

NATIONAL EVALUATION OF SURE START – METHODOLOGY REPORT **EXECUTIVE SUMMARY**

Sure Start represents a unique approach to early intervention for children 0-4, their families, and communities. Rather than providing a specific service, the Sure Start initiative represents an effort to *change existing services*. This is to be achieved by reshaping, enhancing, adding value, and by increasing co-ordination. The Sure Start model presumes that:



In light of this model, three core questions need to guide the overall evaluation of Sure Start.

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?)

The National Evaluation of Sure Start addresses these questions through five components: (1) implementation evaluation, (2) impact evaluation, (3) local community context analysis, (4) cost-benefit analysis and (5) support for local evaluations.

The Implementation Evaluation

The overall purpose of the implementation evaluation we propose is to illuminate the contents of the Sure Start "black box"; it has three primary functions:

1. To provide data on key aspects of design, policy, practice, style and development in each local Sure Start programmes.
2. To produce qualitative information on services as a framework for the impact study.
3. To provide data on programme objectives and facilitate the quantification of inputs for the analysis of cost effectiveness.

The Impact Evaluation

The impact evaluation is designed to examine effects of Sure Start 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

The Local Context Analysis

In partnership with the South East Regional Research Laboratory and the Office of National Statistics, we will examine neighborhood-level data on 260 Sure Start communities longitudinally, and on 50 Sure-Start-to-be control communities cross-sectionally. This illuminates the community context in which Sure Start programmes operate and contributes to the impact study's analysis of the effects of Sure Start on communities. The data collected in the local context analysis will:

1. Provide essential contextualization for implementation, impact and cost-benefit evaluation and also useful information for local evaluation support.
2. Provide a cross-sectional and longitudinal study of communities to be used in the impact evaluation.
3. Enable the evaluation of community variation in control communities.

The Cost-effectiveness Evaluation

The cost-effectiveness evaluation aims to answer three core questions:

1. What is the true cost of a particular initiative?
2. Did the outcomes justify the investment of resources?
3. Was this the most efficient way to achieve the identified outcomes?

Cost effectiveness should only consider the effect of the intervention on the programme's own objectives. Cost-benefit analysis, however, should include all effects, both positive and negative, direct, indirect and spill over, anticipated and unanticipated.

Support for Local Evaluations

Support for Local Evaluations is provided by the national evaluation in terms of:

- Expert guidance and advice.
- Training to programmes on evaluation.
- Information on how the local and national activities interface.
- Strategies for coordinating national and local information collection.
- Identify key themes for local, small scale and in-depth, research projects.
- Provide some supervision for these small-scale projects.

Full methodology reports for the five modules are available up to April 2001. Summaries of these reports and a provisional schedule of outputs for the project follow.

Summary of the Implementation Methodology

This module focuses on both formative evaluation questions, such as ‘what are the main elements of the programme? What are the main characteristics of the programme? – and summative ones such as ‘were deviations from the original design great enough to constitute a different model from what was originally intended?’ Its design takes account of the current policy context, (including emphases on evidence-based practice; the value of preventive interventions) as well as other relevant evaluation approaches such as Early Head Start in the US.

Aims

The overall aim is to produce a comprehensive picture of the first 260 Sure Start programmes in terms of process and components, which is capable of linking programme activity to outcomes for individual users and communities. Specific objectives include:

- To describe the full range of approaches to service design and delivery.
- To produce a systematic analysis of local programmes.
- To produce alternative assessments of programme quality.
- To provide information for sampling programmes for the impact evaluation.
- To assess the extent to which implementation conforms with programme specifications.

Method

The study, which incorporates quantitative and qualitative methods, has three components:

- **A National Survey of all 260 Sure Start programmes.**
- **An in-depth study of 26 programmes.**
- **A series of themed evaluations.**

Each contributes data along **six dimensions of local programme design and delivery**:

1. Management and coordination.
2. Access for families to Sure Start provision and services.
3. Community involvement.
4. Allocation of resources.
5. Quantity of services provided by the programme.
6. Quality of services provided by the programme.

The National Survey - has two elements: (i) **a questionnaire** administered by email to all 260 programmes at annual intervals at least three times; (ii) **document search** of all relevant documents produced by local programmes and by partner agencies.

Data analysis will incorporate comparison across programmes; cluster analysis of classification data; and a series of rating scale developed for each of the six programme dimensions and for ‘joined-up-ness’. The latter follows a similar model to that employed by Love et al. in the Early Head Start Evaluation in the US, who developed rating scales for assessing implementation in key

areas. This part of the analysis is intended to produce an insight into the potential quality of programme implementation, so that 150 programmes can be selected for inclusion in the cross sectional investigation to be carried out by the impact evaluation.

Case Studies of 26 Local Sure Start Programmes - will enable us to explore, in depth, the six programme dimensions outlined above; as well as undertake a comprehensive service audit and collect data that could not be accessed by questionnaire. Methods of data collection will include face-to-face interviews with parents; community members and staff of programmes and partner/local agencies; as well as structured observation. The main elements in the Case Study process will be repeated twice at 18-month intervals to examine change over time.

Thematic Evaluations - will examine issues that arise in all or a substantial proportion of local programmes. They will enable the exploration of issues of over-arching concern that cannot be adequately addressed within an individual programme. The selection of themes will be data-driven by the survey and case study modules; and will be negotiated with the Sure Start Unit. Examples of likely themes include Getting Sure Start Going; Minority Ethnic Communities and Sure Start; Special Needs Provision; Sure Start in rural areas; Child Protection and Sure Start. Methods will include secondary analysis, supplementary data collection through interview and other survey techniques.

Timetable

The timing of selected key stages in the Implementation Study is as follows:

- June/July 2001 - Piloting of the national survey questionnaire in 9 randomly selected areas.
- August - September 2001 - First application of national survey questionnaire to programmes in rounds 1 & 2.
- January 2002 - First application of national survey questionnaire to programmes in rounds 3 & 4.
- July 2003/January 2004 (depending on round) - Final application of the survey questionnaire.
- October 2001 - Selection of 16 case-study areas from round 1 & 2 programmes.
- February 2002 - Selection of further 10 case-study areas from Round 3 & 4 programmes.

Summary of the Impact Methodology

The impact evaluation is designed to examine effects of Sure Start 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. We initiate our impact evaluation 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). This will enable us to undertake a very early evaluation of the impact of Sure Start on infants, two-year olds, and four-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 realized by longitudinally following families with infants in the cross-sectional phase of study when children are 3, 5, and 7 (and beyond). The longitudinally studied children and families will be drawn from the 150 cross-sectionally studied Sure Start communities and compared with each other and similar children and families studied as part of the Millennium Birth Cohort Study

1. 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 communities 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 Sure Start, and (c) their initial progress in implementing Sure Start (as determined by the implementation evaluation). More specifically, we will strategically select for inclusion in the cross-sectional phase of inquiry Sure Start programmes that *systematically vary* in their very early success in implementing their programmes as well as in their programming strategies. This will ensure the inclusion of programmes likely to be effective and increase the power to detect program effects. One hundred and twenty (120) families with infants, and 3-year olds, will be studied in each of 150 carefully selected Sure Start communities and 50 randomly selected Sure Start-to-be (i.e., control) communities. We will thus have data from communities, parents and children of *diverse* ages, as early as the end of the third year of evaluation, and will be in a strong position early in the evaluation to detect effects of Sure Start.
2. 8000 of the infants and families from the 150 Sure Start communities 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, longitudinal study.
3. Data gathered in economically deprived communities as part of the Millennium Birth Cohort on child, family, and community functioning when children are infants, three-year olds, five-year olds, and beyond will function as control data for the purpose of longitudinally evaluating the effects of Sure Start in the 150 longitudinally studied Sure Start communities.

The following table provides a summary of the categories of data to be collected in the impact evaluation.

Individual/Family Data Collections

A. Family Demographics

1. 1. Family and household composition.
2. Income and source and debt burden.
3. Employment: mother, father, other.
4. Education: mother, father, other.
5. Marital status.
6. Ethnicity and religion.

B. Family Dynamics

7. Intimate relationship quality.
8. Domestic violence.
9. Parenting stress.

C. Parental Health*

1. Mental health: Including depression & self-esteem.
2. Physical health.
3. Health behavior: Smoking, drinking, drug use.

D. Parenting/Home Environment

1. Home Environment: H.O.M.E. & Home Learning Environment (EPPE).
2. Parent-Child Relationship.
3. Discipline.
3. Family organization.
4. Father contact and involvement.
5. Parent literacy.

E. Child Care

1. History: when, how much, who, where.

F. Community

1. Local environment quality (checklist & observation, Barnes, 1997).
2. Housing adversity.
3. "Belonging" to neighborhood (Barnes, 1997).
4. Density of acquaintanceship, local social ties.
5. Perceptions of local danger/disorder (Barnes, 1997).

G. Child Development *

1. Infancy
 - a. Birth weight & gestational age.
 - b. Pregnancy/delivery - smoking, screening, antenatal classes, complications.
 - c. Prenatal and infant health.
 - d. Type of feeding and duration of breast feeding.
 - e. Child's developmental status, including delay.
 - f. Child temperament

2. Three years

- a. Social/behavioral/emotional development: Observation & parental rating.
- b. Cognitive/language development.
- c. Health: Feeding, physical dev., immunizations, congenital problems.
- d. Child temperament.

3. Five years

- a. Social/behavioral/emotional development and mental health.
- b. Cognitive development.
- c. Language development & emergent literacy.
- d. Health: Physical dev., immunizations, diet, special needs, hospitalization.
- e. Child temperament.

H. Services Utilized by Children and Parents

1. Visits to GP over recent period and who for.
2. Use of A&E and specialist services, e.g., mental health, including inpatient.
3. Contacts with relevant agencies, including voluntary sector.
4. Views about volume, quality, and appropriateness of services used.
5. Whether service providers dealt with problems or referred them.
6. Speed of response to a problem, and follow-up continuing care.

I. Economic Context

1. Current/recent training/education activities of parents.
2. Benefits family is receiving.

* A more extensive list of child and parent health measures may be obtained from health record systems (e.g., Hospital Episode Statistics, Incapacity Benefit, NHS Trusts, Disability Living Allowance), presuming that parents will give permission so that we can identify individual parents and children in the aggregate record archives.

Summary of Local Context Analysis Methodology

The local context analysis will contribute to answering the following core questions:

What are the socio-demographic characteristics of Sure Start users/communities?

How can the Sure Start neighborhoods be compared to further the understanding of the effectiveness of Sure Start?

Do these profiles and comparisons change over time and, if so, can these changes be related to variation in implementation of Sure Start.

What is the relationship between Sure Start neighborhoods and national or other aggregate level benchmarks?

What are the effects of Sure Start at the community level, on services, levels of disadvantage and disorder, and the general health and well being of local residents?

What is the relationship between Sure Start Programme areas and those for other area based initiatives.

Types of data

Existing administrative databases

The local context analysis will generate area-specific information by applying the specific area definitions of each Programme to a variety of local and national databases.

1. Observations

Interviewers conducting the impact evaluation interviews will complete a short observation schedule and regional fieldworkers will complete more detailed observations in all the programme areas.

2. Questionnaires & Surveys Questionnaires and interviews enable the investigation of the phenomenological processes that intervene between community structure and child or family outcomes.

Methodological Issues

1. Identification of Programme Boundaries

Data will, in most cases, be expressed according to the Sure Start Programme areas. All programmes supply maps of the areas they are proposing to cover. Translating these boundaries into digital form allows the identification of the unit postcodes that fall within them. These postcode lists then form the basis for identifying families with young children from DSS records as well as collecting other indicators linked to the receipt of benefits or the monitoring of low birth weight children.

A number of factors affect the precision of this process. The maps supplied by are of variable quality with boundaries often drawn 'freehand' leaving ambiguity as to their exact location. In addition, these boundaries will cut across some unit postcodes. Although the latest, most accurate Post Code Address File will be used it is inevitable that small parts of postcodes will be either

omitted or wrongly included. On average each postcode contains about 15 households so that these effects are likely to be small.

Ordnance Survey and lists of constituent postcodes provided. Rounds three and four programmes will supply outline maps of their programme areas that will be translated into a digital form using a combination of flat tablet digitizing and the registering of scanned images.

2. Identification of labor market areas in which programmes fall

Identification of the labor market areas in which residents of Programme Areas are likely to search for work will always pose problems particularly within large cities. Our approach is to use the local knowledge of the programme teams to indicate the broad areas that form effective labor markets for their residents and to translate these into aggregations of wards, thus enabling use of the Annual Employment Surveys to monitor local employment change in relation to regional and national benchmarks.

3. Profiling communities

We will examine variability in the various indicators for the wave 1 and 2 programmes. We will develop strategies for looking at variability both within a programme area and between programme areas. Thus the groupings between programme areas will be based on rankings for each or the main dimensions under consideration. Programmes will be ranked by domain and a total index will also be constructed.

Specific Information to be collected

1. Local programmes will be asked about the availability of specific services and their location, so that access for families with children 0-3 can be mapped.
2. While the Index of Multiple Deprivation is ageing rapidly and is not specifically designed to relate to young children, it provides one way of relating conditions within the Sure Start areas to the rest of the country. In particular we will include information on the local concentration of poverty by examining the percentage of poverty in the ward containing the Sure Start Programme and the contiguous wards.
3. The 1991 and 2001 census tabulations will be particularly useful in establishing household/family structures, residents, occupational structure, ethnic composition, local living conditions, demographic profiles and estimates of the flux of migration passing through the areas.
4. The Office for National Statistics (ONS) will provide data to the National Evaluation of Sure Start from two separate sources of data: birth and death registrations, made by members of the public to Registrars of Births and Deaths, collected by ONS; and local child health systems. The data extract from Child Health Systems will include socio-demographic details of the mother, prenatal data items, measures of disability and associated conditions at around age two. The data items to be included are not routinely available on a national basis.
5. In conjunction with Sure Start's contract with the DSS to provide information on children in workless households, additional data will be obtained regarding adults receiving a number of benefits that indicate disabling physical or mental health problems.

6. Using NOMIS data, employment levels will be measured for full time and part time employment for males and for females and will be classified by the nine broad industry groups of the 1992 Standard Industrial Classification. With respect to employment and unemployment the programme areas themselves are generally too small to be regarded as local labor markets but in many cases the Travel to Work Areas in which they are contained are likely to be too large to represent realistically the range of jobs being accessed by local residents.
7. With the knowledge that crime data are not uniformly available, the following information will be sought: rate of burglary and car theft; rate of drug arrests; rate of violent crime; rate of juvenile contact with police.
8. Information from local primary and secondary schools, and nationally held data from the DfEE: achievement test scores; detention and exclusion rates; unauthorized absence rates; children statemented with special needs; participation of young people in full or part-time education and post-16 training.
9. Local Social Services will be asked to provide information on numbers of children on the at-risk register and re-referrals to the at-risk register among residents of the Sure Start programme areas. We will need to clarify whether children on the register, or reregistered, come from families living within the Sure Start programme area or whether they have been moved into the area from families in other neighborhoods.
10. Local surveys of parents and other residents, conducted either by the local programmes or as part of the impact evaluation, will include questions about perceptions of crime and danger in neighborhoods and observations of communities will indicate the level of social and physical incivilities.
11. Impact evaluation interviews will include psychological attachment to the community, the level of perceived danger and disorder, and the level of social networks and social capital.

Summary of Cost-Effectiveness Methodology

The cost-effectiveness evaluation will consider the relationship between the costs of Sure Start and other services for children and families in Sure Start areas and the outcomes achieved, both intermediate and longer-term. During the initial period, this will focus on the cost-effectiveness of the implementation, but as outcome information becomes available, it will move on to the cost-effectiveness of the outcomes achieved and also to the wider cost-benefit analysis.

Cost-Effectiveness of the Implementation

Individual Sure Start programmes have targets to achieve. They also have other outputs that they are choosing to deliver according to local needs and priorities. The cost effectiveness evaluation will consider the efficiency and effectiveness of the use of resources in achieving these targets and other intermediate objectives. These intermediate targets have been chosen because they are believed to be associated with improvements in longer-term outcomes for children. Those outcomes themselves will be assessed as part of the cost-benefit analysis, which will be based on the impact of the programme (see below). At this stage, the evaluation will be considering the extent to which the resources available to support families and young children are being used to achieve the targets that have been set either locally or nationally.

The starting point for the analysis will be data on costs derived from information supplied by each programme to the Sure Start Unit. This information is supposed to include the costs of other services already being provided in the area (in order both to measure deadweight and to recognize that the achievement of individual targets may be influenced by the level of resources which is already being provided to support existing mainstream services). However, this information may be of lesser quality than information about the deployment of the Sure Start grant, so it will need to be double checked and, if necessary, independently verified. Local evaluators are likely to be involved in this.

Programmes themselves assign the costs they incur to individual targets. Although this is unlikely to be fully accurate, it should be sufficiently comparable to provide an indication of the range of resources that go towards achieving different targets using different delivery mechanisms. It is important to stress that the long-term outcomes may vary according to the delivery mechanism. For example, some areas are using health visitors to do home visits, others are using nursery nurses and others are using specially recruited and trained members of the community. Although some of these methods may produce lower unit costs achieved than others, until we have outcome indicators from the impact evaluation, we cannot be certain that the costs of achieving a particular outcome are higher or lower.

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

Cost-Effectiveness of the Impact

The second stage is therefore to draw on the information gathered during the evaluation of the cost-effectiveness of the implementation and relate that to the short-term outcomes. This is relatively complex with an intervention such as Sure Start, since there are multiple outcomes and the assignment of costs is likely to be problematic. 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.

Cost-Benefit Analysis

The cost-benefit analysis will use a human capital framework to consider the potential range of benefits from the intervention. Sure Start has to be seen as an investment in the development of young children and their families. Like any human capital investment, the costs are incurred early on, but the benefits flow through throughout an individual's lifetime. There will be some benefits that will occur early (for example on children's health or parents' employment), but US experience shows that breakeven point tends to come in the late teens or early twenties or even later. Because of the likely time scale the calculation of benefits is likely to be sensitive to the discount rate used. Our analysis will use the rate prescribed by the Treasury (currently 6 per cent), but most US evaluations have used lower rates, which produces higher benefits.

Broadly, there are four groups of beneficiaries:

- Children
- Families
- The local community
- The wider society (both as members of the community and as taxpayers)

The benefits to which monetary values can be attached come under several headings:

- Educational outcomes and skill development
- Health (physical and psychological) outcomes
- Behavioral outcomes (for example involvement with substance use or delinquency)
- Employment outcomes (including earnings)
- The economic and social well being of the area

There are likely to be benefits over the short, medium and long term. There may also be some additional costs (for example from some young people remaining in full-time education longer than they would have done in the absence of Sure Start or from higher levels of referral to specialist health services).

The total net benefits to society as well as the different groups identified above will be estimated. The distribution of benefits is important, as some gains for one group (for example tax savings) will have counterpart losses for others (for example lower receipt of social security benefits).

Summary of Support to Local Programmes Methodology

The Support to local programmes will begin with a mapping exercise that will:

- Identify the contact information about the local evaluation plans
- Their relationship to the local programmes
- The support needs of the programmes and their evaluators

This is being done by contacting every programme and obtaining as much information as possible. A questionnaire will be sent to programmes in June 2001.

The overall structure of the support programme will consist of 4 components:

- Support provided by 6 regional units. The units will provide the frontline support to programmes in their region. Each region will cover approximately 45 programmes. The units are expected to be established in the summer and will be immediately available to offer support.
- A website which will contain information about NESS, local evaluations, and evaluations of other related initiatives, advice and opportunities to communicate with NESS and other local evaluators. The website will be launched in summer this year.
- A Central office, which will co-ordinate the regional offices and provide a link into the administration of the national evaluation. This will facilitate links with the national implementation, local community context, cost effectiveness and the impact evaluations. It will also provide information about a range of methodologies, specific instruments, and the evaluations or a number of other area based initiatives.
- National and regional workshops which will provide information and exchange between the National and Local evaluations. The first national workshops will be held in Leicester and London in early July 2001. Local workshops will concentrate on developing regional themes for small scale, in-depth research projects, and then providing assistance in conducting the projects.

One issue already identified is the sharing of information between NESS and local programmes about impact, implementation and local context findings that may be relevant to local evaluations. Careful thought needs to be given to what information will be shared and how this should be done.

A second issue concerns ethics and confidentiality in local evaluations, and strategies to gain information from collaborating agencies while adhering to all their ethical concerns.

A third issue will be strategies for disseminating local evaluation results within an action research model.