn between what was feasible and how the information was used.

Sure Start and the role of Local Authorities

Sure Start is a cross-Government funded programme in England which was put in place to join up services for families with young children under 5 years old living in poverty and suffering the consequences of social exclusion. It aims to achieve better outcomes for children, parents and communities by:

- increasing the availability of childcare for all children
- improving health and emotional development for young children
- supporting parents as parents and in their aspirations towards employment. (http://www.surestart.gov.uk/aboutsurestart/).

The first local programmes were set up in 1999, and were area-based initiatives. Latest figures indicate an investment of £760 million between 1999-2004 (Moss, 2004), with further funding of £4,317 million projected 2005-8 (Hodge, 2004). Many of the original local programmes now form the basis of the Sure Start Children's Centres, for which the Government is spending £2.2 billion on revenue and over £1 billion on capital between 2004 and 2008 (NAO, 2006). Funding aside, what is key to the national programme is that it is a multidisciplinary intervention programme which is:

- targeted at individual health, learning and social-emotional development and also aims to strengthen families and communities;
- is an agent for change in local service delivery charged with pioneering and then mainstreaming innovative ways of joining up services and working more sensitively and democratically with families; and
- intends to exemplify the relationship between research, evidence and practice (Eisenstadt, 2000; 2002).

The Sure Start, Extended Schools and Childcare Group claims that is currently developing a comprehensive research programme that both adds to and complements existing work for children and families (see: http://www.surestart.gov.uk/research/). The national programme is a long-term, wide ranging study designed to evaluate the efficacy and cost-effectiveness of Sure Start. The first phase of the national evaluation runs

from 2001 to 2008. Sure Start appeared, therefore, to grasp the mettle of evaluation work, and to meet this commitment, and it is this issue which is the main focus of our paper. However, firstly we explore the context of local programmes within Local Authorities.

Local Authorities (LAs) are hemmed in by confusing and at times contradictory messages about how to meet the specific needs of the people in their areas. Stronger and stronger managerial and reporting structures (Sanderson, 2001; Boyne, 2002) and the politics and policies of 'best value' (Geddes & Martin, 2000) make it clear that LAs must identify the most effective and efficient means of providing services – often by following templates offered by 'beacon' (flagship) councils identified and awarded as 'excellent' (Dyson, *et al.,* 2000). In addition, LAs are expected to innovate and respond to the context and needs of their particular area. The pressure to be always innovating and learning – a 'coasting' authority should in fact be 'striving' (Freer, 2002) – will inevitably put pressure on existing resources. There now exists a website dedicated to local Government Performance, where one can compare one local authority with another (http://www.bvpi.gov.uk/pages/Index.asp).

The development of such performance management in the UK within the context of the 'new public management' has been primarily 'top-down' with a dominant concern for enhancing control and 'upwards accountability' rather than directly promoting learning and improvement from the ground level. Additionally, in recent years, LAs have faced a shift in their role as direct service providers, to one where determining policy and strategy is increasingly as important. In the past, however, LAs have tended to rely heavily on short-term project funding, often stitching together funding from a range of initiatives – New Deal for Communities, Education Action Zones, Area Regeneration Grants – in order to support their priorities. This pragmatic use of funding reflects both the ingenuity of local government workers and the extent to which national policy agendas tend to overlap, so that it becomes feasible to represent a single project increasing the communication skills of nursery and reception children in a socially deprived area as fulfilling the objectives of all of the previously cited programmes and thereby ensure the project's longevity. However, in relation to childcare, what is crucial now for LAs is the dramatic shift in their role from implementing these individual programmes or projects, to commissioning and managing the childcare market as a whole. As a result of the Childcare Act 2006, LAs now have new duties to assess the parental demand for childcare; to secure sufficient childcare and to provide an information service for parents of children aged up to 20. Such a shift has particular relevance for Sure Start programmes, and the further development of Children's Centres, as these fit with the pressure from Government for SSLPs to move from *reactive intervention* to *proactive prevention* to meet the needs of the most vulnerable children (DfES, 2004).

What kinds of evidence did we hope for?

Regardless of the good intentions and aims of any initiative or programme, evaluation must be an essential and necessary integrated feature. What is key to the evaluation of Sure Start programmes is the reality that there are two levels: the national evaluation and the local evaluation, commonly referred to as 'dual level' evaluation. Allen and Black (2006) discuss the difficulties and possibilities of the dual level evaluation of Sure Start programmes and describe Sure Start as an example of a Complex Community Initiative (CCI). Sure Start is indeed a perfect example of a large, complex, multi-level, multi-sited and ambitious community initiative programme which inevitably raises issues around evaluation, both nationally and locally. Consequently, evaluations of such community initiatives need to be more sophisticated and there is increasing desire for evaluation approaches that take into account scientific rigour and experimental research while still addressing the complexities of systems change work in community settings (Weiss, 2003). However, there have been criticisms made that review at LA level is more focussed on performance management and monitoring rather than effective evaluation which is integrated into planning and 'best practice' approaches to improving service delivery (Sanderson *et al.*, 2003).

The national evaluation (see: http://www.ness.bbk.ac.uk/) was set up to monitor the implementation and outcomes of the first 260 (of the total 524) local programmes and to collate and analyse the local evaluations that each Sure Start Local Programme (SSLP) had to commission. While the parameters of data collection for the national evaluation were set, the form and content of local evaluations has evolved differently in local programmes (Allen and Black, 2006), providing an opportunity for local programmes to decide on their own evaluative priorities but also presenting some problems of cross-programme comparison. Sure Start guidelines for local programmes state the distinctness between the national and local evaluation:

> "*The national evaluation will answer the questions: 'What difference did Sure Start make to the children in it, and which activities made the most difference'…. Local evaluations on the other hand, will deal with the questions 'what are we doing and how well are we doing it?"* (Sure Start, 2002, Annexe 6, p. 2).

Although evaluation is a part of the management culture and in 'public speech', there appear to be several structural problems which can limit the effectiveness of the evaluation process in improving and spreading innovative practice. These include the quality of information management (Jones and Hughes, 2001); the nature of decision making about innovative practice; the extent to which 'generations of difficulty' and the structural problems in society can be 'solved' at local level (Alcock and Craig, 1998) and the limits to measuring the impact of individual programmes (Gustafsson and Driver, 2005; Boot and Macdonald, 2006).

Experimentation – and Sure Start could certainly be labelled as such - requires the possibility of failure, a failure that can hopefully be analysed to provide better questions and approaches.

Ultimately, Sure Start is about trying new things – some will work, some won't – and we should not berate those whose work has not worked as well as hoped. The benefits of local evaluation should not be undervalued - case studies can illuminate issues that are bound to occur in many other cases and the way in which we have to clarify and reflect upon key issues could be highly relevant to people trying to clarify and reflect upon services in their different contexts. Evaluation should be integral to practice, especially where practice is/should be innovative and we can learn from other local programmes, e.g. Sheffield, where evaluation was part of all Sure Start job specifications and staff were expected to contribute to evaluation work (Weinberger *et al,* 2005).

Evaluation is a required element in most funding packages and guidelines indicate that local Sure Start programmes should allocate between 3% and 5% (capped) of their overall budget to local evaluation (Sure Start, 2002). Whilst local programmes are welcome to co-ordinate and facilitate the evaluation themselves, they can (and usually do) commission external evaluators, such as a HEI or a research consultant. We were commissioned (following competitive tendering) by the LA to conduct a three-year evaluation, costing £39,000 in total, of a local Sure Start programme. Whilst any evaluation has limitations relating to sample representation, incomplete datasets and research design, through this evaluation, we were keen to explore the *processes* of delivery and practice within the area, and provide useful, regular feedback on the projects and the delivery of services within it. We never set out to provide a summative evaluation which focussed purely on numerical and quantitative data; but adopted Cronbach's (1982) view of formative evaluation that, as soon as an innovation is implemented it evolves in relation to the local *and* national context.

The position of the HEI (in our case) was, in theory, to be at the centre of both the national and local programme *and* the evaluation both locally and nationally. We propose that the conceptual model – illustrated in Figure 1 below - demonstrates, the theoretical, or expected, relationship between the HEI and the Sure Start programme.
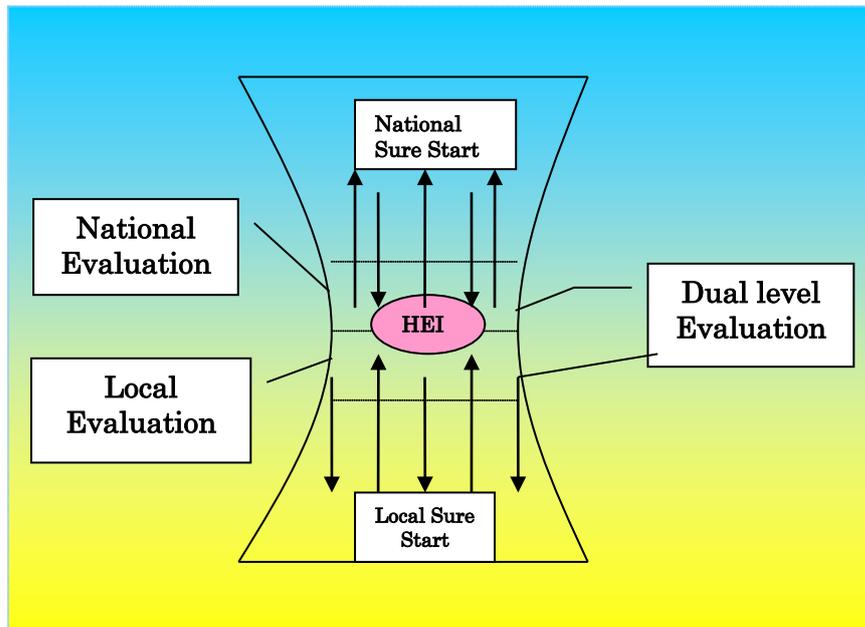
Figure 1: Dual level Evaluation: A theoretical model

We anticipated that our position as researchers would be quite pivotal in the grand scheme of the programme, placed somewhere central between the national and local programme. The dual levels are represented by the two colours (blue and yellow) with a central green area to illustrate some cross-over. Dotted lines represent our perceptions of permeable boundaries between the two levels, and unbroken arrows stand for the anticipated lines of communication and guidance. The two-way directionality of the arrows imply that this guidance would be both 'bottom-up' and 'top-down'.

Our evaluation focus was always planned to be a formative one, which is typically conducted during the development or improvement of a particular programme or project, often with the intention of helping programme staff identify ways in which they can improve delivery. By adopting this approach to the evaluation, we not only acknowledged the developing nature of the Sure Start programme, we also acknowledged the fact that all Sure Start programme staff are – usually informally – constantly doing their own formative evaluation. The reality was, however, that the parameters of the project were set up long ago – and despite our commitment, we would have set it up differently if we had been designing a research project from scratch – and our appearance on the scene was too late to do much more than a post-hoc reflection. Formative though it was in intention, any project where the parameters of success are not mutually agreed before it starts risks being merely summative.

What evidence was feasible?

Although national guidance and training about evaluation assumed that evaluators would be an integral part of the programme, our experiences

were very similar to Allen and Black's who reported they were seen as 'auditors' who were checking up on their work. We were keen to set our formative findings in the context of more quantitative data (e.g. attendance, registration and demographic data). However, the data management systems which would have enabled a monitoring strand within the evaluation had still not been installed by the time of our final meeting in March 2006. Whilst this was disappointing for us as evaluators, it was even worse for programme staff, who were extremely frustrated at being unable to prove that they had satisfied key performance criteria.

The model contained within the national guidance on evaluation assumes that the evaluator will be consistently feeding back findings which are then taken on board by the programmes and fed into project design and delivery, and this would certainly fit with our conceptual model of the theory of evaluating a Sure Start programme. The reality, however, proved to be very different, and a conceptual model of where we felt we actually 'fit' as researchers is illustrated in the second model (Figure 2) below.
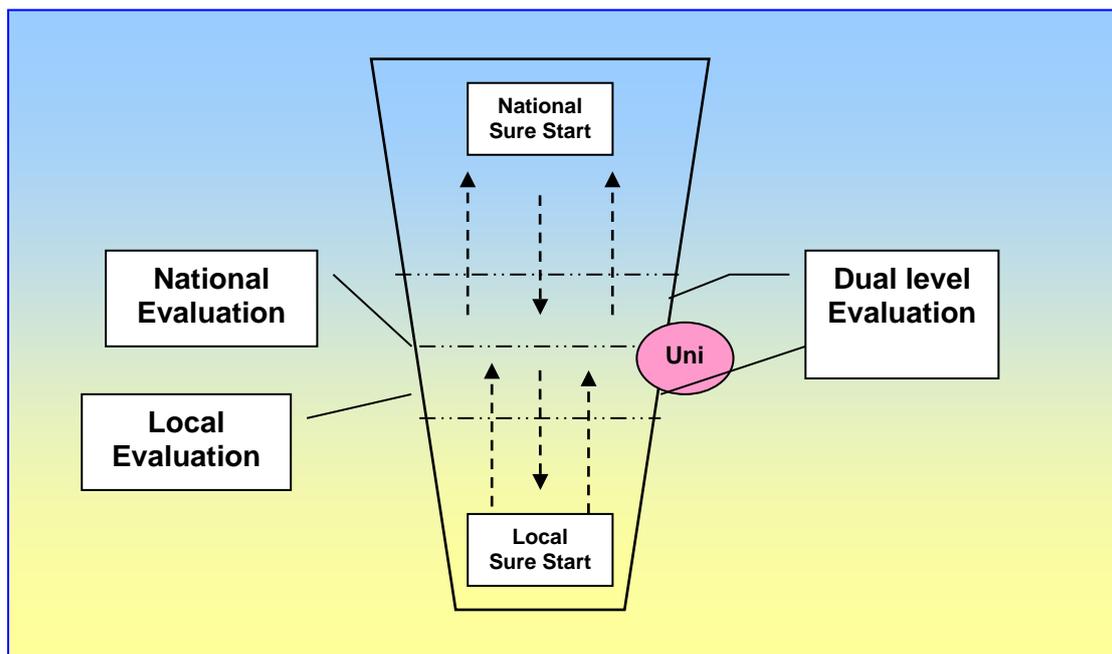
Figure 2: Dual level Evaluation: A realistic model

Considerable delays hampered the progress of the evaluation. Due to start in June 2002, it actually started in February 2003, only to be suspended when the Manager departed from the Programme and then revived in December 2003 when a replacement was appointed. At that time, staffing issues were clearly important for the programme team, and we were informed that the evaluation was not a priority. However, we completed a 'Baseline' survey of local opinion and awareness of the local programme (Hall and Clark, 2004).

In our particular evaluation, there were two levels of reporting and collaboration within the project: one with the local programme team and their Executive Board, which included local councillors and parents from the areas and another which was the strategic management level within the Local Authority. Regular meetings were held with the programme management and the Executive Board at which case study topics were advanced by management and accepted by the research team. Two observational case studies were undertaken of projects within Sure Start focussing on children's creative and language development in 2004-5. In addition, during this time, two 'away-days' were organised and facilitated by the research team, and were held at Newcastle University for the entire local programme team. Part of this process was to allow the research team to explore issues of multi-professional working with the whole group, and to develop new research tools to elicit better understanding of the processes at work. The first session included an exercise to 'map' the multi-professional links within the Sure Start team and beyond, linking with other agencies. The second session included an extended discussion of the role of the multi professional and the competing priorities of community work, in the context externally imposed targets and awareness of the specific needs of the community. Two further case studies, one of a support programme for children with language difficulties and one of working with 'hard to reach' families were planned. Both of these case studies were subsequently cancelled by the programme management and a draft final report was submitted.

At this stage the overt involvement of the strategic managers came to the fore and a period of negotiation over the content and presentation of the final report began, which was finally concluded in March 2006. The main concern for the strategic level managers was the formative nature of the case studies, which raised problems and areas for development as well as highlighting good practice. From their perspective, Sure Start had to be presented as an example of good local management and a springboard to their ambitious development of Children's Centres across the borough and from our perspective, leaving out any comments that were not wholly positive amounted to an unethical use of research evidence. It became explicitly obvious to us, as evaluators, that Government officials be they local or national, want, and indeed *need* to be seen as efficient and effective. In particular, they were unhappy with our report on multi-professional working, specifically the view of the Sure Start professionals that their main barriers were linked to structural and management issues in the Authority. This is a major tension in project evaluation: a main objective of the initiative has not been fulfilled, or has been only partially achieved, because of factors beyond the scope and remit of that project. The evaluators must report this, normally in local evaluations, to the managers who have within their remit the theoretical power to address these factors – though it must be allowed, rarely the actual money, time or other resources to do so.

How was the evidence used?

The complexity of the programmes, their many targets, the difficulties of measurement and information management are all key issues which add to the overall difficulties of evaluating 'success' or impact. Increasing numbers of commentators suggest that it is the very complexity of Sure Start that will doom it to failure (Ormerod, 2005; Factcheck, 2006). Moss (2004:632) questions the contradictory nature of Sure Start as an early years intervention as a whole, and refers in particular to the tensions that then exist between the local and national programmes: *'while there is some latitude for local diversity, local programmes are strongly governed at a distance by "the evaluative state". The scope for significant and transgressive local experimentation is severely constrained'.* Similarly, Willan (2007:23) – after reading emerging reports from the national evaluation team writes:

> *... one gets a feeling of complexity, conflicting expectations, and inherent contradictions between what politicians want (a solution to poor social and educational prospects), what policy-makers want (an umbrella service for 'problem' families/communities), what the centres want (a useful local service centred on the needs of children and their parents) and what families want (a drop-in place to meet, staffed by knowledgeable professionals; an information point; and somewhere for the children to enjoy high-quality play).*

Edgley and Avis (2006:434) further explores this in relation to multi-professional working: *'the combination of Sure Start being both demand-led and a form of co-ordinated service provision also produced paradoxical outcomes and thus collaborative strain for some mainstream professionals.'*

Indeed such tensions were evident in our evaluation. The professionals in the local programme reported that there were tensions between the external expectations of Sure Start (i.e. the national, 'bigger' picture) and their genuine excitement in, and commitment to, the innovative work done in collaboration with one another and with the local community. Other programmes have reported the tension which exists for them of the local agenda versus the national agenda – in particular in relation to (the late) national guidance on local evaluation. The 'changing goalposts' described by Allen and Black (242) on what should or should not be included in local evaluation has caused even more tensions for local programmes and their evaluators, and has resulted in less innovative or action research approaches to be adopted in the evaluations. Equally difficult are the apparent target-setting changes (e.g. smoking cessation) –which are set and changed nationally, which are then left to the local programmes and their management to 're-sell' to their teams (Weinberger *et al,* 2005).

Local programme staff in our study reported feeling confused as to whether we were ultimately working for Sure Start, the Local Authority or

the community served by the programme and this lack of clarity about accountability is reflected in other evaluations of SSLPs (Allen and Black, 2006). For the strategic managers in the LA, the complexity of implementation and the focus on outcomes assessment were of paramount concern, so that the evaluation design – focussed as it was on process – had little to offer them: *"a distraction from the main event"* (Allen and Black, 2006: 242). Even if research and evaluation was viewed as an integral component of service delivery, we feel that ultimately service responsibilities would have taken priority.

Research is starting to emerge which reports the difficulties that many Sure Start programmes face, related to staffing in particular, and this was certainly the case within our evaluation. Staff changes (and resulting shortages) had a significant impact on the delivery of specific activities within the local area. Often with Sure Start programmes, staff have to step out of their 'historical' roles and programmes face losses of key staff. This can be for a variety of reasons - the excellence in training and experience received adds to the CV and staff move on. Most Sure Start staff are on short-term contracts, and so mobility (in both directions) can only be expected. The feeling that Local Programmes have 'poached' the best people, while not entirely justified, is certainly a factor in the difficulties experienced by staff in taking innovative practice back to their colleagues in mainstream (Tunstill, *et al.,* 2005; Hall and Clark, 2005). In these terms, Sure Start could be regarded as either a professional opportunity or a professional *cul de sac*. Edgley and Avis (2006) explored the experiences of professionals working collaboratively within Sure Start areas and reported that this had *not* fostered innovation in their own working practices. Activities and services which already existed routinely in the area were re-presented under the 'umbrella' of Sure Start, but resource constraints prevented innovative working. Key to these professionals was the fact that they were working in deprived areas, not because it was Sure Start *per se*. Conversely, the National Sure Start Evaluation team (http://www.ness.bbk.ac.uk/) have produced a series of synthesis reports on the implementation of the local programmes which conclude that while there has been considerable variation between the circumstances inherited by individual programmes, the key determining factors for successful implementation have been the role of the manager and the ability to build and maintain partnerships with other agencies and to engage with families at the level which was appropriate to their needs (Tunstill, *et al.,* 2005).

Like other Sure Start programmes (see Weinberger *et al,* 2005) the local programme did not have access to basic information and data about families locally, and there were issues about data sharing and confidentiality – all factors which may well have an impact on 'success'. For our partners in the SSLP, these external parameters, delays and imposed agendas were familiar elements from our previous research work, though there was a sense that within Sure Start the levels of surveillance

and the stakes in general were higher. The researchers' role in raising the stakes should not be ignored – we regarded ourselves as partners, as 'critical friends' but at the beginning of any evaluation we must be more aware that our presence is threatening, alien, accepted as a necessary evil but fundamentally unwanted. This 'high stakes' atmosphere worked against iterative development of projects and working relationships and overall, our mutual feeling was that while some changes in practice had taken place, the collaborative learning was not sustained over a long enough period to really embed in practice.

With the general tensions and contradictions related to such a huge, complex programme as Sure Start, it is no wonder that evaluation – at both national and local level – is a massive undertaking. On a macro level, Sure Start is about targeting individuals at risk of social exclusion - a task which is difficult enough: '*categorising, identifying, reaching and engaging with the groups poses major problems, not least of definition*' (Willan, 2006:28). On the micro level, Sure Start local programmes are individual complex community initiatives which are geographically bounded and which offer a specific service delivered within targeted neighbourhood contexts. It does appear that considerable weight and responsibility has been placed on Sure Start local programmes to dramatically change the lives – for the better - of young children and their families. However, how far should this responsibility be taken? The very nature of the contexts of programmes means that they are based in areas with issues/problems that could be regarded as more structural, such as inadequate housing, high unemployment, etc. It is these complex issues which local programmes face in their delivery: structural ambiguity; conflicting priorities; resource shortfalls; local history and co-ordination of a range of professional opinions.

All these can clearly impact on 'success' or 'impact', and this is where evaluation comes in. There is a general consensus that many of the impacts of Sure Start may well not be measurable for ten years (Jowell, 2003, quoted in Tunstill, *et al.*, 2005), or even once the targeted children have grown up, so 20 years (Factcheck, 2006). At present, therefore, it is not possible to say from the evidence presented that Sure Start 'works', though we can say that it has reportedly been implemented more or less according to plan. What research evidence does exist in the academic literature, are examples of dissemination of successful projects (for example, Brown and Liddle, 2005; Bagley *et al.,* 2004; Urwin, 2003; Morris and Leavey, 2006; Weinberger *et al,* 2005) which may signal attempts to 'spread the word', though what we know of professional cultures of learning suggest that grassroots networks would have greater impact.

The idea is that Sure Start is designed to promote innovation, or at least, innovative ways of working, and indeed local programmes have been described as a 'test bed' for developing new and imaginative strategies to promote access to services (Garbers *et al.*, 2006) – so is it more likely to

fail? For each positive finding from Sure Start, there appears to be an equally damning one. A large scale study suggests that the impact on deprived families is not universal, specifically, that *"SSLPs seem to benefit relatively less socially deprived parents (who have greater personal resources) and their children but seem to have an adverse effect on the most disadvantaged children"* (Belsky *et al.,* 2006:1476). While another study of personal and cognitive outcomes (Schneider *et al.,* 2006) found that there were no discernable impacts on educational outcomes and only tentative links to be made between Sure Start and better non-academic outcomes. Contradictory evidence such as these examples as led Belsky to conclude: '*there is enough variation in the results that there is grist for everyone's mill*' (quoted in Factcheck, 2006): the results of the evaluations can be either used to 'support or to trash Sure Start'.

Conclusions

Many local projects have carried out minimal evaluation and evaluators have struggled to support them and to carry out evaluation (Allen and Black: 242), and that was our experience (see also Hall and Clark, 2007). Our relationships with members of the SSLP remained positive and informative on both sides, evidenced by the 'away-days' where understandings from fieldwork were explored collaboratively and new explanations sought. The extent to which we arrived at a mutually acceptable compromise with strategic managers can be judged by the circumstances that the final report has been accepted by the LA, but not published and we are not naming the authority in this paper.

Although this paper is a case study of one local Sure Start programme, it is a focused study which raises many issues which can act as both a theoretical and as a practical resource for evaluators, policy-makers and practitioners as they put together and further develop structures such as Children's Centres. Though too early to see the long-term impact of Children's Centres, a recent report from the National Audit Office (NAO, 2006) makes for depressing reading. They report that 56% of LAs are not monitoring the performance of their Centres and almost as many (52%) are not working to identify the cost or cost-effectiveness of services. The role of Universities as evaluators will continue to be important, and we need to remember that unpublished lessons are more likely to be forgotten than learned.

Several clear messages emerge from our experience. Fundamentally, the purpose of the evaluation needs to be explicitly agreed between local government, national government and the evaluators. All parties will then, hopefully, have a much more realistic idea and expectation of what is actually achievable. Competing agendas may be inevitable in local evaluation work but it would be as well to know who is competing from the start. Secondly, it is vital that all parties understand what is realistic for

the evaluation to explore alongside the extent to which learning and development are goals of the evaluation. Often learning is an 'added extra' of a process which is more about validating implementation processes. Finally, an important consideration is how evidence collected can be potentially be used. The areas of vulnerability need to be openly discussed. Even where there is a commitment to a formative evaluation, some areas of organisation or practice may be deemed 'off limits', often for important operational reasons. This would fit with what Spicer and Smith (2008) refer to as fostering a genuine 'dialogic relationship' – meaning that dialogue can help to ensure ongoing relevance of an evaluation as such programmes naturally develop.

Michael Rutter, a professor at the Institute of Psychiatry who criticised the Government after advising it over its Sure Start programme for families in disadvantaged areas, said: "*The Government definitely doesn't want evidence, although the rhetoric is entirely different.*" (2006). The structures and governance of dual level reporting does not seem to have worked effectively for Sure Start. One could argue that our initial conceptual expectation of being central to the national and local programmes was somewhat naive, perhaps even unrealistic. However, guidance and frameworks did lead us to believe that this would be the case and this was not the reality. Similar experiences are borne out through recent works such as Spicer and Smith (2008) who, with reference to evaluating complex, area-based initiatives, argue: "*independent local evaluations commissioned by the partnerships have had limited influence on programme design and implementation, and indeed have been highly contested in their use*" (p.75).

References

Alcock, P. and Craig, G. (1998) 'Monitoring and evaluation of local authority anti-poverty strategies in the UK', *International Journal of Public Sector Management*, 11, 7, 553-565.

Allen, M. and Black, M. (2006) 'Dual Level Evaluation and Complex Community Initiatives: the local evaluation of Sure Start'. *Evaluation,* 12, 2, 237-249.

Bagley, C., Ackerley, C.L. and Rattray, J. (2004) 'Social exclusion, Sure Start and organizational social capital: evaluating inter-disciplinary multi-agency working in an education and health work programme'. *Journal of Education Policy* 19, 5, 595-607.

Belsky, J., Melhuish, E. Barnes, J., Leyland, A. H., Romaniuk, H. and the NESS Research team (2006) 'Effects of Sure Start Local Programmes on children and families: early findings from a quasi-experimental, cross-sectional study'. *British Medical Journal* 332, 7556, 1476-1479.

Boot, A. and Macdonald, G. 'Parents and Sure Start Evaluation; Suggestions for Partnership'. *Child Care in Practice* 12, 3, 269-281.

Boyne, G. A. (2002) 'Concepts and Indicators of Local Authority Performance: An evaluation of the statutory frameworks in England and Wales'. *Public Money and Management* 22, 2, 17-24.

Brown, B. and Liddle, J. (2005) 'Service Domains – The New Communities: A case study of Peterlee Sure Start, UK'. *Local Government Studies,* 31, 4, 449-473.

Cronbach, L.J. (1982). *Designing Evaluations of Educational and Social Programs.* San Francisco: Jossey-Bass.

DfES (2004) *Every Child matters: Change for Children.* London: The Stationery Office.

Dyson, A., Clark, J., Hall, E., Hall, I. and Roberts, B. (2000) *Increasing the Attainment of Under-Achieving Groups: a report to the independent advisory panel on Beacon Councils,* London: DETR.

Edgley, A. and Avis, M. (2006) 'Interprofessional collaboration: Sure Start, uncertain futures'. *Journal of Interprofessional Care* 20(4), 433-435.

Eisenstadt, N. (2000) 'Sure Start: Research into Practice, Practice into Research'. *Public Money and Management* 20, 4, 6-8.

Eisenstadt, N. (2002) 'Sure Start: Key Principles and Ethos'. *Child: care health and development* 28, 1, 3-4.

Factcheck (2006) *How effective is Sure Start?* Channel 4 News Report: http://www.channel4.com/news/special-reports/special-reports-storypage.jsp?id=2722. Accessed September 13th, 2007.

Freer, S. (2002) 'Comprehensive Performance Assessment in Local Government' *Public Money and Management,* 22, 2, 5-6.

Garbers, C., Tunstill, J., Allmock, D. and Akhurst, S. (2006) 'Facilitating access to services for children and families: lessons from Sure Start Local Programmes' *Child and Family Social Work,* 11, 287-296.

Geddes, M. and Martin, S. (2000) 'The policy and politics of Best Value: currents, crosscurrents and undercurrents in the new regime'. *Policy & Politics*, 28, No. 3, 379-395.

Gustafsson, U. and Driver, S. (2005) 'Parents, Power and Public Participation: Sure Start, and experiment in New Labour governance'. *Social Policy and Administration* 39, 5, 528-543.

Hall, E., and Clark, J., (2007) 'Ghosts at the Feast: The role of research centres in supporting innovative practice in local authorities'. *Studies in research: Evaluation, Impact and Training* 2007, 2, 1-9.

Hall, E., and Clark, J., (2004) *Sure Start Evaluation: Baseline survey*, Newcastle University: Research Centre for Learning and Teaching.

Hall, E. and Clark, J. (2005) *Exploring multi-professionalism and the changing roles of health professionals within Sure Start*, Newcastle University: Research Centre for Learning and Teaching.

Hodge, M. Written answer on Sure Start Funding, *Hansard,* 1st September 2004 http://www.publications.parliament.uk/pa/cm200304/cmhansrd/vo040901/text/40901w29.htm Accessed September 13th, 2007.

Jones, S. and Hughes, J. (2001) 'Understanding IS evaluation as a complex social process: a case study of a UK local authority'. *European Journal of Information Systems* 10, 4, 189-203

Jowell, P. (2003) *Trying it Out: the role of 'pilots' in policy-making,* London: Cabinet Office.

Morris, T. and Leavey, G. (2006) 'Promoting phonological awareness in nursery aged children through a Sure Start Early Listening programme'. *International Journal of Early Years Education,* 14, 2, 155-168.

Moss, P. (2004) 'Sure Start' *Journal of Education Policy* 19, 5 631-4.

National Audit Office (2006) *Sure Start children's centres*, London: The Stationery Office.

Ormerod, P. (2005) 'The Impact of Sure Start' *The Political Quarterly* 76, 4, 565-567.

Rutter, M. (2006) quoted in 'Ministers Vilify Researchers' article in *Times Higher Education Supplement*, published 1st December 2006.

Sanderson, I., Martin, S. and Dowson, L. (2003) *Beyond Performance Management? Evaluation for Effective Local Governance*, End of Award Report. Swindon: Economic and Social Research Council.

Sanderson, I. (2001) 'Performance Management, Evaluation and Learning in 'Modern' Local Government'. *Public Administration*, 79, No.2, 297-313.

Schneider, J., Ramsay, A and Lowerson, S.A. (2006) 'Sure Start graduates: predictors of attainment on starting school'. *Child: Care Health and Development* 32, 4, 431-440.

Spicer, N. and Smith, P. (2008) 'Evaluating complex, area-based initiatives in a context of change' *Evaluation*, 14, 1, 75-90.

Sure Start (2006) *Guidance on Evaluating Sure Start: Planning and delivering Sure Start – 6th Wave Edition*. Sure Start.

Tunstill, J., Allnock, D., Akhurst, S., Garbers, C., Meadows, P. (2005) *Implementing Sure Start Local Programmes: An In-Depth study,* London: DfES.

Tunstill, J., Allnock, D., Akhurst, S., Garbers, C. and the NESS Research team (2005) 'Sure Start Local Programmes: Implications of Case Study Data from the National Evaluation of Sure Start'. *Children and Society* 19, 2, 158-171.

Urwin, C. (2003) 'Breaking ground, hitting ground: A Sure Start rapid response service for parents and their under fours'. *Journal of Child Psychotherapy* 29, 3, 375-392.

Willan, J. (2007) 'Delivering Sure Start in rural communities'. *Early Years*, 27 (1) 19-31.

Weinberger, J., Pickstone, C. and Hannon, P. (2005) *Learning from Sure Start: Working with young children and their families,* Maidenhead: Open University Press.

Weiss, H. (2003) 'Editorial of 'The Evaluation Exchange', *Harvard Family Research Project*, IX, 3, 1-20.