

Evaluating Local Implementation: An Evidence-based Approach

Paper submission to:
Policy and Society

Originally presented at:
IRSPM XIII
Copenhagen, April 2009

Adapted from Conference Paper Entitled:
Problem-inspired or problem-solving?
Evaluating evidence-based policy

Tracey Wond
Researcher
Teesside Business School
University of Teesside
Middlesbrough
UK
TS1 3BA
+44 (1) 642 342 694
t.wond@tees.ac.uk

Dr Michael Macaulay
Reader in Governance
Teesside Business School
University of Teesside
Middlesbrough
UK
TS1 3BA
+44 (1) 642 342 894
m.macaulay@tees.ac.uk

Evaluating Local Implementation: an Evidence-based Approach

Abstract

This article, based on data collected from a year-long study, investigates the evaluation of a UK local government policy implementation and the use of evaluation data as an evidence base for public policy (McCoy and Hargie, 2001; Schofield, 2004; Bovaird and Loeffler, 2007; Stern, 2008). Our case study highlights a number of issues. First, uncertainty and ambiguity of policy direction can lead to barriers in establishing clear evaluation goals, which, second, results in frustration among stakeholders at a perceived disparity between what we term *problem-inspired* policy and *problem-solving* policy. Finally, this perception can be compounded by a lack of consideration for local variations of, for example, specific cultures, geographies or historical contexts. In responding to these problems our article argues that regardless of where policy control and decision-making occurs, the importance of the experiences of implementers of policy at a local level (where subject/geographical/cultural specialism and familiarisation exists) is crucial.

Keywords

Evaluation; evidence-base; public policy; local government; implementation

Introduction

There is nothing a government hates more than to be well-informed; for it makes the process of arriving at decisions much more complicated and difficult (Keynes cited in Parsons, 1995:393).

This article investigates an evaluation of a local government implementation in the North East of England. Drawing upon data from a year-long ethnographic study it seeks to highlight a number of issues surrounding policy development and local policy implementation, including: the disparity between central guidance and local implementation; perceptions of data as policy evidence; and the utilisation of evaluation data as evidence to inform policy design.

The study highlights a number of key features concerning evaluation as an evidence-base. First, ambiguous policy direction can obstruct the setting of well-defined evaluation goals, resulting in frustration among stakeholders. Second, this frustration can manifest itself as a perceived disparity between what we term 'problem-inspired policy' and 'problem-solving policy' approaches. Finally, problems can be compounded by a lack of consideration of local variations in terms of specific cultural, geographic and historical contexts.

In highlighting these issues, this paper echoes Stern's (2008:250) argument that evaluators recognise 'policy debates and the world of policy-makers' and confirms that there is a great potential for influences, political or otherwise, to impinge on

policy development and implementation. Evaluation needs to be sensitive to such influences and the disparate perspectives of policy-makers and policy-implementers; regardless of where policy control and decision-making occurs, the importance of the experiences of implementers of policy at a local level (where subject/ geographical/ cultural specialism and familiarisation exists) is crucial.

The rise of evaluation

Evaluation studies have long been given prominence in determining the effectiveness of policies and programmes (Barbier, 1999; Bovaird and Loeffler, 2007), attracting 'increasing attention from researchers, policy-makers and theorists' (McCoy and Hargie, 2001:325). This evidence is required:

to tell us what works and why, and what types of policy initiatives are likely to be most effective (and)...must vastly improve the quality and sensitivity of the complex and often constrained decisions we, as politicians, have to make (DfEE, 2000:24,25)

Sadly, as we argue, the detail of *why* a policy initiative is successful or not is often not always gained due to the type of data requested for national evaluation purposes. Nonetheless, almost a decade ago Davies (1999:150) noted 'the widespread expectation that policy and resource allocation decisions be based on sound information about the performance of public programmes', an argument that has been consistently reconfirmed since (for example see McCoy and Hargie, 2001; Schofield, 2004; Stern, 2008). Several government publications also sustain this point (Cabinet Office, 1999a, 1999b, 2000). The attempt at public service

modernisation, intended to provide a more responsive, efficient and strategic provision (Cabinet Office, 1999), reaffirmed further a renewed emphasis on evidence-based policy paradigms (Shaw, 1999). The development and rise of evaluation to fill this need was therefore a logical response, and despite a lack of evidence demonstrating policy effectiveness, Albaek (1995) argues that the use of evaluation data is increasing. Rational models of policy fulfilment see evaluation as a tool for determining whether policies or programmes have met their objectives (Sanderson, 2000) clearly a task capable of satisfying numerous stakeholders when assuming that all are agreeable to the policy and programmes objectives and targets.

One immediate issue is that, as Biott and Cook (2000:410) argue, although 'a range of data sources' are invited at local levels (perhaps to demonstrate engagement), in practice national solutions/frameworks have usually predetermined. With numerous purposes for evaluation, a further local-national dichotomy emerges with local evaluation implementation argued to be for learning and formative purposes, whilst evaluation nationally concerned with accountability (Martin and Sanderson, 1999). Martin and Sanderson's argument that 'what matters is what works' (1999 in Sanderson, 2000:433) echoed New Labour's 1997 manifesto pledge, which was itself a cornerstone for the alleged development of evidence-based policy. There is also suggestion that local evaluators may encounter issues around control at a national level, another phenomenon that can be mirrored in our policy-related findings and a concern conveyed through this paper (Biott and Cook, 2000; Grimshaw and Stewart, 1999). In her study of the 'LEADER' community initiative, Saraceno (1999) made clear that issues existed between national and local policy,

and evaluation; and also noted significant differences in evaluation approaches at national and local levels which make comparability difficult.

Our case study demonstrates all of these issues and more, particularly including timing issues, varying stakeholder interests and an overreliance on quantitative measurement. Fundamentally, however, it reveals a central-local tension that informs the way that implementation evaluations are strategically viewed.

The evaluation case study

This paper is based on an ongoing evaluation programme that the authors are currently undertaking. The evaluation is being conducted on a three-year local economic development programme, funded centrally by the Department of Communities and Local Government and implemented by an English local authority.

Nationally, there are twenty such programmes, worth around £300 million collectively, and intended to develop sustainable economic growth in under-developed and deprived areas; in some cases neighbouring local authorities have united to deliver joint programmes in their areas. The case programme began in 2007 and is due to end in 2010 (although this has been subject to much change), the programme attempts to kick-start enterprise by encouraging new business start-ups, as well as aiming to help existing businesses grow. As previously stated, the authors are providing the local evaluation for the case study programme.

Each local authority (or partnership) was free to choose its own pattern of delivery and evaluation. With a lack of national evaluation direction in the first 18-months of funding release, local evaluation activity was uncertain and many programmes chose to perform more of an output-driven quantitative performance measurement strategy than formative evaluation activity pending further guidance, such a delay was also seen by Biott and Cook (2000) with a 12 month gap between local and national evaluation initiation. However, within the case programme an evaluation strategy was devised, agreed and began to be delivered, despite this lack of national activity.

The evaluation strategy was designed by the authors of this article, and contained two components; a qualitative element and an objective-based quantitative framework reflecting the specific objectives highlighted by the case authority: occupational segregation; environmental regeneration; and, business activity adversities (such as failure rates) were included. Initially, programme delivery sought to subdivide the programme into six projects each concentrating on a specific economic issue, this was adjusted as the programme progressed although was limited by the proposal on which funding was released. The beneficiaries of the projects exhibited considerable diversity in terms of their social, economic and demographic status, which led to significant discrepancies in modes of both programme and evaluation delivery and monitoring. To an extent therefore a flexible delivery approach was required to adjust to the contextual demand although this was done with some hesitation by the programme implementers who felt confined by national guidelines and chance of reprisals for such changes

In one sense, flexible delivery can be beneficial, providing bespoke and geocentric interventions to communities but can have negative consequences for evaluation. Data is more difficult to initially capture, for example, which in turn provides obstacles for measurement, communication and replication in future programmes. In addition, there is the ever present spectre of generalisability and transferability: there is no guarantee that one successful project could work in a different context. It also appears that this degree of flexibility will diminish with the late establishment of the national evaluation (a point also noted by Biott & Cook in their study of Early Years Excellence Centres).

The existence of both national and local evaluations within this programme fit with Allen and Black's (2006) notion of 'dual-level' evaluations, 'dual' being used to refer to the existence of both national and local level evaluations. As mentioned previously, with this bringing inherent issues such as control, purpose and flexibility.

Evaluating local implementation

As previously stated, programme evaluation findings and the impact of interventions may not materialise for many years, or even decades and therefore 'what works' is often an unknown quantity simply because policy initiatives are not given sufficient time to be evaluated properly. This problem of quick policy turnover was an immediate issue to the case evaluation. Although ostensibly it aimed towards a long-term, cultural transformation in an economically and socially disadvantaged area, the reality was based on immediate results. A vignette from the field-notes journal maintained throughout the study states:

Policy and Society

“The work of the programme team in schools appears greatly in line with the programme aims and objectives, and meets some of the outputs desired by the Local Authority and nationally. However, the real outcomes won’t be seen for a long-time – have the interventions had an effect in 6 years time when primary school age students decide which path they choose – further education, work, to be self-employed, or if unsuccessful to continue to what is now third generation unemployment for many?...by this time the programme and I (the evaluator) will be gone, a new programme will have been funded, and little regard will be paid to the good (or bad) outcomes that are beginning to fruit from this historic effort” (Extract from Fieldnotes Journal, 12 June 2008)

This policy included interventions aimed to transform the attitudes of individuals who have encountered two generations of unemployment within their families; yet evaluation data only considers the three-year life of the programme and is therefore significantly short against the potential time span for these interventions to materialise. Sabatier (1986) describes this as ‘premature assessment’ and subsequently calls for process evaluation that looks more specifically at the actual implementation of policy rather than just what outcomes or achievements are produced. To complicate this issue further ongoing deliberations over programme length, and whether or not programmes should actually be retracted, suggest that rather than extend programmes (which would allow for longer and more comprehensive evaluation studies) there is a prevalence to adopt a quick-win, or abandonment strategy, in the UK economic policy arena.

The effectiveness of the latest programmes which replicate the intentions of programmes before it, must therefore be questioned, given that the evaluation

results of expiring programmes are often not even published by the time the new programmes are implemented:

In several recent cases (e.g. the Employment New Deals and Best Value in local government) the UK government has proceeded with policy commitments even though evaluation evidence from pilots has not demonstrated whether they work (Sanderson, 2000:448)

Subsequently, data is neither reviewed nor considered as evidence in the design of these substitutions (Pawson and Tilley, 1997; Sanderson, 2000); when evaluation data is reviewed for use in future programmes, the key points that the evidenced programme sought to address may no longer be relevant to the current context – highlighting expiration and generalisation risks.

It is also worth considering the overreliance on quantitative data in policy evaluation (at the request of national government office). This quantitative emphasis has received much criticism and calls over the years from evaluation theorists encouraging a move away from positivism to a more process-driven approach (Pawson and Tilley, 1997; Sabatier, 1986):

Many activities in the public policy realm, by their very nature, are complex and intangible and cannot be reduced to a numerical figure (Perrin, 1999 cited in Davies, 1999:155)

At first glance, for example, the quantitative indicators requested by local and central government would make our case study appear to be failing and even accelerating some adverse business scenarios such as survival rate, growth rate and employment. This would be unfair, if not wrong: the effects of the economic

downturn on these factors, however, have not even been considered as a contributory factor. Quantitative measures alone, therefore, do not tell the whole story. In contrast, a qualitative reflection finds that the programme has sufficiently accommodated its target groups by providing a comprehensive delivery strategy, aligned with other key support providers and which sees many business-based stakeholders being saved, strengthened or started. Again relating to good timing, this qualitative perspective requires the patience of policy-makers to allow qualitative evaluation research to be conducted and presented. Qualitative data allows a formative evaluation to contribute to programme delivery (Tessmer, 1993) a feature which quantitative, experimental evaluation design cannot offer. It is our belief that the quantitative nature of information required to form evidence within this case study stands as a barrier promoting, most simply, the identification of whether the programme works through these specified outputs but requiring very little rich detail to describe the processes integral to that success. This shortcoming is very much like collecting the ingredients required for a recipe but failing to gain the instructions on how these go together to create a successful, replicable product. The qualitative efforts of the local evaluation in this case programme are of no interest to the national evaluation efforts, despite their overwhelming ability to inform policy-makers of the successful processes implemented in this programme locally. The complex nature of the case programme also lends itself to a more qualitative approach (although that is not to say that a triangulated approach encapsulating both qualitative and quantitative data is less adequate).

Ultimately, policy is misinformed by these timing and methodological points leading to a seemingly futile evaluation effort, and reaffirming critics who question the political motivations for evaluation for accountability rather than learning purposes.

In addition, there is the rather context-dependent judgement of 'what works'? What is important (and relevant) to one group of stakeholders may compete with the interests of another group. This phenomenon was evident in the evaluation case study where there were several notable attempts by various stakeholders to influence the nature and purpose of the intended evaluation. Some, for example, wanted the evaluation to inform learning; others merely for auditing and accountability purposes; whilst some stakeholders saw its existence for no other reason than to check the box to say an evaluation had been done. The clash of interests seems to follow on from the way in which such programmes are initially implemented:

Most policy evaluations are commissioned by government departments and agencies and publicly-funded bodies accountable to them for the effective development and implementation of the policy. This has given a strong top-down orientation to evaluation with a major concern to promote accountability and control. (Sanderson, 2000:438)

Accountability and control, while crucial, potentially promote a competition of interests at the local level. Within the case programme, the interactions between policy-implementers and potential beneficiaries led to an interest in the optimal development of the programme to best serve these intended beneficiaries (attachment to programme participants morally and emotionally may perhaps feature within this interest). On-the-other-hand, local authority managers, as the accountable

body pressured the implementers to remain objective, focusing solely on those outcomes that influenced the indicators specified. Whose interests should be considered and prioritised for delivery and evaluation purposes in this case? - The potential beneficiaries, programme-implementers, Central Government and accountable body (in our case the local government authority) all have specific, yet varied, interests in the policy's implementation. Even those who benefit from the programme may have competing views of what actually works; primary data from the case evaluation clearly demonstrates a broad spectrum of opinions as to the success of the programme so far. Small and new start-up businesses (particularly those in the creative industries), for example, have welcomed the inclusion of action-learning and networking sets as part of the evaluation method, whereas medium-sized businesses and small start-ups in manual trades have resisted these approaches. Action learning, therefore, only appears to 'work' for particular beneficiary groups, a problem exacerbated in the case evaluation by geographical variations, collectively hampering attempts to define categorically 'what works'; perhaps conflicting with our own argument for process evaluation, yet still an observation far overlooked in current evidence-based policy and practice.

These twin problems – a short-term strategy coupled with a multiplicity of competing interests at the implementation level – potentially gives rise to substantial problems for organisational learning. Collective learning by policy-makers through evaluation may be thwarted by a failure to properly utilise evaluation data:

(evaluation serves two main functions) enabling accountability and collective learning. Both of these - and their combination - run into diverse complications when applied in the complex multi-actor policy processes. (Van Der Meer and Edelenbos, 2006:201)

At a policy-level, evaluations may actually have perverse effects on accountability, with a concentration on quantitative quick win outputs rather than substantive outcome-driven successes. There continues to be scope for a selective process of evaluation data whereby those datasets deemed measurable may not be required or even bear 'little resemblance to what is relevant' (Perrin,1999, cited in Davies, 1999:155). Again, this was reflected in the case evaluation, this time at the national/central level. Whereas programme implementers at the local authority sought ongoing and reiterative learning from the evaluation there was a notable lack of interest from regional or national government offices. Formative interim findings of our evaluation, for example, were not responded to and instead there was repeated emphasis on quick-wins and evidence-gathering for accountability purposes. Stringent attempts to remain objective-based both whilst evaluating interventions, implementing policies and reviewing evidence post-implementation may hinder flexibility and lose sight of the grander picture of fostering economic prosperity amongst our case programme; we suggest that these attempts to foster objectivity, are in fact inflexible and unresponsive, and most ironically may lead to a lack of objectivity at a macro level. For this reason, this paper calls for holism, with a focus on the bigger picture rather than on objectives, proposals and bid promises made at the programme start and which may no longer require the type of interventions first discussed.

A further observation within the case study was that uncertainty over policy direction led to frustration among internal stakeholders, as well as creating barriers in establishing clear evaluation goals. The potential complexity for this to change adds

further speculation as to the purpose and potential changing purpose of the evaluation, with a deviation from original policy objectives not uncommon, and as Pollitt highlights: 'the purpose of evaluation at policy-makers level may change with evaluation findings being overturned in favour of these new intentions' (2003:114). The evaluation had to respond to shifts in both central policy and local 'ad hoc' implementation. For example, one of the programme's initial local objectives was to develop certain specialist sectors, notably the chemical industry and renewable sector. Within 16-months of the programme being launched this objective was refined and the programme began actively pursuing the tourism sector. Clearly this impacted upon the evaluation; participants of the tourism strand were very much different to the chemical industry and renewable sector. Therefore the methodology had to undergo substantial redesign in terms of familiarisation with the sector and stakeholders and how success (or otherwise) could be measured in terms of performance indicators, etc.

Further ambiguity was encountered though the paucity of central guidance over the relative importance evidence-based and accountability perspectives. Potentially, the data collected locally could contain no comparable, consistent or common indicators as a result of this guidance deficit. Schofield (2004:283) also addresses such policy and evaluation contentions highlighting 'ambiguous policy guidance' in a health service context. Taken together we suggest that this forms two distinct policy approaches: *problem-inspired* and *problem-solving*.

Competing approaches to evaluation: problem-inspired or problem-solving?

The perceived disparity between what we term problem-inspired policy and problem-solving policy resulted again in frustration by multiple stakeholders, triggered by ineffective and ill-informed policy formulation based upon the false premise that policy is founded upon a reliable evidence-base, and under the assumption that 'operationalizing such policy was doable' (Schofield, 2004:284).

Problem-inspired approaches refer to the reaction of policy makers who are inspired by the evaluation data and evidence of past programmes and social research and as a result initiate policies to redress these. We also suggest however that locally there is a better alignment on local needs, this approach we term a problem-solving approach.

Approach:	Problem-Inspired	Problem-Solving
Characteristics:	Predefined and decide among policy-makers	Adhoc and responsive approach taken by policy-implementers – e.g. programme team
	Designed with best intention to meet socio-economic problems using an evidence-base (albeit flawed?) and generalised	Interventions taken to observed socio-economic problems within a specific, 'real', immediate and local context
Evaluative consequence:	Fails to identify 'what' and 'why' an intervention succeeds (or fails)	Qualitative evaluation which includes consideration of process and 'how' the programme/policy works
	Measured in accordance with national guidelines or to meet	Difficult to measure due to extensive possibilities of

Policy and Society

	local authority agreed indicators - quantitative	interventions being implemented
	Fails to encourage qualitative evaluation	Greater learning purpose
Future policy impact/ formulation	The experiences and ad-hoc learning that a problem solving approach establishes never comes to light and so hinders learning	Acceptance of such a response beneficial for future programme delivery
	Programme continues to be required to work off a quantitative evidence-base which fails to learn from previous process	Can be used to develop future policies and programmes

Figure 1: Problem-inspired and problem-solving approach overview

The problem-inspired approach, taken by policy-makers and forced upon policy-implementers, although well intended, frequently fails to understand what does (or does not) work and does little to encourage qualitative evaluation, which enables an understanding of pragmatic responses by programme teams and an appropriate use of evaluation findings in a timely and constructive manner. Local level implementers perceived as *problem-so/vers*, adopting a pragmatic and responsive approach, so as to generate improvements that are context-specific. Although this is an understandable response, there is a strategic weakness in data conceptualisation, analysis and utilisation by policy-makers so that these real and pragmatic solutions to policy problems can be properly understood.

The problem-inspired/problem-solving dichotomy can be used to legitimise this pragmatism and ad-hoc decision-making: policy-implementers consider there to be

no choice but to react to unexpected pressures and events. As shown in the case study, there were many instances where informal and unplanned activity was undertaken with ensuing success, but due to the mainly quantitative reporting requirements of the government these 'success factors' will never be considered for future policy and programme formulation.

The familiar debate regarding the capability of policy-implementers to translate these policy goals and apply policy (Barrett and Fudge, 1981; Brodtkin, 1990), may also be relevant at this point. Although arguably the regulatory involvement of internal evaluation and performance teams act to bridge this gap:

the institutionalization of evaluation has systematized evaluation commissioning and procurement, and installed skilled internal evaluation managers within many agencies who form a tier of interlocutors between evaluation and policy-makers and a channel for the involvement of many more policy interests in the evaluation process. (Stern, 2008:250)

Both evaluators (in terms of the need to consider the following issue in methodological design) and policy-implementers (as local knowledge source) can face issues in terms of local diversity (mirroring theories in international business relating to local reflection and risks of ethnocentrism). A lack of knowledge of specific local variations, for example specific cultures, geographies, historical contexts by policy-makers or evaluators, and with it an over-generalised approach, could hinder the ability of the programme to meet its target beneficiaries and stakeholders. Our study found that both the evaluator and implementers needed to understand the historical context of the area and residents, on researching the area more the impact of mining and steel industry collapses were apparent:

“It was interesting to discover that the old industries in the area provided plentiful work and the expectation that people would work in these companies on leaving school, the career path for people in the area particularly in the mining villages was essentially pre-made. On talking to X it seems that females filled the administrative roles and the males committed to manual tasks. Without the industries there is a) less alternative work without travelling, b) no preset career path, c) a dying expectation and I suppose peer pressure to work as many generations before did in those industries.” (Extract from field notes journal, 13th February 2008)

We therefore maintain that regardless of where policy control and decision-making occurs, the importance of the experiences of implementers of policy at a local level (where subject/geographical/cultural specialism and familiarisation exists) is crucial. The ethnocentrism upheld by policy-implementers needs to give way to a more geocentric approach which appreciates the localism of target areas and gives programmes the ongoing discretion and flexibility to adjust their programme. Policy-implementers, including programme staff, managers and specialists in the local policy field, arguably know more about the issues relevant to the areas that the policy is being applied, therefore can respond and align the programme to fit accordingly and therefore combined with this internal evaluation support are able to implement, measure and align delivery with policy far more closely. The bid process for the case programme allowed for such local factors to be considered and put forward by bidders, namely local authorities, although, it is 'often suggested that a policy has failed when the outcome deviates from the original policy regardless of whether the eventual outcome exceeded the initial problem that the policy or programme set about to redress (Schofield, 2004:283).

Thus, we also suggest that generalisability may be an issue for consideration (Pawson and Tilley, 1997; Shaw, 1999; Sanderson, 2000). In suggesting so we must also highlight the difficulty for evaluation to provide data which reflects adequately the 'real' situation in that geography. The quantitative indicators specified by national government, call for generalisations and common measures to be reported – context-free of factors more fitting to that area such as history, or for the more responsive programme which may appear to miss its national targets but that addresses geocentrically the 'real' problems that matter at a local level to that particular area.

Conclusion

Evaluation needs to be sensitive to these influences and the differences in perspective between policy-makers and policy-implementers. Fundamentally, policy designers and implementers are different and it is conceivable that neither entirely understand one another:

It is unlikely that policy designers can anticipate the operational consequences of their initiatives because they are too far removed from operational management (Schofield, 2004:284)

Our case confirms that there is a great potential for influences, political or otherwise, to impinge on policy development and implementation, and echoes Stern's call for evaluation to ensure familiarisation with 'policy debates and the world of policy-

makers' so as to attempt to influence and improve the current evaluation role in evidence-based policy and practice (2008:250).

In local implementations, the evidence-base used by policy-makers cannot be taken as Gospel: evidence collection may have to be a fluid and flexible commodity, responding to the changing policy needs and ad-hoc delivery that programmes find themselves taking. Furthermore, the ambiguity at a policy level in terms of programme length, newer or complementing programmes (for the purpose of ensuring there is no replication) can hinder both an evaluator and programme with understandable frustrations on both parts.

Although, we originally set about to judge the reliability of this evaluation evidence, on further thought and the development of our response this was found to judge the evaluation community on their efforts:

There has been extensive soul-searching amongst the evaluation community about the perceived lack of impact evaluation research on the world of politics and practice (Sanderson, 2000:435)

It appears that the issue is more complex and contingent on effective policy practice and process, outside the direct responsibility of the evaluator.

Under-utilisation of such evidence at a policy level, calls for quantitative rather than qualitative measurements, and the pragmatism practiced but perhaps not formalised by implementers at a local level restrict the potential for learning and refining programme delivery and policy design. Furthermore, we posit that policy is

principally rhetoric and theoretically ideal; drawing upon evidence to ensure that well-fitted solutions to the problems encountered are addressed satisfactorily, which does indeed sound perfect and a contrast to the observed reality of evaluation evidences' journey - improper, underused, and when used then with delay and thus allowing possible expiration risks.

We would suggest that ours is not an isolated case. The contributions of others involved in local implementation and evaluation of economic regeneration and employment programmes also suggest that such issues are common (those we have exemplified include Sure Start, Health Action Zones and LEADER initiatives).

To dismiss the use of evidence in policymaking entirely is asinine and foolish, for 'without soundly based analysis and modelling, those involved in the formulation of policy and the delivery of services will work in the dark' (Cabinet Office, 2003). What is needed is more far-reaching approach to evaluation evidence, alongside a better utilisation strategy that considers the points we raise would assist in addressing many frictions. As Keynes suggests 'There is nothing a government hates more than to be well-informed; for it makes the process of arriving at decisions much more complicated and difficult' (Keynes cited in Parsons, 1995:393) a wealth of evidence does not make either implementation or evaluation any easier.

References:

- Albaek, E. (1995), Policy Evaluation: Design and Utilization, in Rist, R.C. (ed.) Policy Evaluation: Linking Theory to Practice. Aldershot: Edward Elgar
- Allen, M. & Black, M. (2006) Dual Level Evaluation and Complex Community Initiatives: The Local Evaluation of Sure Start, Evaluation, 12, pp.237-249

Barbier, J. (1999), Inter-governmental Evaluation: Balancing Stakeholders' Expectations with Enlightenment Objectives, *Evaluation*, 5, pp.373-385

Barrett, S. and Fudge, C. (1981), *Policy and Action: Essays on the Implementation of Public Policy*, London: Methuen

Biott, C. and Cook, T. (2000) Local Evaluation in a National Early Years Excellence Centres Pilot Programme: Integrating Performance Management and Participatory Evaluation, *Evaluation*, 6, pp.399-413

Brodkin, E. Z. (1990). "Implementation as Politics" in D.J. Palumbo & D.J. Calista (Eds.), in *Implementation and the policy process: Opening up the black box*. New York: Greenwood Press.

Bovaird, T. and Loeffler, E. (2007), Assessing the Quality of Local Governance: A Case Study of Public Services, *Public Money & Management*, Sept 2007, pp. 293-299

Cabinet Office (1999a) *Modernising Government*. White Paper. Cm 4310. London, HMSO. (<http://www.policyhub.gov.uk/docs/modgov.pdf>)

Cabinet Office (1999b) *Professional Policy Making for the 21st Century*. A report by the Strategic Policy Making Team. London, HMSO. (<http://www.civilservant.org.uk/profpolicymaking.pdf>)

Cabinet Office (2000) *Adding it up: improving analysis and modelling in Central Government*. A Performance and Innovation Unit Report. London, HMSO. (<http://www.policyhub.gov.uk/docs/addingitup.pdf>)

Cabinet Office (2003) *Magenta Book: Guidance Notes for Policy Evaluation and Analysis*, July 2003, Chapter 1, http://www.nationalschool.gov.uk/policyhub/downloads/Chapter_1.pdf

Davies, I.C. (1999) Evaluation and Performance Management in Government, *Evaluation*, 5, pp.150- 159

DfEE (2000) *Influence of Irrelevance: Can Social Science Improve Government?* Speech by the Secretary of State to the ERSC. London: Department for Education and Employment.

Grimshaw, L. and M. Stewart (1999) 'Joined Up Evaluation: the Evaluation of Area Based Initiatives', paper presented to the United Kingdom Evaluation Society Annual Conference, Heriot-Watt University, Edinburgh, December.

Keynes, J.M. cited in Parsons, W. (1995) *Public Policy*, Edward Elgar: Cheltenham

Martin, S. and Sanderson, I. (1999) Evaluating Public Policy Experiments: Measuring Outcomes, Monitoring Progress or Managing Pilots? *Evaluation*, 5(3), pp.245-58

McCoy, M. and Hargie, O.D.W. (2001) Evaluating Evaluation: Implications for Assessing Quality, *International Journal of Health Care Quality Assurance*, 14(7), pp.317-327

Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*, London: Sage

Perrin, B. (1999) in Davies, I.C. (1999) *Evaluation and Performance Management in Government*, *Evaluation*, 5, p.155

Sabatier, P. (1986) What Can We Learn from Implementation Studies? In Kaufman, F., Majone, G., Ostrom, V. And Wirth, W. (eds) *Guidance, Control and Evaluation in the Public Sector*, Walter de Gruyter: London

Sanderson, I. (2000) Evaluation in Complex Policy Systems, *Evaluation*, 6(4), pp.433-454

Saraceno, E. (1999) The Evaluation of Local Policy Making in Europe: Learning from the Leader Community Initiative, *Evaluation*, 5, p439-457

Schofield, J. (2004) A model of learned implementation, *Public Administration*, 82(2), pp.283–308.

Shaw, I. (1999) *Qualitative Evaluation*. London: Sage

Stern, E. (2008) Evaluation: Critical for Whom and Connected to What? *Evaluation*, 14(2), pp.249-257

Tessmer, M. (1993) *Planning and Conducting Formative Evaluations*, Kogan Page, London.

Van Der Meer, F-B. and Edelenbos, J. (2006) Evaluation in Multi-Actor Policy Processes, *Evaluation*, 12(2), pp.201-218