

2004

The national evaluation of Sure Start local programmes in England

Angela Anning
University of London

Mog Ball
University of London

Jacqueline Barnes
University of London

Jay Belsky
University of London

Beverley Botting
Office for National Statistics

See next page for additional authors

Follow this and additional works at: <https://ro.uow.edu.au/sspapers>



Part of the [Education Commons](#), and the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Anning, Angela; Ball, Mog; Barnes, Jacqueline; Belsky, Jay; Botting, Beverley; Frost, Martin; Kurtz, Zarrina; Leyland, Alastair H.; Meadows, Pamela; Melhuish, Edward; and Tunstill, Jane, "The national evaluation of Sure Start local programmes in England" (2004). *Faculty of Social Sciences - Papers*. 1975.
<https://ro.uow.edu.au/sspapers/1975>

The national evaluation of Sure Start local programmes in England

Abstract

In the latter part of the 20th century evidence was accumulating about the effectiveness of various intervention programmes for young children in disadvantaged families. Some small-scale interventions were evaluated by randomised control trials as in the case of the Abecedarian project (Ramey et al., 2000), the High/Scope Perry Preschool Project (Schweinhart, Barnes, & Weikart, 1993) and evaluations of home visiting (e.g. Olds, 1997). Others were evaluated by quasi-experimental methods as with the large-scale Head Start project (e.g. Barnett, 1995; Kresh, 1998). Despite some weaknesses in the evidence for large-scale interventions, the accumulating evidence of the benefits of early intervention was very influential in the planning and launch of the Sure Start programme (Glass, 2003), now known as Sure Start local programmes.

Sure Start local programmes are central to the UK government's policy for combating the adverse effects of poverty and disadvantage on young children and their families. Hence they have potentially great importance in the lives of hundreds of thousands of families and young children. By 2004 there will be 524 Sure Start local programmes in disadvantaged areas in England that will be serving over 300,000 children and their families. Scotland, Wales and Northern Ireland have their own Sure Start programmes. This initiative represents a unique approach to early intervention for children aged 0–4, their families, and communities. The aim of Sure Start is:

To work with parents-to-be, parents and children, to promote the physical, intellectual and social development of babies and young children – particularly those who are disadvantaged – so that they can flourish at home and when they get to school, and thereby break the cycle of disadvantage for the current generation of young children.

Sure Start local programmes are designed to be comprehensive, community-based projects adapted to local needs, and making maximal use of local expertise and enthusiasms. Such a focus on local autonomy may capitalise on shared concerns of people for their community (Oliver, Smith, & Barker, 1998). The issue of whether preventative interventions should be targeted or universal in application is dealt with by Sure Start by targeting distinct areas of disadvantage and making the services universal within those areas. In this way there are the advantages of economy associated with targeted interventions and also the stigma-free advantage of universal provision.

While some research has reliably documented the benefits of early intervention, there has been much prevention work that has either not been evaluated or where evaluations are so flawed as to preclude meaningful conclusions (see Mrazek & Brown, 2002; Webster-Stratton & Taylor, 2001). The Sure Start programme recognised the need for rigorous evaluation and after a competitive commissioning process the National Evaluation of Sure Start (NESS) was initiated. Rather than providing a specific service, Sure Start local programmes involve changes to existing services. Improved services and community functioning are presumed to lead to enhanced family and community functioning that in turn enhance child development. This approach raises the following questions for NESS.

1. Do existing services change? (How and, if so, for which populations and under what conditions?)
2. Are delivered services improved? (How, and if so, for which populations and under what conditions?)
3. Do children, families, and communities benefit? (How, and if so, for which populations and under what conditions?)

NESS addresses these questions through four core research components: (1) implementation evaluation

that considers how programmes are operating and changing; (2) impact evaluation that considers effects of Sure Start local programmes upon children and families; (3) local context analysis that considers communities as units of analysis and how they function and change over time; and (4) cost-effectiveness analysis that examines economic return on investment of the Sure Start local programmes. The NESS team also provides technical support to smaller scale evaluations being conducted by local Sure Start programmes themselves. The four core research modules of the evaluation are described in turn.

Keywords

evaluation, england, sure, start, local, programmes, national

Disciplines

Education | Social and Behavioral Sciences

Publication Details

NESS Research Team (2004). The national evaluation of Sure Start local programmes in England. *Child and Adolescent Mental Health*, 9 (1), 2-8.

Authors

Angela Anning, Mog Ball, Jacqueline Barnes, Jay Belsky, Beverley Botting, Martin Frost, Zarrina Kurtz, Alastair H. Leyland, Pamela Meadows, Edward Melhuish, and Jane Tunstill

The National Evaluation of Sure Start Local Programmes in England

The NESS Research Team*

Introduction

In the latter part of the 20th century evidence was accumulating about the effectiveness of various intervention programmes for young children in disadvantaged families. Some small-scale interventions were evaluated by randomised control trials as in the case of the Abecedarian project (Ramey et al., 2000), the High/Scope Perry Preschool Project (Schweinhart, Barnes, & Weikart, 1993) and evaluations of home visiting (e.g. Olds, 1997). Others were evaluated by quasi-experimental methods as with the large-scale Head Start project (e.g. Barnett, 1995; Kresh, 1998). Despite some weaknesses in the evidence for large-scale interventions, the accumulating evidence of the benefits of early intervention was very influential in the planning and launch of the Sure Start programme (Glass, 2003), now known as Sure Start local programmes.

Sure Start local programmes are central to the UK government's policy for combating the adverse effects of poverty and disadvantage on young children and their families. Hence they have potentially great importance in the lives of hundreds of thousands of families and young children. By 2004 there will be 524 Sure Start local programmes in disadvantaged areas in England that will be serving over 300,000 children and their families. Scotland, Wales and Northern Ireland have their own Sure Start programmes. This initiative represents a unique approach to early intervention for children aged 0–4, their families, and communities. The aim of Sure Start is:

To work with parents-to-be, parents and children, to promote the physical, intellectual and social development of babies and young children – particularly those who are disadvantaged – so that they can flourish at home and when they get to school, and thereby break the cycle of disadvantage for the current generation of young children.

Sure Start local programmes are designed to be comprehensive, community-based projects adapted to local needs, and making maximal use of local expertise and enthusiasms. Such a focus on local autonomy may capitalise on shared concerns of people for their community (Oliver, Smith, & Barker, 1998). The issue of whether preventative interventions should be targeted or universal in application is dealt with by Sure Start by targeting distinct areas of disadvantage and making the services universal within those areas. In this way there are the advantages of economy associated with targeted

interventions and also the stigma-free advantage of universal provision.

While some research has reliably documented the benefits of early intervention, there has been much prevention work that has either not been evaluated or where evaluations are so flawed as to preclude meaningful conclusions (see Mrazek & Brown, 2002; Webster-Stratton & Taylor, 2001). The Sure Start programme recognised the need for rigorous evaluation and after a competitive commissioning process the National Evaluation of Sure Start (NESS) was initiated. Rather than providing a specific service, Sure Start local programmes involve *changes to existing services*. Improved services and community functioning are presumed to lead to enhanced family and community functioning that in turn enhance child development. This approach raises the following questions for NESS.

1. *Do existing services change?* (How and, if so, for which populations and under what conditions?)
2. *Are delivered services improved?* (How, and if so, for which populations and under what conditions?)
3. *Do children, families, and communities benefit?* (How, and if so, for which populations and under what conditions?)

NESS addresses these questions through four core research components: (1) *implementation evaluation* that considers how programmes are operating and changing; (2) *impact evaluation* that considers effects of Sure Start local programmes upon children and families; (3) *local context analysis* that considers communities as units of analysis and how they function and change over time; and (4) *cost-effectiveness analysis* that examines economic return on investment of the Sure Start local programmes. The NESS team also provides technical support to smaller scale evaluations being conducted by local Sure Start programmes themselves. The four core research modules of the evaluation are described in turn.

The implementation module

It has become clear that in order to evaluate satisfactorily a specific programme in terms of its impact at individual, family, or community level, comprehensive data must be collected on the programme itself. The issues of treatment fidelity and implementation quality are crucial in understanding the effectiveness of intervention programmes, as is their interaction with overall organisation of service delivery (Kanter, 1983; Pressman & Wildavsky, 1984). While the traditional image used to symbolise the implementation research task is

*NESS research team: A. Anning, M. Ball, J. Barnes, J. Belsky, B. Botting, M. Frost, Z. Kurtz, A. Leyland, P. Meadows, E. Melhuish and J. Tunstill.

the 'opening up of a black box' (which can often imply a set of relatively straightforward activities), much recent evaluation literature stresses the complexity of the challenges that face the 'implementation researcher' (Patton, 1997; Ghate, 2001). The NESS implementation module faces all these challenges. Firstly, it is one of four inter-locking modules; and secondly, this particular 'black box' cannot easily be isolated from the complexities of post-1997 policy development (New Labour) around social inclusion.

Some interventions have a clear model or set of models for the form that the intervention programme should take, e.g. the Better Beginnings project (Peters, Petrunka, & Arnold, 2003) and Early Head Start (Love et al., 2002). Also, the accumulated evidence about interventions indicates that particular methods of intervention are more effective (e.g. Schweinhart et al., 1993; Peters et al., 2003; Love et al., 2002; Olds, 1997). Despite this evidence, Sure Start did not include specification of the form that local programmes should take, although it did specify aims, targets and objectives. The priority given to local control meant that local partnerships had a very large measure of autonomy over the form that the programme would take in their area. This has resulted in the emergence of great diversity in Sure Start local programmes and this diversity provides an interesting challenge for evaluators in that the diversity of interventions will need to be documented and the effectiveness of the programme is likely to be highly specific to the nature of individual local programmes. Hence the documentation of the characteristics of Sure Start local programmes is crucial to understanding whether, and under what conditions, Sure Start local programmes are effective.

The overall aim of the implementation module is to produce a comprehensive picture of the first 260 Sure Start programmes, in terms of both process and components, which is capable of linking programme activity to outcomes for individual users and communities. To this end, its detailed objectives include the description of the full range of programme approaches to service design and delivery; the creation of a typology from these of implementation styles; the assessment of potential programme quality, derived from a range of quality indicators; an assessment of the extent to which programme implementation has conformed with programme specifications; and crucially, the provision of data on 'inputs' at programme level to the impact module.

The implementation module incorporates a range of quantitative and qualitative methods in order to achieve these objectives. There are three inter-linked components to the broader implementation study:

- A national questionnaire survey of all 260 Sure Start programmes Rounds 1 to 4, repeated three times at annual intervals
- An in-depth exploration of a representative 10% subset of 26 programmes as case studies, undertaken twice, with an 18 month interval
- A series of themed evaluations, dealing with 'over-arching' issues, which require exploration across more than one programme (e.g., father involvement, quality child care, prenatal/maternity services).

The actual methods vary, but they share a common focus on six programme dimensions:

- Management and co-ordination
- Access for families to Sure Start provision and services
- Community involvement
- Allocation of resources
- The quantity of services provided by the Sure Start programme.
- The quality of services provided by the Sure Start programme.

Challenges are posed by the public policy context within which the Sure Start initiative was conceptualised, and flexibility in methods is necessary to accommodate these challenges. Key elements in this context with a bearing on the implementation of the study include the emphasis on *preventive* rather than *reactive* work; the need for 'joined-up working'; the engagement of a wide range of professional and non-professional community members; and, perhaps most importantly, the government's 'managerial ethos' with its central emphasis on targets and performance indicators (Cutler & Waine, 2000; Newman, 2001). Many of the methodological challenges faced by the implementation module derive, directly or indirectly, from this context. For example, as with any attempt to measure preventive work, the range of activities, projects and actors to be encapsulated in the data is potentially vast. Therefore the national survey should elicit the necessary information yet be of a length that would not deter responses. Similarly, the extent of inter-agency working can entail very complex data. In order to maximise the scope, sensitivity and detail of the questionnaire data, a telephone hot-line provides advice and guidance where required, for the local Sure-Start programme managers with responsibility for completion of the survey questionnaire. This has been well used and underpinned positive and productive relationships with the programmes. Indeed, the response rate for the first application of the questionnaire was 92%, i.e. 118 out of 126 programmes.

If it is possible to select one over-riding challenge that affects every aspect of the implementation study, it is the government's emphasis on meeting targets, and the attendant managerial monitoring and evaluation, which currently characterises policy implementation nationally. This can result in both anxieties on the part of programme personnel completing the national survey, and a sense of evaluation and monitoring overload. At every stage of the study, the implementation team has sought to minimise the research burden imposed on programmes and, where possible, to demonstrate the value to programmes of the data they enable us to collect. Only with the cooperation of local programme personnel can the aims and objectives of the module be realised.

The results of the first application of the implementation national survey have been reported by Tunstill et al. (2002). Also, a themed study on getting local programmes started is reported by Ball (2002). Both of these reports include documentation of the difficulties

and successes in setting up local programmes and the fact that achieving full delivery of services is taking considerably longer than originally anticipated.

The impact module

The impact evaluation is designed to examine effects of Sure Start local programmes on children, families, and communities and to identify the conditions under which Sure Start proves most effective in enhancing child, family, and community functioning. The impact study is designed around an integrated cross-sectional, longitudinal framework. With data on implementation and community characteristics being secured by other modules, the impact evaluation starts by means of a cross-sectional study of 150 carefully selected Sure Start and 50 Control communities (randomly selected from Sure Start-to-be communities, i.e. communities in which Sure Start programmes will occur later). This enables an early evaluation of the impact of Sure Start on infants and 3-year olds and their families, and communities (via the cross-sectional design), while securing the advantages of a longitudinal study of a single birth cohort for the medium- and long-term evaluation of bedded-down Sure Start programmes. The latter is realised by longitudinally following families with infants in the cross-sectional phase of study when children are 3 and 5 years of age (and beyond). The longitudinally studied children and families are drawn from 100 of the 150 cross-sectionally studied Sure Start communities and compared with each other and similar children and families being longitudinally studied as part of the Millennium Cohort Study (Smith & Joshi, 2002), a cohort study involving over 18,000 children and families. Coordination with the Millennium Cohort research team ensures that there is a range of equivalent outcome measures present in both studies.

In a first round of strategic sampling, which forms the basis of the cross-sectional stage of the impact study, a subset of 150 of a total of 260 Sure Start local programme areas are selected for intensive investigation of child, family, and community functioning on the basis of (a) their geographic, demographic, and ethnic characteristics, (b) their approach to developing the local Sure Start programme, and (c) their initial progress in implementing Sure Start. Specifically, the sampling strategy selects Sure Start local programmes that systematically vary in their early success in implementing their programme as well as in their programme strategies, for inclusion in the cross-sectional phase of inquiry. This ensures the inclusion of programmes likely to be effective and increases the statistical power to detect programme effects. Eighty families with a 9-month-old infant, and 25 families with a 3-year-old, are studied in each of 150 carefully selected Sure Start communities and 50 randomly selected Sure Start-to-be (i.e., control) areas. Thus within 3 years, the evaluation provides data from communities, parents and children of diverse age, and will be in a strong position early in the national evaluation to detect effects of Sure Start.

The second round of strategic sampling takes place at the end of the cross-sectional data collection. Information on child and family functioning collected

in the cross-sectional study of 150 Sure Start local programme areas will be used to identify those in which children and families are functioning better and worse than would be expected on the basis of background information. Thus, of the 150 local programmes selected in the first round of strategic sampling, 100 will be further selected for longitudinal follow-up on the basis of three sets of factors. Two of these three factors were used in the first stage of strategic sampling, as indicated above, but the third is importantly different.

1. Geographic, demographic, and ethnic characteristics of Sure Start communities
2. Intervention strategies of Sure Start local programmes
3. Cross-sectional data indicating whether children and families in specific Sure Start communities are being positively affected by Sure Start.

After ensuring that we have a good representation of communities on the basis of inputs (i.e. factors (1) and (2) above), communities that systematically vary from excellent to poor in terms of detected effects of Sure Start are selected. This selection on the basis of outputs is made possible by establishing 'community effects' within a multi-level modelling analysis (Goldstein, 1995). Hence, the second round of strategic sampling is based on a combination of inputs and (actual-not inferred) outputs.

The infants and families from the 100 Sure Start local programme areas selected in the second round of strategic sampling and studied cross-sectionally will be followed up when children are 3- and 5-years of age (and beyond). In other words, children and families first studied as part of a cross-sectional investigation become participants in a long-term longitudinal investigation as part of our strategically staged, birth-cohort longitudinal study.

Data gathered in the Millennium Cohort Study from deprived families not receiving Sure Start on child, family, and community functioning when children are infants, 3-year-olds, 5-year-olds, and beyond will function as control data for the purpose of longitudinally evaluating the effects of Sure Start local programmes in the 100 longitudinally studied Sure Start communities.

Data collected in the cross-sectional and longitudinal studies cover the following range of topics:

- Family demographics, e.g. parental occupation, education, family structure
- Family dynamics, e.g. parental relationship quality
- Parental physical and mental health
- Parenting, e.g. discipline, parent-child relationship
- Home environment, e.g. activities of children, learning experiences
- Child care – type, timing and quantity, both historical and current
- Community characteristics, derived from parent report, observations and the administrative Local Context Analysis data

- Child cognitive and language development, from the British Ability Scales (Elliot, Smith, & McCulloch, 1996)
- Child social and behavioural development from the Strengths and Difficulties Questionnaire (Goodman, 2001) and other parent and teacher report measures
- Child health, e.g. immunisations, injuries, accidents, illnesses, hospitalisations
- Services utilised by children and parents
- Economic context; employment, income, debt, etc.

These data are collected through parental interview, parental and teacher questionnaires, systematic observation, and individual child assessments. Where possible, the data are collected in a manner similar to that of the Millennium Cohort Study in order to facilitate comparisons. One problem that arises with regard to the use of the Millennium Cohort Study as a form of control or comparison group is the possibility of historical change in the period between the MCS and NESS data collections, as the MCS was launched a year before the Impact study. For example, should a major recession or other major social change occur in this intervening period then the societal context within which data are collected changes, and the comparability of data for purposes of illuminating effects of Sure Start local programmes becomes problematic. Ideally, both sets of data would be collected synchronously but pragmatic considerations preclude this. While Sure Start versus non-Sure Start comparisons are one way of evaluating the effects of Sure Start, the great diversity between Sure Start local programmes provides another method. In some respects the diversity of Sure Start programmes is a problem: it makes the questions of what is Sure Start and what is the effectiveness of Sure Start much more complicated to investigate. However, it does provide the opportunity to study a variety of intervention strategies and characteristics. The strategy of analysing impact data through the use of multi-level modelling allows for the establishment of specific local programme effects. Hence, those local programmes that are particularly effective as compared with other local programmes can be identified and their distinguishing characteristics noted. Such within Sure Start comparisons may prove to be most useful in answering questions of effectiveness, given the diversity between Sure Start local programmes.

Local context analysis

The Sure Start local programme areas were designed to be small, so that people could access services on foot, and would 'make sense' to the local community. This concept drew directly on the research papers produced for the Cross Departmental Review of Provision for young children, particularly a review of the effectiveness of early interventions that noted that interventions with high levels of community involvement were able to draw upon and stimulate the shared concern of local people for each other and for their environment. (Oliver et al., 1998). It also reflected Government concern with 'pockets of intense deprivation where the problems of unemployment and crime are acute and hopelessly tangled up

with poor health, housing and education' (Social Exclusion Unit, 1998). However, obtaining definitions of communities based on the perceptions of residents is not a straightforward process (Chaskin, 1997).

The local groups of key practitioners and professionals from both the voluntary and statutory sectors responsible for preparing the local Sure Start development plans defined their Sure Start area by drawing the boundaries on a map. While this method of area selection makes very good sense from a service provision perspective, it means that most Sure Start local programme areas do not conform to standard administrative boundaries such as electoral wards or NHS districts. Data about areas are generally organised by ward, district, local authority or similar standardised boundaries (Noble et al., 1999). Thus the task of collecting information about the idiosyncratic Sure Start areas becomes especially complex. To cope with the challenges posed by this complexity, geographical information systems have been used to obtain data from relevant agencies using lists of post-codes and in some cases digitised boundaries that can be matched to data systems.

With a variety of such data in hand, the aims of the Local Context Analysis (LCA) study are to address the following questions:

1. Are Sure Start local programme areas well located to reach the greatest concentrations of children under 4 years, at risk of poor outcomes?
2. What do Sure Start areas look like compared to other disadvantaged areas without a Sure Start local programme and compared to national norms?
3. Are there differences between Sure Start local programme areas, and how will these differences influence the effectiveness of Sure Start?
4. How have Sure Start local programme areas changed over time, in relation to the implementation of Sure Start?

The LCA module provides the backdrop against which Sure Start local programmes are put into place and is designed to document change over time in the Sure Start and control communities. While Sure Start local programme areas have similarities, it may be the differences between them that prove most relevant when it comes to understanding the (differential) effects of Sure Start. Detailed information about community characteristics in other evaluation studies has proved pivotal in understanding variability in outcomes between seemingly similar schemes (Barnes, 1997; Coulton, 1995). The information provided by the LCA is integral to the design of the whole evaluation. It enables the data concerning infant and child development and family functioning collected in Sure Start programme areas to be interpreted in relation to their environmental context. Thus, for example, the progress of children in a mixed social class area (relative to control children) can be compared with the progress of children in an area with more consistent deprivation. Parenting in families living in neighbourhoods characterised by high levels of crime can be compared with programme areas that have less recorded crime. The variability in programme area characteristics will also be used to

interpret information about the extent to which programmes have been able to implement the stated goals and aims of Sure Start. For instance, those in areas with low levels of employment opportunities may be less (or more) successful in engaging parents in management board activities than programmes in areas with high levels of employment opportunities.

The LCA data will also be important, both in the short term and long term, in understanding the cost effectiveness of the initiative. It is possible that developmental progress and readiness for school are associated with higher levels of achievement only in programme areas where the overall level of achievement reaches a certain level in Key Stage results, for example. Additionally, or alternatively, it may be that greater employability and lower levels of delinquency are only achieved in areas with the lowest rates of recorded crime.

Not only will the LCA data be used in the interpretation of evidence from other modules, but also it will provide information about change over time in the communities themselves and thus illuminate how Sure Start local programme services affect communities – at the level of community. Constructs used in the LCA describe the population of programme areas, the extent of deprivation, the availability of employment and the proportion of the population who are in employment, the extent of crime and other disorder, the general health of adults and children in the community, and the developmental progress and academic achievement of children living in programme areas. Information on parenting and on the extent to which families experience parenting problems is also collected, based on information from Social Services departments. Provision of local services relevant to families with young children such as libraries, GP surgeries and infant schools are being documented, as well as their accessibility for families living in the Sure Start programme areas, on the basis of average distance. Repeated collection of such data over time will provide insight into how Sure Start communities change.

Sure Start areas are all chosen to be within the 20% most deprived in the country. Hence the variability in area characteristics will be more restricted than that of the country as a whole. The question arises as to whether the restriction of area diversity is such as to limit the influence of area characteristics upon evaluation results. Early results from the local context analysis reveal that this is unlikely to be so, in that the data reveal considerable diversity in area characteristics amongst Sure Start areas, even though they are well within the 20% most deprived in the country. Also, the diversity amongst Sure Start areas appears to be linked to particular clusters of area characteristics, providing a possible basis for a typology of deprived areas (Barnes et al., 2003).

The cost-effectiveness module

Cost-benefit analyses of intervention programmes, whilst rare, have been very influential politically, as in the case of the High/Scope Perry Preschool Project (Barnett, 1996; Karoly et al., 1998) and the Abecedarian

project (Masse & Barnett, 2002). They are also critical for informed decisions by policymakers. Hence a cost-effectiveness module was seen as essential for NESS. This module is considering the relationship between the costs of Sure Start local programmes and other services for children and families in Sure Start programme areas and the outcomes achieved, both intermediate and longer-term.

The starting point for the analysis is data on service costs derived from information supplied by each local programme to the Sure Start Unit. In principle, the data on the costs of Sure Start services themselves are relatively straightforward. In addition, many Sure Start local programmes receive resources from other government initiatives (such as the New Deal for Communities), from the European Social Fund, and from charitable trusts and business sponsorship. Many also receive services such as finance, human resources and accommodation support from their parent organisations (e.g. local health authority) for which they do not have to pay, but which represent additional resources going into the programme. The implementation module national survey collects information about these resources for all local programmes. In addition, the implementation case studies are collecting information about the resources provided by mainstream services. In sum, every effort is being made to chronicle the total set of economic resources devoted to local Sure Start programmes, not only those directly provided by each programme's local Sure Start grant.

The evaluation will also consider the extent to which costs are understated because resources have been diverted from other areas into Sure Start areas. This is particularly relevant in skill shortage occupations, where recruitment by Sure Start may result in unfilled vacancies elsewhere.

The second stage of the cost-effectiveness evaluation is to relate the cost information gathered during the evaluation of the cost-effectiveness of the implementation and relate it to the short-term outcomes that emerge from the impact module. This is relatively complex with an intervention such as Sure Start, since there are multiple outcomes (e.g. see Feinstein, 2000). Nevertheless, it should be possible to have some indicators of the costs of achieving particular outcomes through different routes, taking account of the local community context, the baselines in terms of the scale of the challenge, and the level of resources already being deployed.

The cost-benefit analysis will use a human capital framework to consider the potential range of benefits from the intervention (see Gramlich, 1990; Heckman, 1998). Sure Start has to be seen as an investment in the development of young children and their families. As with any human capital investment, the costs are incurred early on, but the benefits (hopefully) flow throughout an individual's lifetime (Gregg & Machin, 1998). There are four potential groups of beneficiaries: the children, their families, the local community and the wider society. The potential benefits relate to educational outcomes, health, behaviour, employment and earnings and the economic and social well-being of the area. There will be some benefits that may occur early (for example, on children's health or parents' employment), but US experience (e.g.

Karoly et al., 1998) shows that the total benefits from early childhood interventions generally only exceed the costs when the children have reached their late teens or early 20s (indeed if some go on to higher education it will take longer still).

Summary

In summary, the National Evaluation of Sure Start local programmes in England is a multi-faceted research evaluation designed to accurately describe (a) what local Sure Start programmes look like in terms of their organisation and service delivery; (b) how children and families in these programme areas are functioning so that the effect of Sure Start on child and family development can be estimated; (c) what local Sure Start communities look like in comparison to each other and the rest of England and whether, and how, they change over time; and (d) what the economic costs of this major effort to promote the well-being of children under 4 years and their families are and to what extent the return on investment covers the costs of the programme over the long term. Not only because of its comprehensive nature, but because it takes time for children, families and communities to develop and change, it will take time before NESS fully realises its promise to (a) describe what is going on in Sure Start local programmes, (b) assess its impact on children, families and communities, and (c) determine whether the investment represents good value for money.

For policy oriented work, timing and context, geographical, social and political, are critical. One example is that the circumstances of the roll-out of Sure Start local programmes precluded the possibility of a randomised control trial as the basis for the evaluation of impact, therefore requiring a creative approach to research design. Also the programmes are prone to rapid change and development, which can only be captured with the right timing and extended investigation. The research study has therefore had to be adjusted as it became apparent that setting up Sure Start programmes had taken rather longer than expected. This suggests that besides the services offered to children and families, and the quality of those services, the way they are established, administered and run is significant. With changing social, economic and political circumstances the context of Sure Start local programmes will change over time. Such changes may affect the evaluation process. For example, if Sure Start services alter in their availability to children and families chosen to be in Sure Start or non-Sure Start groups, either through the reduction or expansion of the Sure Start programme, the evaluation of the effects of such services becomes more problematic. Indeed, the inclusion of Children's Centres under the umbrella of Sure Start is designed to provide some services to children and families who do not live in Sure Start local programme areas. Similarly, should there be major structural changes in the communities being studied, then this could overwhelm any effects due to the programme. There are also indications to suggest that where there were innovative approaches to services for young children before Sure Start began, progress towards establishing the intervention has been most rapid. The nature and impact of interventions depends

critically upon the nature of communities to which they are applied. Hence, the Local Context Analysis and an awareness of more general political and social change is crucial to understanding how Sure Start works.

References

- Ball, M., & the NESS Research Team (2002). *Getting Sure Start started. National evaluation of Sure Start*. Report No. 2, July. Available at <http://www.ness.bbk.ac.uk/documents/Activities/Implementation/18523GettingStarted.pdf>
- Barnes, J. (1997). Measurement issues: Methods of measuring community characteristics. *Child Psychology and Psychiatry Review*, 2, 163–169.
- Barnes, J., Broomfield, K., Frost, M., Harper, G., Knowles, J., Leyland, A., McLeod, A., & the NESS Research Team (2003). *Characteristics of Sure Start local programme areas. Rounds 1 to 4. National evaluation of Sure Start*. Report No. 3, June. Available at: <http://www.ness.bbk.ac.uk/documents/Activities/LCA/FinalLCAReporJuly2003.pdf>
- Barnett, W. S. (1995). Long term effects of early childhood programmes on cognitive and school outcomes. *The Future of Children: Long Term Outcomes of Early Childhood Programmes*, 5, 94–114.
- Barnett, W. S. (1996). *Lives in the balance: Age 27 benefit-cost analysis of the High/Scope Perry Preschool Program*. Monographs of the High/Scope Educational Research Foundation, No 11.
- Chaskin, R. J. (1997). Perspectives on neighborhood and community: A review of the literature. *Social Service Review*, 521–547.
- Coulton, C. (1995). Using community-level indicators of children's well-being in comprehensive community initiatives. In J. Connell, A. Kubisch, L. Schorr & C. Weiss (Eds.), *New approaches to evaluating community initiatives* (pp. 173–199). Washington, DC: The Aspen Institute.
- Cutler, T., & Waine, B. (2000). Managerialism reformed? New Labour and public sector management. *Social Policy and Administration*, 34, 318–332.
- Feinstein, L. (2000). *The relative economic importance of academic, psychological and behavioural attributes developed in childhood*. Centre for Economic Performance Discussion Paper No. 443. London: London School of Economics.
- Elliot, C., with Smith, P., & McCulloch, K. (1996). *British Ability Scales—BAS II* (2nd ed.). Windsor: NFER/Nelson.
- Ghate, D. (2001). Community-based evaluations in the UK: Scientific concerns and practical constraints. *Children and Society*, 15, 23–33.
- Glass, N. (2003). *Sure Start five years on: The early years of an early years programme*. http://www.city.ac.uk/barts/events/2003/06_jun/presentation_files/frame.htm
- Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). London: Arnold.
- Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 1337–1345.
- Gramlich, E. M. (1990). *A guide to benefit-cost analysis* (2nd ed.). Prospect Heights, IL: Waveland Press.
- Gregg, P., & Machin, S. (1998). *Child disadvantage and success or failure in the youth labour market*. Centre for Economic Performance Discussion Paper 397. London: London School of Economics.
- Heckman, J. (1998). Human capital investment policy. *Fiscal Studies*, 19, 103–119.
- Kanter, R. M. (1983). *The change masters*. New York: Simon & Schuster.
- Karoly, L. A., Greenwood, P. W., Everingham, S. S., Houbi, J., Kilburn, M. R., Rydell, C. P., Sanders, M., & Chiesa, J.

- (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions*. Santa Monica, CA: RAND.
- Kresh, E. (1998). The effects of Head Start: What do we know? *National Head Start Association Research Quarterly*, 1, 112–123.
- Love, J., Kisker, E. E., Ross, C. M., Schochet, P. Z., Brooks-Gunn, J., Paulsell, D., Boller, K., Constantine, J., Vogel, C., Fuligni, A. S., Brady-Smith, C. (2002). *Making a difference in the lives of infants and toddlers and their families: The impacts of Early Head Start*. Volume 1: Final Technical Report. Princeton, NJ: Mathematica Policy Research Inc. Available at <http://www.mathematica-mpr.com/PDFs/ehsfinalvol1.pdf>
- Masse, L. N. & Barnett W. S. (2002). *A benefit cost analysis of the Abecedarian Early Childhood Intervention*. New Brunswick, NJ: National Institute for Early Education Research. Available at <http://nieer.org/resources/research/AbecedarianStudy.pdf>
- Mrazek, P. J., & Brown, C. H. (2002). An evidence-based literature review regarding outcomes in psychosocial prevention and early intervention for young children. In C. C. Russell (Ed.), *The state of knowledge about prevention/early intervention* (pp. 42–144). Toronto, Canada: Invest in Kids Foundation.
- Newman, J. (2001). *Modernising governance: New Labour, policy and society*. London: Sage.
- Noble, M., Penhale, B., Smith, G., Wright, G., & Owen, T. (1999). *Index of multiple deprivation 1999 review. Final consultation. Report for formal consultation. Stage 1: Domains and indicators*. Oxford: University of Oxford, Social Disadvantage Research Group.
- Olds, D. L. (1997). The Prenatal Early Infancy Project: Preventing child abuse in the context of promoting maternal and child health. In D. A. Wolfe, R. J. McMahon & R. DeV. Peters (Eds.), *Child abuse: New directions in prevention and treatment across the lifespan* (pp. 130–156). Thousand Oaks, CA: Sage.
- Oliver, C., Smith, M., & Barker, S. (1998). *Effectiveness of early interventions*. In *Treasury Supporting Papers, Cross Departmental Review of Provision for Young Children*. London: HM Treasury.
- Patton, M. Q. (1997). *Utilization-focussed evaluation*. New York: Sage.
- Peters, R. DeV., Petrunka, K., & Arnold, R. (2003). The better beginnings, better futures project: A universal comprehensive, community-based prevention approach for primary school children and their families. *Journal of Child Psychology and Psychiatry*, 32, 215–227.
- Pressman, J., & Wildavsky, A. (1984). *Implementation: How great expectations in Washington are dashed in Oakland etc.* Berkeley, CA: University of California Press.
- Ramey, C. T., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. L. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science*, 4, 2–14.
- Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). *Significant benefits: The High/Scope Perry Preschool Study through age 27*. Ypsilanti, MI: High/Scope Press.
- Smith, K., & Joshi, H. (2002). The Millennium Cohort Study. *Population Trends*, 107, 30–34.
- Social Exclusion Unit (1998). *Bringing Britain together: A national strategy for neighbourhood renewal*. London: Office of the Deputy Prime Minister.
- Tunstall, J., Allnock, D., Meadows, P., McLeod, A., & the NESS Research Team (2002). *Early experiences of implementing Sure Start. National evaluation of Sure Start*. Report No. 1, June. Available at <http://www.ness.bbk.ac.uk/documents/Activities/Implementation/18523EarlyExperiences.pdf>
- Webster-Stratton, C., & Taylor, T. (2001). Nipping early risk factors in the bud: Preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (0–8 years). *Prevention Science*, 2, 165–192.