

BIROn - Birkbeck Institutional Research Online

Myers, P. and Barnes, Jacqueline (2005) Measuring outcomes: guidance on outcome evaluation for Sure Start Local Programmes. Documentation. DfES Publications.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/55538/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html> or alternatively contact lib-eprints@bbk.ac.uk.

**Measuring Outcomes: guidance on outcome evaluation for
Sure Start Local Programmes**

**Patrick Myers
Jacqueline Barnes**

December 2005

NESS

NATIONAL EVALUATION OF SURE START

CONTENTS

Introduction	3
1. Why is outcome evaluation so important?	6
1.1 From inputs to outputs	
2. Identifying short-term and long-term outcomes	12
3. Research designs for outcome evaluation	18
3.1 Using a baseline	
3.2 Case Studies	
3.3 Before and after designs	
3.4 Control or comparison group	
3.5 Reflective approaches	
4. Methods for measuring outcomes	22
4.1 Qualitative	
4.2 Quantitative	
4.3 Observation	
4.4 Other, less formalised, methods	
5. Conclusions	32
Appendices	33
1. Measurement scale references	
2. NESS Support team contact details	

Introduction

The Support Module of the National Evaluation of Sure Start (NESS) has developed a range of guidance documents to assist Sure Start Local Programmes (SSLPs) in undertaking evaluations of their services and projects. These are intended to help programmes conduct worthwhile and appropriate evaluations from which internal and external stakeholders are able to assess the impact that the programme is having within its locality and on the programme's beneficiaries. This guidance focuses on the rationale and methods to measure the outcomes from SSLP services and activities. Outcomes have the power to answer the question 'What difference is this programme making?' or 'What difference is one particular service making?' It is important for local evaluation strategies to ensure that questions such as these, that are looking at outcomes, are addressed, and that the methods employed are appropriate. Box 1 provides a definition of potential outcomes.

Box 1. Towards a definition of 'Outcomes'

For Sure Start local programmes, outcomes are the changes that have been made as a result of the programme's activities. They document the programme's progress towards the long-term goal of 'better life chances for children in areas of greatest challenge and need'.

Outcomes can be:

- Changes in the **people** the programme comes into contact with.
- Changes in the **organisations** that the programme comes into contact with.
- Changes in the **environment** in which the programme operates.

Outcomes in the context of early years

Early Years services are now expected to orientate activity to outcomes. '*Every Child Matters, Change for Children*' details an array of outcomes that all services for children are expected to focus upon, contribute to and realise change as a result of service resource and effort. SSLPs, defined as local change programmes, are expected to work with a clear focus on improving outcomes for children. These outcomes are:

- Be healthy
- Stay Safe
- Enjoy and Achieve
- Make a positive contribution
- Achieve economic well-being

Working towards improving outcomes for children and families is not a new concept for SSLPs and those who work in them. The underlying theory behind the reasons for the provision of extensive resource, such as that seen in SSLPs, directed to those in areas of greatest challenge and need, is that beneficial change can be made for family and child outcomes. Evaluation therefore must assist programmes in validating their contribution to meeting objectives and outcomes such as those described above.

This guidance is for programme staff and evaluators. It aims to illustrate how to identify pertinent outcomes and how changes can be detected. It should be added that other forms of evaluation such as formative and process are not to be dismissed. They are appropriate to measure certain aspects of programme development and implementation. Together with monitoring and output evaluation they play an important role in demonstrating what has happened in the programme and where the programme is in terms of its plans and activities. As a reminder of the possibilities, Table 1 has definitions of the various types of programme evaluation.

Table 1: Types of Evaluation

Formative: Evaluation that can be used to discover if there is a need for a particular service, or to map where there are gaps in existing services. This often is called an evaluation of need.

Process: Evaluation looks at the way the programme and the services provided have been implemented and delivered and can often be used to assess how well the programme has achieved its delivery plan ambitions

Output/Monitoring: Evaluation to measure the productivity of the programme. This often relies on collecting and reporting reach data such as attendance at activities, number of families reached and the number of new contacts compared to existing contacts over a certain time period.

Outcome/Summative: Evaluation that asks questions about what has changed as a result of the programme and its activities. Outcomes can be either short-term or long-term and identifying such outcomes will be an integral part of demonstrating the value of a service, activity or programme.

All the types of evaluation described in Table1 are important in terms of communicating to stakeholders the programme's operations. However, it is sometimes not enough to simply describe the programme's activities. They are provided with a focus on achieving some form of outcome. Programmes and evaluators need to develop strategies and mechanisms by which appropriate outcomes can be identified and once identified collect data about them in a reliable and credible way.

Outcomes are important because they provide a mechanism by which programmes are able to assess the impact that they have had on their beneficiaries. After describing the implementation and process of delivering services, at some point programmes and services need to produce evidence to document what they have realised for the populations with whom they have been working. That way, observers of the programme are able to attribute value to the work that has been undertaken.

The first section of this guidance discusses why outcomes are important in the context of what Sure Start programmes measure and then provides illustrations of some of the outcomes that programmes should be able to detect. Subsequent sections provide illustrations of some methods to access data on outcomes.

The guidance provides ways to meet the expectation that if something has worked well, the result will be discernable as benefits, impacts and changes that are amenable to measurement and documentation¹. A range of methods can be used to facilitate outcome evaluation. Thus it could be said that outcome evaluation is more of an approach than a particular method, since it relies upon such a wide range of data collection techniques, both qualitative and quantitative. The task of outcome evaluation is to provide evidence of changes, which can be attributed to programme activity, changes that allow the programme to learn and therefore influence service delivery through the dissemination of good practice.

¹ This guidance needs to be read in conjunction with other guidance prepared by NESS, particularly the documents "Conducting ethical research" and "Implementing and managing your local evaluation".

1. Why is outcome evaluation so important?

*It is no longer good enough to say that something works;
the evidence must support that assertion.
(Statutory Agency Representative)*

Evaluation is about providing evidence of what works and what does not work. Evaluation also seeks to place a value (monetary and in terms of improving circumstances) upon what is being achieved and to that end it becomes very important that SSLP evaluations consider what has changed because of the programme's presence and efforts. There are many reasons why this is important including:

- Effective decision making, such as allocation of resources.
- Reshaping and programme improvement
- Accountability for resources used
- Developing an effective evidence base
- Delivering better services
- Building an evidence base of what works

Evaluation in SSLPs must have some utility for programme operations. The evidence that is provided should have the power to inform programme development, the allocation of resource and the delivery of appropriate services for the potential beneficiaries. Outcome evaluation is therefore a mechanism by which an assessment can be made of changes within the programme. Consideration of outcomes thus becomes integral to the process of performance management. Evaluation, and particularly outcome evaluation, therefore validates the programme's inputs and activities by linking them to specific impacts.

It is important not to attempt too much. Evaluation within SSLPs will never be able to capture all the practice that occurs in the programme and at best can focus on a limited range of services that occur at any given time. Programmes have repeatedly requested ideas about how they are able to incorporate some of the practice that they see as innovative and worthwhile into their evaluation outputs. An evaluative culture coupled with a focus towards outcomes is such a way in which evidence of what works can be recorded.

When planning services, practitioners usually have clear ideas about the changes that they wish to see as a result of the activity they are implementing. Change is the key factor in outcome evaluation. Typically, many services anticipate changes in behaviour, skills, knowledge, attitude and status of the beneficiaries of the activity planned. Developing methods that practitioners can use to capture these outcomes as part of an evaluative culture has benefits for programmes and for practitioners.

An approach to evaluation that focuses on outcomes benefits practitioners greatly. It provides information that can help them improve their work with programme participants, learn more about programme priorities, participants' needs and most of all documents the successes that they achieve. All programmes, when committed to evaluating outcomes, will need those

individuals delivering services to be committed to the process of an outcome focused approach that an evaluative culture can engender as they are often involved in collecting vital information, and recording it appropriately.

This also means that evaluation should not necessarily be seen as simply proving something. It should also be seen as contributing to the programme dynamic by which services are continually reviewed so that improvements can be made in delivery and outcomes. However, without some attempt to link activities to outcomes this becomes a hit and miss task. Developing a credible description of the programme and the success or otherwise of its provision relies upon a systematic approach to capturing the changes, benefits and impacts that are the outcomes. One such approach is to use a logic model to see how outcomes directly attributable to the programme come about (see Figure 1).

1.1 From Inputs to Outcomes

There has been an increasing interest in measuring outcomes as part of any evaluation strategy. However it has also been recognised that the measurement of outcomes for programmes is more difficult than other forms of evaluation. This is because outcomes take time to become evident since they are often linked to long-term objectives. However various methods have been described that attempt to join together the different aspects of programme delivery and outcomes. Various names have been given to such approaches; perhaps the most common is that of the logic model². At its simplest a logic model is a systematic and often visual way of demonstrating relationships between resources used, the activities those resources facilitate and the changes that you hope to achieve. The basic logic model is detailed below in Figure 1.

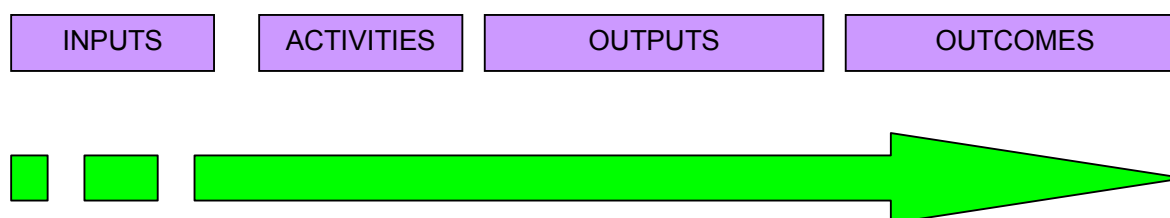


Figure 1: Basic Logic Model

The basic logic model seeks to explore and connect a variety of information available to a programme in order that an overall assessment can be made about what works and why. The model allows for myriad types of information to be collected through a variety of evaluation approaches and methods, bringing together information to develop a coherent picture of how the programme is performing in a number of areas.

However, many of the outcomes that SSLPs are seeking to achieve are long-term by their nature. Local evaluation therefore also needs to capture a range of outcomes that reflect what is happening in the shorter term in order to demonstrate progress. This is like measuring the stepping-stones that create a longer-term impact or in certain instances measuring the contribution that SSLPs make to targets and outcomes shared with others. The basic logic model can then be enhanced to incorporate some of the outcomes that programmes are able to measure and report. Figure 2 and Box 2 illustrate some of the additional outcome dimensions that can be measured by programmes to indicate progress towards longer-term outcomes and goals.

² Rossi PH, Lipsey MW, Freeman HE.(2004) Evaluation: A Systematic Approach. Seventh Edition. Newbury Park, CA: Sage
Institute for the Study of Children, Families and Social Issues, Birkbeck,
University of London

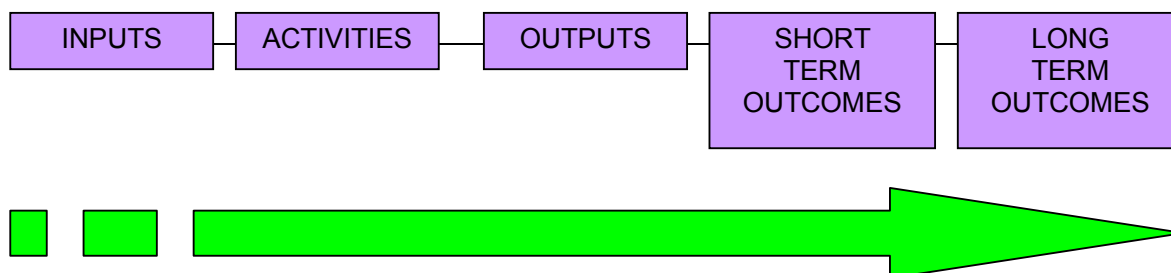


Figure 2: Enhanced Logic Model

This model provides a visual representation of the outcomes pathway by which the short-term outcomes contribute to long-term goals. Through this type of model programmes can see the ways in which the resources that have been allocated to activities according to the long-term goals are mediated by outcomes that are closer to the programme's activity. Making the connection between resources (inputs) and activities and long-term outcomes then permits the programme to reflect upon how it may be able to assess its contribution to such targets. For example the *Every Child Matters* framework has a stated target of reducing the number of children under eleven who are classified as obese. Much of the healthy eating and play activity being undertaken in SSLPs will contribute to this long-term ambition. It is the task of outcome evaluation to assess and measure those aspects of the programme that contribute to this goal.

Box 2: Components in the Logic Model

INPUTS: These are the resources that are used to plan, implement and provide a programme and the services it delivers. In SSLPs these would be: revenue and capital funding, staff, volunteers, facilities, partners etc.

ACTIVITIES: The activities and services that the inputs realise. The activities should be delivered with outcomes in mind. One example could be the 'stay and play' that is provided in the Sure Start locality. The activity of stay and play where parents are encouraged to play with their children as part of a learning process, has the power to contribute to an array of outcomes, including better cognitive development and socialisation skills.

OUTPUTS: This refers to the productivity of activities in terms of the numbers of families, children or whatever the output criterion is, and the volume of activities undertaken. These are most often objectively quantified measures such as attendances, number of families accessing a service etc. They can rely heavily upon high quality monitoring data.³

³ Please see NESS Guidance "Using Existing Data in Sure Start Local Evaluation" for additional information on this important aspect of programme evaluation.
Institute for the Study of Children, Families and Social Issues, Birkbeck,
University of London

SHORT TERM OUTCOMES: These are the stepping-stones by which progress towards the longer-term goals can be assessed. These can rely on both qualitative and quantitative documentation and are vital in programme planning. They are intermediary steps in the realisation of the longer-term outcomes. They focus on events that are necessary to achieve long-term outcomes and impacts.

For example a SSLP may be working towards a reduction in the number of households with children 0-4yrs where no one is working. Measuring the steps along the way towards that target may require some evaluation of the softer outcomes. These could include for example changes in confidence about interview situations after a back to work course, increase in computer literacy or the benefits that volunteering has had in developing skills for the workplace.

LONG TERM OUTCOMES and IMPACTS: Much of the work of SSLPs may not be realised in the early years of the project. For example the programme's goals include the raising of attainments with particular reference to the Foundation Stage Profile. While changes may not be evident in the first years it is useful to put into place strategies for examining FSP scores in local schools as soon as possible. Another example would be the reduction in social exclusion experienced by children as a result of speech and language difficulties. By introducing a systematic strategy for documenting referrals to speech and language therapists, and then following up these children over time, long-term outcomes will gradually emerge.

These outcomes are by their nature, long-term, but should remain the focus of programme planning and implementation. They may require close liaison with other services for children, especially local schools, so that information is collected in a systematic manner over time. Clearly the National Evaluation of Sure Start is examining this type of outcome across the country, but local populations of special interest (e.g. refugee children, those with English as a second language) may be studied in local evaluations, which have the capacity to study outcomes in the context of local need.⁴

The logic model is just one way of conceptualising the evaluation pathway to tease out outcomes. Other evaluation models also provide ways to assess the outcomes. These include 'Objective Based Evaluation' and 'Goal Free Evaluation'. 'Goal Free Evaluation' is characterised by avoiding predetermined goals that narrow the focus of evaluation. This then enables evaluators to focus on actual outcomes rather than intended programme outcomes. As such this type of evaluation is particularly adept at identifying unanticipated outcomes⁵ 'Objective Based Evaluation' also focuses on outcomes by specifying the purpose of a particular programme and then determines if or to what extent, those stated objectives have been attained.⁶

⁴ Please see the NESS Impact Module section on www.ness.bbk.ac.uk for some additional information about long term outcomes and their measurement in the context of SSLP

⁵ Caulley DN (1997). What is Goal Free Evaluation? Evaluation News and Comment Magazine 6: 2

⁶ Luo M & Dappen L (2005) Mixed methods designs for an objective based evaluation of a magnet school assistance project. Evaluation and Programme Planning 28, 1
Institute for the Study of Children, Families and Social Issues, Birkbeck,
University of London

However, the use of logic models in planning provides a mechanism by which a wide range of data sources can be considered. They can facilitate triangulation, the process of confirming findings by examining different data or perspectives. Equally the logic model takes into account the expressed outcomes that the programme is attempting to influence and change. As such it becomes important that evaluation in SSLPs must incorporate a range of data sources that reflect the different aspects of the model described.

Monitoring and output data will only make sense if outcomes are associated with such a record of programme productivity, and likewise outcomes will only be meaningful if they are presented in the context of inputs and activities. Remember, the aim of programme evaluation endeavours to provide information that has utility for the programme whilst at the same time providing information to those who wish to know something of programme operation and success.

2. Identifying short-term and long-term outcomes

Inputs and activities are the building blocks of change and outcomes. Unless these changes are measured the value of the activity and input cannot be determined. At the planning stage of any programme or service the intended outcome should be identified. Rather than thinking about activity the first consideration is what the programme is attempting to achieve. In other words what changes do you anticipate as a result of the programme's work? The next task is to devise the methodology to document these changes. Below are some questions that you should ask about your intended outcomes.

- Do the outcomes represent impacts that occur as a result of the programme's activity and services?
- Are the outcomes within the programme's scope of control and sphere of influence?
- Are the outcomes perceived as valid by those involved in the programme?
- Are they conceived in terms of change, benefits and impacts?
- Are they measurable?
- Are they important to the programme?
- Are they important to the wider context of Change for Children?
- Will they enable the programme to influence the mainstream services?

At this point it is worth looking at an example using the logic model previously discussed to provide an illustration of the difference between outputs, short-term outcomes and long-term outcomes and impacts. For this purpose, aspects of service delivery aimed at reducing the number of emergency admissions to hospital of young children with accidental injury are examined (see Figure 3).

Measuring Outcomes

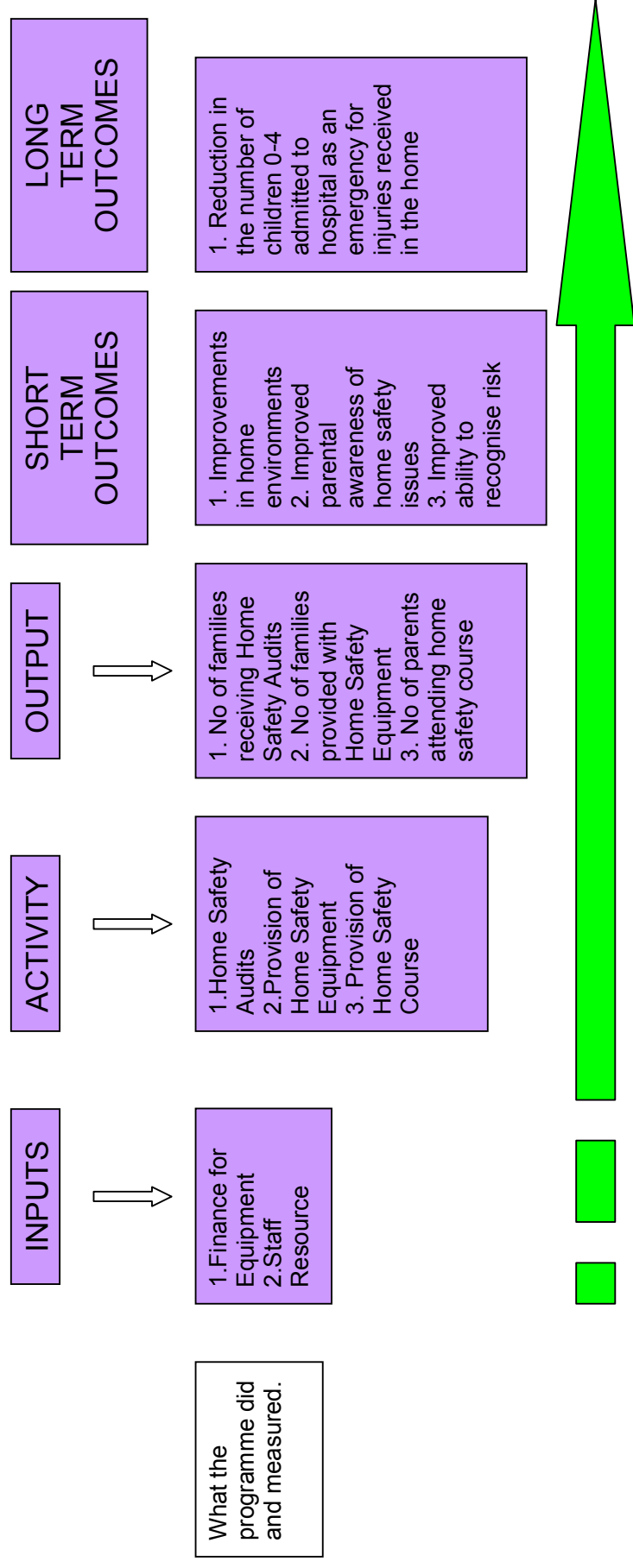


Figure 3: Example of outcome evaluation – reducing number of children 0-4 hospitalised for injury in the home

In the above example it is possible to see the way that a programme uses the resources at its disposal to create activities that anticipate outputs and outcomes. The output provides some notion of the productivity of the programme and also can assist in information needed for cost effectiveness and cost benefit evaluation. Outputs are part of the process of measuring programme activity that is directed to outcomes.

In the example it is clear that the evaluation will need to focus on what participants of the various activities associated with the outcome have realised because of their involvement. This may be what the participants know, think, or can do, which is different following the activity provided by the programme. The task is for evaluators to continually consider ways, both qualitatively and quantitatively, to capture these short-term outcomes, both for themselves and for those who have embraced an evaluative culture.

Often, measuring change is seen as relying simply upon hard data, numbers, percentages etc. However 'soft' outcomes (those not so easily defined or assessed) are equally important in the process of measuring change and can be seen as evidence of working towards long-term outcomes. Table 2 gives some of the 'softer' outcomes that may be associated with a return to work training programme run within a Sure Start setting, and the kinds of evidence that might be collected.

Table 2. Soft outcomes for a return-to-work programme⁷

'Soft' Outcomes	Evidence
Key Work Skills	Acquisition of key skills such as IT
Attitudinal Skills	Increased motivation Increased levels of confidence Recognition of existing skills Increased levels of self-esteem Higher personal aspirations
Practical skills	Ability to complete forms Ability to write CV Improved ability to manage money Improved awareness of rights

The key to effective short-term outcomes is that they are measurable. Whether they are measured quantitatively or qualitatively they remain amenable to detecting that, as a result of the programmes inputs, activity and outputs, changes have taken place. Short-term outcomes then become the stepping-stones to which longer-term changes are realised.

⁷ Adapted from Guide to Measuring Soft Outcomes and distance travelled, IES 2000
Institute for the Study of Children, Families and Social Issues, Birkbeck,
University of London

To summarise the key aspects of outcomes: they should be SMART:

- **S**pecific
- **M**easurable
- **A**ction Orientated
- **R**ealistic
- **T**imed.

The most important of these features is being realistic. SSLPs are under pressure to provide evidence of their impact within time frames that are short in research terms. Selecting outcomes that will require evidence that is expensive to collect (such as individual testing of children) or that will require detailed analysis (e.g. detailed qualitative interviews) should be avoided.

Softer outcomes such as those detailed above are important. Table 3 provides some illustrations of how these outcomes can contribute to the longer-term outcomes that they seek to address. Many programmes are keen to evidence their contribution through the assessment of changes in soft outcomes. When using soft outcomes as short-term measure towards longer-term goals it should be remembered that a credible and evidential pathway by which the long-term outcomes will be affected must be articulated.

Measuring Outcomes

Table 3. Programme Activity Outputs and Outcomes routes

ACTIVITY & OUTPUT	SHORT-TERM OUTCOMES	LONG-TERM OUTCOMES
Number of new mothers who attended breast feeding support group	Mother's intention to breast-feed is translated into successful breast-feeding.	Increasing % of mothers breast-feeding. Better infant nutrition leading to better health in the long term
Number of teenage mother groups held in various locations	Consistent attendance at teenage mothers groups (note that in some instances an output is also an outcome)	As a result of group content a reduction in the number of teenage mothers who have second children in their teens
Number of early years settings receiving speech and language training	Increase in early years practitioner's knowledge about communication. Development of language enriched environments	Reduction in the number of children commencing school with undetected communication problems
Number of families offered appropriate behavioural support workshops	Development of improved skills in dealing with conflict with children and increased understanding of strategies of negotiation	Improved family circumstances derived from less conflict
Evidence of collaborative working with voluntary sector and Sure Start teams. (No of joint initiatives undertaken)	Improved co-operation between teams	Seamless pathway between services
Number of food hygiene courses offered	Improved knowledge of basic food hygiene	Reduction in instances of gastroenteritis associated with poor hygiene practices
Number of parents attending play and stay sessions	Increased co-operative play between parents and child. Improved knowledge of learning as play	Impacts upon the foundation stage profile

Measuring Outcomes

Number of childcare places supported through Sure Start Revenue	Opportunity for parents to undertake training and/or participate in other programme activities	Parents return to work or training reducing the number of families with children 0-4 out of work
Number of play settings supported by Occupational Therapist	Increase in the number of activities associated with key developmental markers	Children exhibit appropriate behaviours and abilities once at school thus contributing to better learning outcomes
Number of parents attending return to work course	Increase in confidence in interview situation Improved CV writing skills Participants learn job seeking skills	Parents are job ready when they are able or decide to return to work or training
Number of parents receiving debt advice in the drop in sessions	Improved understanding of money management	Reduction in emotional and mental pressure brought about by anxiety and improved family circumstance

3. Research Designs for Outcome Evaluation

This section covers some of the different types of designs and approaches that can be used to detect change. Change has been discussed as a key concept when considering outcomes. It is worth noting here the importance of defining the desired outcomes of any project – much as one would identify research questions before undertaking a research project. Once outcomes have been identified it makes the evaluator's task easier by being able to match the approach and method to more reliably measure the anticipated changes

The later section provides some basic reminders of the methods that can be employed by evaluators. There is more detailed guidance on the NESS website covering specific areas of methodology⁸ so this section will only offer a brief overview of how programmes can go about detecting changes.

3.1 Using a Baseline

All SSLPs have made some attempt to establish baselines for a range of indicators at a point in time. For example many programmes used secondary data to provide baselines of library membership or referrals to social services at the commencement of the programme. Programme delivery plans were often used in this context so that programmes could assess changes that were occurring as a result of activity and outputs. Revisiting the baselines, where it is feasible and realistic to do so can offer useful information of how and why programmes and activities are responsible for the differences, if indeed any are detected.

What is a baseline?

Baseline Information describes the starting position of a programme or initiative across a range of indicators, ideally before any services were offered. To measure the impact of SSLP activity it is important to have local baseline statistics that can be used to measure improvement across a range of objectives. Effective baseline indicators must be appropriate to the programmes activity and must be measurable.

Indicators can be qualitative and quantitative. They are often derived from existing datasets that have recorded a key variable at a point in time. For example it is possible to measure unemployment rates in a locality at the onset of the project and then regularly assess changes that occur. Indicators cover such things as demographics, perceptions, physical facilities etc. Their use is a very important way to measure performance, and potentially a cost-effective way if the information is already being collected.

⁸ See in particular <http://www.ness.bbk.ac.uk/documents/GuidanceReports/395.pdf> for guidance on using existing data in your evaluation
Institute for the Study of Children, Families and Social Issues, Birkbeck,
University of London

3.2 Case Studies

Case studies are often used to illustrate the progress and changes that have been experienced by an individual (sometimes it can be more than one individual) in a particular programme. For instance this has been reported on the basis of the 'distance travelled' by a parent from their first encounter with the programme to where they are now. They can be powerful illustrations of the impact that the programme has had on a family, parent or child. In order to maximise the effectiveness of this strategy it is worthwhile considering if you can include both qualitative and quantitative methods. For example if you are intending to describe a family's or parent's pathway through the programme it would be appropriate to include some detailed statistics of the range and number of activities accessed by the case study participant. Typically case studies draw upon interview and observational information as well and as such they are able to offer a rich description of what changes individuals have experienced for themselves and their children. Case studies can also be conducted with staff in particular disciplines (e.g. health visitor).

3.3 Before and After Designs

This approach makes some attempt to compare the target population, or conditions before and after a particular service or activity has begun, similar to baseline, but more closely linked with the population using a specific service.

Before and after studies are very useful at measuring short term impacts in that they can detect changes that occur as a result of a service or activity. At the very least it is able to consider the association with participation and a desired change and outcome. Many before and after studies attempt to measure change very close to the end of an activity or participants involvement with a service. In order to detect whether the changes have been sustained it is worthwhile incorporating three data collection points, before, after and perhaps 3-6 months later. The sustained benefits of any change can then be reported (see Box 3 for an example).

If the intention of a particular activity was to improve knowledge about safety in the home, then some form of assessment before a workshop begins and another after should be able to tease out how knowledge has changed about aspects of safety.

Before and after measures can be either undertaken using quantitative or qualitative methods. For example if the programme evaluation was examining the impact a speech and language project was having on parent to child interaction it may be that observations of eye contact and turn taking is undertaken. At the same time interviews with the parent can examine more wide-ranging aspects of the parent child relationship. Changes from the beginning of the project to later can then be recorded and some associations can be drawn between the project and the outcomes.

Box 3. Example of evaluating the outcome of a SMOKING CESSATION service

Smoking cessation has been recognised by many programmes as a difficult target to meet. There is a dearth of reliable statistics detailing the prevalence of smoking in pregnancy and after the baby is born. The short-term outcomes of the service may not be smoking behaviour itself but attitudes to smoking, with the longer-term outcome changes in smoking behaviour.

For example if programmes set the initial outcome of a smoking cessation class as moving people from not even considering giving up, to what the literature calls beginning to contemplate, then a short term outcome has been achieved.

If at the beginning of a smoking cessation course parents are unaware of the dangers of smoking around children and by the end of the course they are aware and have decided to smoke away from children, another short-term outcome has been achieved.

The art of outcome evaluation is to consider what can be effective in the short term, how that relates to the longer term and how effectively each step can be measured.

3.4 Comparisons or Control Group

Evaluators may decide to compare the outcomes for one population – receiving a service - by comparing it with another group who have not received the service. For example if the outcome of interest were to be a reduction in referrals to acute speech and language therapy services, evaluators could compare rates between those of the SSLP and the rest of the PCT area.

This approach can also be combined with before and after studies in that a comparison or control group is measured at the start of the intervention in the SSLP area and the impact compared later between those whom were involved in the initiative and those who were not. This can then show how the 'experimental' group differed from the comparison group.

A study with a control group would involve a more systematic means of selecting individuals for the service. Ideally all would be eligible and those receiving the service would be chosen on a random basis. In this way the 'real' effect of the service can be identified. However in real life this is an unusual research plan, and more often the control group may be those on a waiting list.

There are different types of control group, some more difficult to attain than others. Matched control groups are difficult to achieve but a comparison group design can more easily be attained.

1. Experimental and Control

This involves randomly assigning a group of participants to an experimental or a control group. Using SSLP evaluations as an example, a group of parents and/or children randomly assigned to have access to the programme's activities and services would be the *experimental group*, whereas a group of randomly chosen individuals, whose characteristics are similar to the experimental group, but who are not given access to the services that the programme offers, would be the *control group*. Outcomes can then be compared between the two groups. Issues around ethics and methodology make this approach very difficult to achieve in the local evaluation of SSLPs

2. Comparison Groups

The problems with the above make the use of existing comparison groups a more appropriate method for local evaluation. Comparison groups are groups of individuals whose characteristics are similar to those of the programme participants, but these individuals are not exposed to any of the services that programme participants have access to. However, there would be no need to randomly assign participants to the comparison group, as the group would already exist. Outcomes can then be compared between the two groups.

3.5 Reflective Approaches

SSLPs place great emphasis on the role that parents, staff and other members of the community play in developing and delivering services within their locality. It seems sensible then to explore ways in which parents and staff can contribute to the knowledge of what has changed. Methods to achieve this will be discussed later but for now it is enough to say that asking people to reflect back to what it was like before, is an appropriate evaluation strategy.

4. Methods for Measuring Outcomes

Having outlined some of the general research designs used to consider outcomes, a brief overview of measurement methods and how they can be used is now presented.

A successful outcome evaluation needs careful planning. The outcomes that are to be measured should be identified as early as possible in the evaluation planning process. Having clearly identifiable outcomes makes the task of matching the evaluation method to the task easier. The strategies detailed above allow for the use of both quantitative and qualitative approaches to data collection. What is vital is that the methods and implementation of the evaluation strive to establish credibility and reliability through the application of systematic data collection and data analysis.

4.1 Document Review/Using Existing Data

This could include examination of:

- Activity Records
- Participation rates
- Local Context Analysis Data
- Budgets
- Delivery Plans
- Meeting Records

Advantages: This information has generally been collected for another purpose and is readily available to the evaluator. It is also often in a quantitative format that is easy to use and it is less resource intensive than collecting new data.

Outcome Use: This information, whilst providing some information around outputs, also provides some of the contextual information that is needed to understand how outcomes have come about. However it is also able to offer some insight into changes that have occurred across time in terms of organisation's outcomes.

4.2 Qualitative Methods

These include:

- Interviews
- Focus Groups

- Diaries (see Box 4)
- Multimedia Methods (see Box 5)

Advantages: This type of method usually involves naturalistic types of information and when collected in a systematic way, and analysed according to good qualitative research standards, provides an enriched account of people's experiences. The methods can be particularly useful for people who do not have good reading or writing skills or where English is not the first language, or for work with children.

Outcome Use: Qualitative approaches often tease out some of the unexpected outcomes that more structured approaches do not facilitate. By asking participants to identify what has changed for them because of contact with the programme or a particular service, personal stories can be constructed from the data. Asking participants to relate their experiences in terms of benefits, outcomes and impacts provides the necessary information to document the changes being experienced.

However, unless multiple accounts are analysed in a systematic and thematic way the evidence is little better than a case study. This may be extremely useful when trying to understand the stepping-stones, the short term outcomes, that lead to the longer-term impacts that a range of individuals have experienced.

Box 4. Example of qualitative method - Food Diaries

Assessing the impact of a nutritional programme can prove difficult. The long-term goal may be a reduction in childhood obesity, but a programme wanted to show how the work it was carrying out was contributing to such an aim. Theoretically it is possible to link early nutrition to worse outcomes in terms of obesity so it seemed sensible to measure family dietary habits after an intervention to increase fruit and vegetable intake. A two-week diary was compiled in which parents were asked to list contents of meals, lunch boxes etc. Accepting the limitations that may exist about self-reporting the programme felt confident in documenting the influence the intervention had on family eating habits. The programme also made an attempt to assess the sustainability of the intervention by following up participants at a later date.

Box 5. Example of qualitative method - Photography

Before and after photography can illustrate changes that have occurred in the community. For example the improvements that a programme may have made in physical play equipment can be documented. When this is coupled with other data sources such as questionnaires and surveys about frequency of use and parental perceptions of improvements it becomes another tool by which programmes can illustrate the impact that they are having in their particular community.

4.3 Quantitative Methods

These include:

- Surveys
- Questionnaires
- Feedback Forms
- Visual Analogue Scales (see Box 6)
- Measurement Scales (see Box 7)
- Existing Data

Advantages: These approaches generally collect uniform data that are straightforward to analyse. Often surveys and questionnaires are administered anonymously so it *may* be that participants will respond more openly than with other methods. They are less time consuming than qualitative methods and can provide information on large numbers of people.

Outcome Use: The standardised nature of questionnaires and surveys make them very useful for before and after studies. Asking a question about a particular soft outcome and revisiting the same question at a later date gives some insight into what has changed as a result of some programme intervention. Causality may not have been established through this method but some associated benefits will be able to be identified.

Box 6 provides an illustration of a self-constructed measurement scale. Visual Analogue Scales (VAS) are scales that have been tested for reliability and validity in terms of detecting and measuring change. It is also worth noting that there are many existing scales and measurement instruments that can be used to see what the impact has been of a certain activity or range of services. Box 7 provides some illustrations of existing scales and measures that can be used to assess change.

For example programmes assessing the impact they have made on reducing levels of post natal depression may wish to apply the Edinburgh Postnatal Depression Scale before and after an intervention and then return and reapply 6 months later to assess the sustainability of any change detected immediately after the intervention. Other scales that could be used include the Rosenberg Self Esteem Scale to measure changes in self-concept. The advantage of using such scales is that they have been subject to much testing to ensure they are effective at measuring what they intend to measure.

Box 6. Example of a quantitative method - Self Constructed Visual Analogue Scales

One criticism of questionnaires and surveys is that often they do not represent the key dimensions that may be pertinent to individuals. Sometimes short-term outcomes vary between people and standardised questionnaires sometimes are not able to register these nuances. Visual Analogue Scales are used to determine how someone feels between two constructs on a particular continuum. They can be tailored to individual expectations of service or project delivery in 4 easy steps.

Before a workshop, project or programme of work, practitioners can ask individuals some questions about how they feel regarding some of the content of the project. For example, if the project was looking to improve key skills associated with returning to work the following questions could be asked.

1. How confident do you feel about a work-based interview?
1. How nervous are you about returning to work?
2. How well do you feel you could compile a CV?
3. Do you have any idea about how much working tax credit you could get?

The answers to these are positioned at one end of a line usually about 10cm long and the participant is then asked to state what the opposite of their reply would be, this is placed at the opposite end of the line. Thus, if the answer to question 1 was that the participant felt 'not very confident' the opposite would be 'very confident'.

Question 1

Not very Confident   Very Confident

Question 2

Very Nervous   Not very nervous

Question 3

Not very well   Very well

Question 4

No idea at all   A good idea

Participants are asked to place themselves on the line indicating where they consider themselves to be at that moment in time. The red responses above represent this first data measurement point. After the intervention the scale is applied again where the green responses represent this second data measurement point, hopefully detecting some change on position on the continuum from the first response.

Box 7. Using Existing Measurement Scales.

Many scales have been devised that measure perceptions and characteristics. Below are some scales that are useful in the early years settings. Some can be used as before and after measures to detect changes in status etc.

Strengths and Difficulties Questionnaire. (Goodman, 1991).

This is a behaviour rating scale to measure mainly problematic behaviour (e.g. fighting, over activity), but it also includes some pro-social behaviour (e.g. helping peers). This is available in two versions: for children of 2-3 years, and for children of 4-16 years.

Child Well Being Scale (Magura and Moses, 1986)

Rating scale that examines children's needs, physical psychological and social)

Pleasure in Parenting (Fagot, 1995).

Parent-Toddler interactions and parental perceptions measurement scale.

Family Support Scale (Dunst, Trivette and Deal (1988)

This scale asks questions about support in relation to how helpful it has been in terms of raising your children

Adapted from Outcomes and effectiveness of Family Support Services⁹ See appendix 1 for references

Existing data can also be used to document change. This guidance stresses the importance of outcomes and in certain circumstances outputs can be seen as outcomes. If the aim of a teenage parents group was to develop trust, mutual support and social networks then it may well be that continued and sustained attendance at the group can be seen as a short-term outcome. Monitoring data then becomes a vital tool in evidencing these and programmes can make great use of the internal programme data to support these types of analysis.

This can be true of an evaluation examining what the capital programme has realised. Monitoring data will give an indication of changes in attendance that may have been facilitated with the programmes newly built premises. Questions such as 'has attendance increased and what impact has that had on staff and children?' will rely on both existing data and creating new data from other forms of evaluation.

Changes can be detected in data collected by others, such as health and local authorities. For example over time it is possible for programmes to see the impact that an oral health programme may have by requesting data on filled

⁹ Statham J (2000) Outcomes and Effectiveness of Family Support Services. Institute of Education. London.

and extracted teeth in children 0-4 and above in certain cases.¹⁰

4.4 Observation

Observing the Community (Community Change)

Observing facilities

Observing group activities

Observing participants

Advantages: Observations are generally non-intrusive and as such do not require much participation from those being observed. Observations can, if they are undertaken in a professional manner, be very naturalistic, capturing data that occurs without any researcher effects.

Outcome Use: Changes in behaviour can be detected through the use of observation. By taking baseline observations it is possible for example to document the level of parent child interaction in a given setting. Other observations could include the amount of co-operative play, concentration or sharing that occurs between children. Changes in facilities and environments are also usefully observed. For example, if the desired outcome from speech and language activity was an improvement in language enriched environments, then observations could validate those changes. Using multimedia-recording techniques would also provide other stakeholders with pertinent information of the change. If a project sought to regenerate a particular area in the community with either a horticultural project or a new play area using observation again with visual evidence can show some of the impacts that may have occurred. They will be seen as stepping stones perhaps towards reduced obesity in the long term through physical play or healthier eating.

There is every reason to mix the methods highlighted above and often one form of data compliments the other. However evaluators must always keep in mind the purpose of the evaluation and decisions about methods should always reflect the type of question that is being answered. The method should be seen as a way to deliver confidence in the evaluative process and the results that are presented.

Box 8. Example of Observations - Changes in Shopping Habits

Labelling of food packaging can be misleading and confusing. With the co-operation of a supermarket, one programme's nutritional advisor took a group of parents on a label trail. This involved demonstrating the best way to identify foods that were really lower in fat, salt and sugar levels. The tour of the supermarket also encouraged parents to think about trying other alternatives to certain foods, done in conjunction with some cook and eat sessions. The desired short-term impact was to assess the changes in buying habits that contribute to the longer-term goal of better life outcomes through better

¹⁰ Please see NESS guidance Using Existing Data for further examples.
Institute for the Study of Children, Families and Social Issues, Birkbeck,
University of London

nutrition. Observation of weekly shopping was used to assess the products purchased and the choices parents were making. This was supported with a questionnaire.

4.5 Other, Less Formalised Methods.

Evaluation has now become such an integral part of service delivery, not just for Sure Start but a wide range of funded initiatives, that it becomes necessary to be inventive about the way data is collected. This can avoid evaluation fatigue in terms of your programme population but can also reduce the bias that may arise from familiarity by participants with certain evaluation methods. There is a range of innovative methods that can be used to detect outcomes and some are detailed below. However, because some of these can be interpreted as art based, do not let that influence the need to carry out appropriate analysis of the data. Indeed, sometimes when innovative methods are used, it becomes more crucial that the method of analysis is thorough and explicit. For example, one programme used wallpaper to draw a time line and parents were invited to write and draw comments on the timeline to reflect what the programme had achieved for them and their children. Five groups of parents contributed five time lines. The evaluator then undertook a content analysis coupled with some thematic analysis to develop a narrative of the parent's views.

Other innovative and less formalised methods to assess outcomes may include:

- Quizzes to assess knowledge acquisition
- Story Writing
- Drama
- Reflective Accounts

Earlier we discussed some of the changes that programmes may seek to achieve. Table 4 matches some of the evaluation methods mentioned above to the types of short-term outcomes suggested earlier.

Measuring Outcomes

Table 4. Outcomes and measurement strategies

ACTIVITY & OUTPUT	SHORT TERM OUTCOME	POSSIBLE EVALUATION METHOD
Number of new mothers who attended breast feeding support group	Mother's intention to breast-feed is translated into successful breast-feeding.	Questionnaire with follow up interviews
Number of teenage mother groups held in various locations	Consistent attendance at teenage mothers groups (note that in some instances an output is also an outcome)	Focus Group. The aim would be to ascertain some outcomes for the group such as increased social networks, or changes in attitude if appropriate
Number of early years settings receiving speech and language training	Increase in early years practitioners knowledge about communication. Development of language enriched environments	Observations of early years settings and staff child interactions. Questionnaire to assess changes in confidence and recognise communication problems
Number of families offered appropriate behavioural support workshops	Development of improved skills in dealing with conflict with children and increased understanding of strategies of negotiation	Questionnaire or interview to assess changes in knowledge, skills and attitude
Evidence of collaborative working with voluntary sector and Sure Start teams. (Number of joint initiatives undertaken)	Improved co-operation between teams	Staff focus group
Number of food hygiene courses offered	Improved knowledge of basic food hygiene	An end of course quiz
Number of parents attending play and stay sessions	Increased cooperative play between parents and child.	Observation Questionnaires

Measuring Outcomes

	Improved knowledge of learning as play	
Number of childcare places supported through Sure Start Revenue	Opportunity for parents to undertake training and/or participate in other programme activities	Focus Group asking what the childcare opportunity has allowed parents to achieve.
Number of play settings supported by Occupational Therapist	Increase in the number of activities associated with key developmental markers	Observation
Number of parents attending return to work course	Increase in confidence in interview situation Improved CV writing skills Participants learn job seeking skills	Self Devised VAS
Number of parents receiving debt advice in the drop in sessions	Improved understanding of money management	Questionnaire, Quiz

5. Conclusions

Outcomes matter. They are signs of the progress programmes have made in the delivery of services to their target population. Working in an outcome focussed way conveys the programme's serious intent in making sure those who are the recipients of services can experience benefit and positive changes. As such, outcomes should be an essential part of any evaluation.

Outcomes are integral in documenting how programmes are utilising resources effectively to achieve the desired improvements. The focus of this guidance has been to bring those outcomes closer to programmes by suggesting that a focus on short-term outcomes provides insight into the contributions being made to longer-term targets and goals.

Research methodologies have been discussed and this guidance also described a range of methods that can be incorporated into programme evaluations in order to demonstrate progress towards outcomes. Both qualitative and quantitative methods are useful in addressing the questions of what has changed for programme beneficiaries. It is a matter of always keeping in mind the methods that are able to detect and measure change, be that from existing data, self report from programme participants or documentary evidence. The importance of outcome evaluation in terms of developing an evidence base of what works cannot be overemphasised. The advent of the *Every Child Matters* framework gives some structure to the anticipated long-term goals for all children and families. Outcome evaluation is a route by which programmes are able to evidence their contribution to these aspirations.

Outcome evaluation also provides the opportunity for programmes to:

- Clearly communicate the success of the programme
- Make informed decisions
- Allow for reshaping and programme improvements
- Be accountable

Input, activities and outputs come together to produce measurable changes, benefits and impacts, both in the short and long term for those communities that programmes serve. As such, outcome evaluation becomes an integral part of the continued development of services to children and families and when done well has the power to change services for the better.

Appendix 1. Measurement Scales References

Dunst, C., Trivette, C. & Deal, A. (1988). *Enabling and Empowering Families: Principles and guidelines for practice*. Brookline Books. Cambridge MA.

Fagot, B. (1995). Development of a pleasure in parenting scale. *Early Development and Parenting*. 4, 75-82.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire. A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.

Magura, S. & Moses, B. (1986). Outcome Measures for Child Welfare services. Child Welfare League of America. Washington DC.

Appendix 2. NESS Support for Local Evaluation Team

National Evaluation of Sure Start, Support to Local Programme Evaluations
Institute for the Study of Children, Families and Social Issues
Birkbeck, University of London,
Tel: 020 7079 0823 Fax: 020 7323 4738
Email: NESS-support@bbk.ac.uk

Professor Jacqueline Barnes, Director
Tel: 020 7079 0837
Mobile: 07855 308311
Email: jacqueline.barnes@bbk.ac.uk