REES CENTRE
Research in Fostering and Education
University of Oxford Department of Education

UNIVERSITY OF OXFORD

DfE CHILDREN'S SOCIAL CARE INNOVATION PROGRAMME: EVALUATION GUIDE

Purpose of the Guide

Where it has come from and how it will develop

This short guide to the evaluation requirements for the DfE Children's Social Care Innovation Programme is aimed at Project Leaders and is designed to assist them and their coaches from the Spring Consortium to develop their evaluation plans with the support of the Evaluation Teams registered with the Evaluation Framework. The guide has been produced by the Evaluation Coordinator - the Rees Centre at the University of Oxford (http://reescentre.education.ox.ac.uk/) - together with the Evaluation Teams to ensure that the principles espoused are those with which everyone involved can agree. This is the first version of the guide and it is intended that as the Programme progresses additions, revisions and examples will be added in the light of the experience of the Projects and those working with them.

What the guide is for

The three main purposes of this guide are to:

- Set out the principles of effective evaluation;
- Describe different approaches to evaluation that might be suitable for assessing the outcomes of individual projects;
- Provide some consistent standards of evaluation across the Programme.

In order to maximise the use of this guide, we have kept it short. For those that wish to find out more a list of further resources is provided at the end of the guide with links or access information.

Why evaluation is needed

There are several reasons why both the individual Projects and the overall Programme need to be evaluated:

Need for on-going 'formative' feedback

The strengths and weaknesses in the Projects and Programme should be identified as the projects progress. On-going feedback to individual Project Leaders will enable them to make adjustments for example, in the resource allocation or clarification of outcomes in the light of information about how the implementation is being received. Similarly, the overall Programme will benefit from discussions that draw on on-going evidence indicating what improvements might be made to achieve better outcomes.

Justification of investment of resources
The significant investment of public resources in the Programme requires that outcomes are assessed and available in the public domain in order to justify this expenditure. Costs will need to be considered in relation to shorter and medium term outcomes and what these might suggest about longer term outcomes in children's social care. An attempt to assess the cost benefit of the innovation should be included in the evaluation plan.

Ensuring interventions in children's lives are informed by evidence
The best possible available evidence should inform interventions in the lives of children and their families. There is a moral obligation for all those involved in the provision and implementation of services to ensure that this is so. We anticipate and expect it to be so in medical services but this same expectation has been less evident in other services such as social care and education.

Challenges in Evaluation of the Children's Social Care Innovation Programme

Given these justifications for evaluation, what are the main challenges that are faced in evaluating the projects and overall impact in the Children's Social Care Innovation Programme? Research on evaluation demonstrates that it generally could be of a higher standard and that knowledge and acceptance of evidence does not guarantee that it is used. We are looking to the Evaluation Teams to help the Evaluation Coordinator to improve the quality of evaluation evidence in this Programme. The Evaluation Teams will need to challenge projects and help push them up the 'ladder' or scale of evaluation standards (see below).

The need for high quality evaluation
Service managers and practitioners need findings that they can trust. If the findings of evaluation are to be trusted, they need to be robust. This means taking an approach that will lead to findings that are:

- *Valid*: do the findings address what the evaluation intended to address or might they be due to the effect of something else?
- *Reliable*: Would the findings of the evaluation be the same if they were collected by different researchers or at different times?
- *Generalisable*: The extent to which the findings are applicable to other contexts, settings or times – this is crucial given the intention of the programme to encourage scaling up of innovations.

The guide suggests approaches to evaluation that maximise validity, reliability and generalisation while recognising the need for practical and timely evaluation.

The competing demands for evaluation
Service managers and practitioners have priorities other than evaluation, limited resources and sometimes lack the relevant skills for robust evaluation. The level of certainty they require in a short time without the disruption to their daily work may conflict with approaches to evaluation that will provide robust findings and carry wider conviction.

Managing this potential conflict between the requirements of the evaluation and of practice will need strong partnership between Project Teams, Evaluation Teams and the Evaluation Coordinator.

Levels and types of evaluation

Whether the innovation meets its objectives
Project proposals mostly include a 'logic model' or 'theory of change' that provides objectives against which the evaluation can describe the emerging outcomes. A clear description of what actually happens when the innovation is implemented and how far it appears to be meeting its goals should be possible for all projects. This should include quantitative information that documents changes.

Process evaluation: how the people involved are affected
At the next level, evaluation can provide a more reflective description of how the professionals and young people and families targeted by the innovation reacted to the innovation. This should include the issues it raises for the participants, the degree to which it seems to be reaching the people intended and the resources required. This should be possible for all projects and is likely to involve some qualitative data such as interviews.

Impact evaluation: whether the innovation is more effective than previous provision or than services in other areas
In order to address whether the innovation is doing any better than 'services as usual' or what the effects would be were it rolled out more widely, a more robust design is needed. Answering this kind of question involves comparison, ideally by randomising those targeted by the intervention as is done in controlled medical trials. This is challenging given staff in children's services are understandably committed to providing the best possible service to all children and families as soon as possible, but it is even more challenging in complex interventions involving multiple strands of innovation.

The challenge in evaluating complex interventions is partly that of assessing which aspect or element of the intervention is having which effect. However, it is also hard to assess which parts of the context are necessary for a good effect. For example, Projects in this Programme will have the support of skilled coaches and be implemented by services that are signed up to the approach and these conditions may or may not be necessary for success. This combination of complex interventions and contexts can lead to uncertainty over whether an innovation will succeed when rolled out more widely.

It will be helpful if Projects are aware of these challenges and try to address them. Sensitive and intelligent descriptions are central to this task, as is good design. Before and after measures of key outcomes should be collected for all Projects. It may also be possible to exploit the variety of contexts and interventions that exist in the Programme as a whole in order to understand the effect of the differences. For this to happen it will be necessary for all projects to collect some common data, and perhaps to develop an overall account of why their project worked as it did that can then be compared with those of others.

Economic evaluation: assessing the benefits in relation to the costs
Economic evaluation is not just about reducing the costs. The benefits accrued from the innovation must be considered in relation to reduction in costs. Sometimes reducing costs in one area, for example bringing less young people into care, increases the costs in another for example providing more support to families. Both should be considered in an economic evaluation of the innovation.

Formative and Summative evaluation
Individual Projects require formative evaluation that means providing on-going feedback

based on evidence of 'how it is going'. In addition, the overall Programme will need formative feedback to ensure it can respond by changing the support provided or criteria used to approve Projects or assess impact.

Summative evaluation involves analysing existing data or new data collected to make judgements about the impact at the 'end' of the intervention. In this Programme, most innovations are intended to continue beyond the period over which the evaluation team will be directly involved (unless Projects decide to extend their arrangements with Evaluation Teams). However, the point at which DfE is required to report on the Programme to others is a key summative point in the evaluation.

The Principles of Effective Evaluation

The following principles should guide each Project evaluation and thereby establish high standards and increase consistency across the Programme:

Principle 1: Engage stakeholders throughout the evaluation process

It is crucial to enlist the involvement of management, practitioners and users at all levels from the start of any evaluation. This will maximise the validity of the evaluation by ensuring that the context is understood, relevant questions are asked and there is a 'buy-in' to using the findings. The purpose should be made clear from the outset with a commitment to reporting back to those involved so that they can benefit from their contribution.

Principle 2: Identify incentives and establish commitment from project providers

Make clear from the outset the purpose of the evaluation and how both service providers and adolescents or practitioners will benefit from a better understanding of the impact of the innovation. A more demanding evaluation design that involves for example random allocation of participants to the innovation may be less attractive to local authorities but may be required to enable more robust evaluation to be undertaken which, if fully explained, may become more acceptable.

Principle 3: Keep things simple and minimise additional burdens on those involved

Evaluation of complex interventions is very challenging but over-complicating the evaluation may lead to either more data being collected than can be used or reducing the potential for accurate interpretation that can inform future development of the service. Local authorities in general now collect more data than they use so checking what existing measures might be used is a good starting point. Evaluation Teams will work with the Projects to identify measures/data that are already being collected and which can contribute to the evaluation and encourage sustainability of evaluation once the Evaluation Team's work is completed. Minimise the additional time required by practitioners or service managers to complete the evaluation tools. Many will find a short telephone interview more acceptable than a lengthy questionnaire.

Principle 4: Use mixed methods

The evaluation should explore as much of the variety that exists as possible, to use the strongest design that's possible and to have some baseline and follow-up description and measurement. Using an evaluation design that combines different methods (see approaches to evaluation below) such as statistical analysis and interviews and multiple sources (e.g. data on the children and the social workers) rather than a single source is known as 'triangulation' and will strengthen the findings. This is because it provides information that

helps us interpret findings. For example, the statistical analysis may suggest that social worker turnover is increasing in the service but interviews reveal that job satisfaction has increased and that the figures reflect greater rates of promotion. Triangulation also enhances validity of the findings since it reduces the likelihood of drawing incorrect conclusions.

Principle 5: Adhere to established standards of evaluation ethics

Evaluation ethics are about conducting the evaluation in a way that respects others – their rights to understand the purpose of the evaluation, confidentiality, anonymity, data protection and informed consent. Those participating in the evaluations in the Children's Social Care Innovation Programme will need to know that views that they express to the evaluators will not be used for purposes other than the evaluation, such as for appraisal or promotion. They need to feel able to withdraw from the evaluation at any time and to know that they will not be identified in the evaluation report. The relevant external professional ethical committee such as that in the LAs, ADCS, Health Service or University may need to provide clearance.

At minimum the following ethical areas should be considered in planning any evaluation:

Safeguard the interests and rights of those involved or affected by the evaluation

- consider the well-being, feelings and best interests of those involved or affected;
- obtain written and signed consent;
- inform participants of their right to refuse to be involved or to withdraw at any time;
- explain the purposes and processes of the evaluation including any implications for participants of time, cost and the possible influence of the outcomes;
- using data collected only for the explicit purpose stated to participants unless further consent is sought;
- offer anonymity and confidentiality (noting that this might not always be possible);
- verify findings with respondents and offer feedback on them.

Consider the possible impact of the evaluation from the outset and its use or misuse for those involved in the study and other interested parties

- consider the short and long term consequences of the research for the participants and others;
- identify the costs of the evaluation to participants or their institutions/services;
- ensure evaluators know the procedure for dealing with a disclosure made in the context of an evaluation interview or other process;
- identify agencies or professionals able to follow up on any possible unsettling effects of the evaluation itself.

Ensure relevant legislative (including the DPA) and policy requirements are met

- ensure those interviewing young people have current DBS clearance;
- store and transfer data confidentially using encryption or password protection;
- label participants with numbers which are stored separately to other identifying details;
- remove potentially identifying details in publicly disseminated materials.

Principle 6: Report findings in an accessible way drawing out key implications

Stakeholders and managers who have been engaged throughout an evaluation will not receive any major surprises when findings are reported as they will be aware of emerging findings during the process. A clear description of the context within which the intervention occurred and the characteristics of the families and young people involved will support any claims about generalisability. Communication of evaluation outcomes needs to take account of audiences, maximise accessibility and clarity and draw out implications and recommendations as explicitly as possible. Evaluators may need to 'mediate' these findings to support service providers in translating them into appropriate decisions about future service developments. Using a wide range of media, social networking and webinars will be likely to engage more of the young people who are the target of the Programme.

Providing consistent standards of evaluation across the Programme

The evaluation of the Projects and Programme require the highest possible standards of practices including in research design, collection and storage of research material, analysis, interpretation and reporting. This includes:

- appropriately acknowledging previous relevant evidence and evaluations;
- selecting an evaluation approach that will be most likely to provide robust findings, yet appropriate and fit for purpose in relation to the innovation;
- making use of any existing data;
- minimising burdens in the collection of new data;
- selecting appropriate samples (including control or comparison groups where possible);
- ensuring engagement of people who are the most vulnerable/hard-to-reach;
- checking possible interpretations of the findings with service users and providers;
- considering how the specific context and population might affect the results;
- ensuring follow up measures are taken sufficiently long after the innovation (and possibly after the completion of the DfE funded Programme) to provide robust evidence of sustainability.

Approaches to evaluation for assessing the outcomes of individual projects

It is important to remain open-minded and keep 'fitness for purpose' as the principal criterion in selecting approaches to evaluation. There is a wide range of rigorous methodological approaches available and an eclectic attitude to selecting the appropriate approaches for each evaluation is helpful.

Selecting and engaging participants

Control groups: Identifying participants for the evaluation will depend on the overall design. The strongest design where appropriate, will randomly allocate some participants to the intervention while ensuring that a group similar in relevant characteristics (e.g. age, needs) who do not receive the intervention nevertheless contribute to data collection as a 'control' group against which robust comparisons in outcomes can be drawn.

Comparison groups: Where random allocation is not possible for example because the participants are an existing group receiving a service, a 'comparison' group may be identified that are broadly similar to those receiving the intervention but are not involved and pre and post measures from them compared to those involved in the intervention. However, the

idea of randomisation is to maximise the chances that the sample involved in the intervention is representative of all possible individuals who might have been included.

Sample size: In general, the larger the numbers involved, the more robust are likely to be the conclusions that can be drawn. However, the evaluations of individual Projects in the Programme must be realistic so massive numbers are unlikely to be involved except perhaps where existing or routinely collected data is being analysed. There are methods for calculating effect sizes according to sample size but this is beyond the scope of this introductory guide.

Recruitment: Participants in an evaluation are likely to be contributing time and resources and at minimum when invited to participate should receive clear information on the purpose of the evaluation, the expectations of their contribution and a commitment to them to feedback the findings.

Engagement of the young people: A small number of stakeholders and in particular young people expected to benefit from the innovation should be invited to comment on the evaluation design (e.g. is it asking the right questions), the proposed data collection and the interpretation of emerging findings. Consideration should be given to the opportunities to engage young people as researchers with appropriate training, in the evaluation in particular, in collecting data from other young people as this can sometimes increase the validity and authenticity of the information received.

Quantitative approaches

RCTs and other experimental designs: Where impact assessment is the most important factor and where feasible and practical, a randomised control trial (RCT) provides the most rigorous evaluation and should be used if at all possible. Many of the Projects in the Programme involve complex multiple interventions, some parts of which are already being implemented. In these situations, an RCT will be unlikely to be possible for evaluation of the whole Project but might be feasible on an individual element such as a discrete intervention. Where RCTs are not used, 'comparison' groups should be identified if possible and strong pre and post intervention measurements identified.

Secondary data analysis: The LAs and many of the other organisations involved in the Projects collect data routinely for example on characteristics of staff and young people and outcomes. Many of these data could be helpful in evaluating the outcomes of the Projects and consideration of which questions about the innovation these data might address should be an early step in the evaluation. LAs also have access to data on 'statistical neighbours' that might be used for comparative purposes perhaps where a RCT is not appropriate. In addition, national longitudinal datasets exist for example on looked after children, children in need and on all children's educational outcomes and these could provide useful comparative data.

Specialist assessments: Some Projects are addressing very specific issues such as mental health. These Projects might wish to utilise existing specialist assessments that attempt to measure changes over time in behaviour, mental health or other aspects. If evaluators are considering using such measures it will be helpful if they consult with other partners in the Programme such as the Evaluation Coordinator, so that the efficacy of the proposed measures can be checked. Some assessments have not been validated or their reliability established which means they may provide spurious information if used for evaluation.

Large-scale surveys using interviews or questionnaires: Interviews and questionnaires are generally regarded as qualitative approaches (see below) but they can be employed in large-scale surveys and when used in this way may produce simple frequency counts (descriptive statistics) or even more complex statistical associations that provide useful information on changes over time. Surveys are useful for collecting broad-brush information from large numbers of people at low cost (if done on-line for example) but cannot provide in-depth information and may produce information that cannot be used (due to non-responses, not following instructions or deliberate winding up).

Qualitative

Interviews: Interviews are essentially conversations that allow in-depth exploration of issues and can be used to explore the perceptions of the person interviewed about the innovation, its delivery, its effects on different groups and how it might be improved. Interview schedules range in their level of structure from most structured which have the same questions for everyone and can include questions that require both open and closed answers. Semi-structured interviews allow for prompts and follow-ups and unstructured interviews leave open the opportunities for the person interviewed to talk about whatever they want around a general theme. As such, unstructured interviews are very hard to compare across respondents and are not usually used in the kinds of evaluation involved in this Programme.

Observational or ethnographic techniques: Detailed case construction, video diaries, video stories and mobile auto-ethnography can be considered alongside traditional social research techniques to provide additional insight into people's behaviour. It usually involves observing people in their natural, real-world setting, rather than in an artificial environment with the aim of gathering insight into behaviours. Ethnographic approaches seek to record the actions and experiences of people, with less interference and prompting than interviewing. The data gathered captures actual behaviours but can also help to develop a closer understanding of lifestyles as they are lived and the factors influencing people's decision making.

Focus groups: Focus groups are essentially group interviews in which a question is posed and the group encouraged to discuss it. While useful for encouraging those reticent to talk in a one-to-one situation and for people's ideas to spark off one another, they are more difficult to control and more challenging to analyse. However, they are helpful in engaging larger numbers of people in giving feedback and are often a very productive means of getting young people's views provided they are well chaired.

Deliberative engagement approaches: these tend to take the form of reconvened workshops that can enable people to discuss more complex and controversial issues than might be possible using traditional focus groups. Respondents are provided with more information, given the time and space to fully reflect upon and engage with the issues at hand, and have the opportunity to come to informed views. The benefits include going beyond initial perceptions, and can include how they form views and debate issues in more depth.

On-line group discussions and webinars: Online via synchronous (or live) group discussions and asynchronous bulletin boards and online communities can be a useful (and a cost-effective) way of gaining a good geographic sample, and of engaging time-poor participants or those who would struggle to attend a discussion at a central location. Similarly mobile platforms now enable researchers to task participants with various activities which can be

used to augment more traditional face-to-face research as part of iterative research design; respondents can reflect on their 'real data' to have in-depth reflective discussion.

Documentary analysis: As part of an evaluation in which one source of data is triangulated against another, documentary analysis can be very helpful and minimises burdens on others. It involves considering in what documents you might expect to find relevant information about the innovation being evaluated and then seeking permission to access these in order to search out policies, records or illustrations relating to the issue being evaluated. For example, we might expect to use documentary analysis in evaluating whether bureaucratic burdens on social workers have decreased as an outcome of changing the ways in which they work.

Cost-benefit analysis

The HM Treasury (2014) guide on cost benefit analysis suggests that evaluations should consider more than just the benefit-cost ratio of the Project. Interventions should be evaluated from a range of perspectives, including qualitative feedback, strategic contribution and capacity to deliver. Beecham and Sinclair (2007) concluded that all evaluations should start from the premise that economic evaluation will be included and consider the associations between costs, needs and outcomes. The Treasury guide provides technical approaches on how to do this as does the University of Loughborough's cost calculator[1] and Beecham and Sinclair provide many examples of applications in children's social care.

[1] http://www.lboro.ac.uk/research/ccfr/research/exploring/project---the-cost-calculator-for-childrens-services.html

Further resources

- Beecham, J. and Sinclair, I. (2007) Costs and Outcomes in Children's Social Care. London: Jessica Kingsley Publishers.

- Early Intervention Foundation: http://www.eif.org.uk/wp-content/uploads/2014/06/Guidebook-final-slides.pdf

- HM Treasury (2011) The Magenta book: Guidance for Evaluation. London: HM Treasury. Available at: https://www.gov.uk/government/publications/the-magenta-book

- HM Treasury (2014) Supporting public service transformation: cost benefit analysis guidance for local partnerships. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/300214/cost_benefit_analysis_guidance_for_local_partnerships.pdf

- Morrow, V. (2009) *The Ethics of Social Research with Children and Families in Young Lives: Practical Experiences* Oxford: Young Lives

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

- Social Research Unit, Dartington, evidence standards: http://dartington.org.uk/projects/what-works-evidence-standards/

- Thomas, G. (2009) *How to do your research project*. London: Sage