Evaluability Assessment: a systematic approach to deciding whether and how to evaluate programmes and policies

Peter Craig and Mhairi Campbell
What Works Scotland (WWS) aims to improve the way local areas in Scotland use evidence to make decisions about public service development and reform.

We are working with Community Planning Partnerships involved in the design and delivery of public services (Aberdeenshire, Fife, Glasgow and West Dunbartonshire) to:

- learn what is and what isn’t working in their local area
- encourage collaborative learning with a range of local authority, business, public sector and community partners
- better understand what effective policy interventions and effective services look like
- promote the use of evidence in planning and service delivery
- help organisations get the skills and knowledge they need to use and interpret evidence
- create case studies for wider sharing and sustainability

A further nine areas are working with us to enhance learning, comparison and sharing. We will also link with international partners to effectively compare how public services are delivered here in Scotland and elsewhere. During the programme, we will scale up and share more widely with all local authority areas across Scotland.

WWS brings together the Universities of Glasgow and Edinburgh, other academics across Scotland, with partners from a range of local authorities and:

- Glasgow Centre for Population Health
- Healthcare Improvement Scotland
- Improvement Service
- Inspiring Scotland
- IRISS (Institution for Research and Innovation in Social Services)
- Joint Improvement Team
- NHS Health Scotland
- NHS Education for Scotland
- SCVO (Scottish Council for Voluntary Organisations)

This Working Paper is one of a series of papers that What Works Scotland is publishing to share evidence, learning and ideas about public service reform.

Pater Craig is a Senior Research Fellow at the MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, where he leads the Informing Healthy Public Policy Programme, and co-Director of What Works Scotland.

Mhairi Campbell is a systematic reviewer, working within the Informing Healthy Public Policy Programme at the MRC/CSO Social and Public Health Sciences Unit, University of Glasgow

June 2015

What Works Scotland is funded by the Economic and Social Research Council and the Scottish Government

www.whatworksscotland.ac.uk
Contents
Summary ........................................................................................................................................... 1
Introduction: what is evaluability assessment? ............................................................................... 2
How has evaluability assessment been used in the past? ............................................................... 2
Evaluability assessment: what are the core elements? ................................................................. 5
How might evaluability assessment be used in Scotland? ............................................................ 7
Conclusions ...................................................................................................................................... 9
Appendix ......................................................................................................................................... 10
1. Evaluability assessment: scoping review .................................................................................. 10
2. Bibliography ........................................................................................................................... 11
Summary
Evaluability assessment (EA) is a systematic approach to planning evaluation projects. It involves structured engagement by researchers with stakeholders to clarify intervention goals and how they are expected to be achieved, the development and evaluation of a logic model or theory of change, and provision of advice on whether or not an evaluation can be carried out at reasonable cost, and what methods should be used.

To date, EA has been relatively little used in the UK, but it has begun to attract attention as a way of balancing the growing demand for evaluation with the limited resource available. As well as providing a sound basis for making decisions about whether and how to evaluate before resources are committed, EA can improve the translation of research into practice by ensuring that policy-makers and practitioners are involved from the beginning in developing and appraising evaluation options.

Two EAs have recently been conducted in Scotland, which provide a model that can be applied to a wide range of interventions, programmes and policies at national, regional and local levels. What Works Scotland is keen to work with Community Planning Partnerships (CPPs) to identify opportunities for EA.
Introduction: what is evaluability assessment?

Evaluability assessment (EA) is a systematic approach to planning evaluation projects. It involves structured engagement with stakeholders to clarify intervention goals and how they are expected to be achieved, development and evaluation of a logic model or theory of change, and provision of advice on whether an evaluation can be carried out at reasonable cost or further development work on the intervention should be completed first.

Although a wide range of approaches are evident in the EA literature, a useful general characterisation of EA is as a ‘low-cost pre-evaluation activity to prepare better for conventional evaluations of programmes, practices and some policies’ (Leviton et al, 2010). EA offers value by sharpening the focus of interventions that are put forward as candidates for evaluation, and establishing the likelihood of measurable impact, before resources are committed to a full scale evaluation. It can forestall commitments to evaluate programmes where further development is required, or where there is little realistic expectation of benefit, and make the evaluations that are undertaken more useful. It also provides a basis for constructive engagement with stakeholders, whether or not a full scale evaluation is undertaken. This should encourage the translation of research findings by ensuring that policy-makers and practitioners are involved from the beginning in developing and appraising evaluation options.

To date, EA has been relatively little used in the UK, though a number of EAs have been commissioned in recent years by the Department for International Development (Davies, 2013). More recently, EA has begun to attract attention from public health researchers as a way of balancing the growing demand for evaluation with the limited resource available (Ogilvie et al, 2011). An example is the EA of the Responsibility Deal commissioned by the Department of Health (Petticrew et al, 2013). Over the past year, two EAs have been conducted on behalf of the Scottish Government (Beaton et al, 2014; Wimbush et al, 2015), and there is scope for the method to be applied much more widely, at both a national and a local or regional level.

What Works Scotland has identified EA as one of its key approaches to improving the use of evaluation and evidence by Community Planning Partnerships (CPPs) and is keen to work with CPPs to apply EA to their improvement work. This paper sets out the essentials of the EA approach, describes how it has been used to date, with a focus on examples relevant to Scotland, and suggests how it may be used in future.

How has evaluability assessment been used in the past?

The idea of evaluability assessment was first set out in 1979, by Joseph Wholey, an official in the US Department of Education and Welfare (later the Department of Health and Human Services), although the approach appears to have been in use before then (Smith, 1981). Its development was spurred by disappointment with the results of the large investment in evaluations of the US ‘Great Society’ initiatives of the 1960s. Many of the resulting studies
showed no effect, leading to a backlash against public spending on social programmes. When the evaluation studies themselves were reviewed by Wholey and others, they found that many of the programmes had such unclear goals, or were so badly implemented, that evaluation was uninformative (van Voorhis and Brown, 1997). EA, as developed by Wholey and colleagues at the Urban Institute in Washington DC, had the twin purpose of improving the quality and usefulness of the evaluation studies, and the quality and effectiveness of the programmes being evaluated: ‘Although evaluability assessment (as the name implies) explores the feasibility of programme evaluations, another important focus ... is the likely usefulness of evaluation in improving programme performance’ (Wholey, 1987).

EA was widely used by some US Government Departments in the 1970s and early 1980s, before interest waned. Its use within Government declined after Wholey left the DHHS, and it failed to attract widespread interest within the academic evaluation community. The reasons for this are not entirely clear, but since most of the early EAs were published as reports by sponsoring agencies (rather than as academic papers), and as a pragmatic approach to improving decision-making (rather than a scientific method), EA may simply have been eclipsed by other, more formal evaluation techniques (Trevisan, 2007). Interest picked up again in the 1990s, driven in part by US legislation requiring federal agencies to report on performance, and a 2010 review identified a wide range of evaluability assessments carried out at federal, state and local levels (Leviton et al, 2010). Much of this work continues to be published in grey literature reports, rather than scientific journals, although a 2007 review (Trevisan, 2007) found 22 journal papers published between 1986 and 2006.

In the UK, a review commissioned by the UK Department for International Development (DfID) identified over 70 reports of EAs, and over 50 other publications, including methodological guidance, reviews and protocols (Davies 2013). Around half of all publications identified in the search were from US Government agencies with most of the rest commissioned by international development agencies. The increased use of EA in recent years largely reflects a growth in their use in connection with the evaluation of aid projects. Our own rapid scoping review (Appendix 1) identified nearly 30 journal articles reporting EAs, and many other reports from sponsoring agencies. Most were of programmes implemented in the US or Canada (Table 1). A further 20 papers described EA methods, again largely by US-based authors. The published EAs cover interventions in a wide range of policy areas, including public health, health and social services, education, criminal justice and economic development, and levels of public administration, from national to school district. A comprehensive search would be likely to identify considerably more papers, especially in the grey literature. However, this brief survey has established that EA is a sufficiently flexible method to be used in a wide variety of settings, and is likely to be especially useful in cases where either the goals or design of an intervention are unclear, or where evaluation needs to be dovetailed with an ongoing programme of implementation and monitoring.
Table 1 Examples of completed evaluability assessments

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Setting</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Asthma Programme</td>
<td>35 US States and Puerto Rico</td>
<td>Hester et al., 2013</td>
</tr>
<tr>
<td>Public health ‘Responsibility Deal’</td>
<td>England</td>
<td>Petticrew et al., 2013</td>
</tr>
<tr>
<td>Local wellness policies</td>
<td>US school districts</td>
<td>Pitt Barnes et al., 2011</td>
</tr>
<tr>
<td>More kids in the woods initiative</td>
<td>US Forest Service</td>
<td>Zint et al., 2011</td>
</tr>
<tr>
<td>Rural immunisation improvement strategy</td>
<td>Nouna, Burkina Faso</td>
<td>Sanou et al., 2011</td>
</tr>
<tr>
<td>Pharmaceutical services for people living with HIV/AIDS</td>
<td>Brazil</td>
<td>Esher et al., 2011</td>
</tr>
<tr>
<td>Healthy Community Challenge Fund</td>
<td>England</td>
<td>Ogilvie et al., 2011</td>
</tr>
<tr>
<td>National driver retraining programme</td>
<td>Canada</td>
<td>Joanisse et al., 2010</td>
</tr>
<tr>
<td>Community college retention programme targeting African-American males</td>
<td>USA</td>
<td>Mckinney 2010</td>
</tr>
<tr>
<td>Nutrition promotion for newly arrived refugees</td>
<td>Perth, Western Australia</td>
<td>Durham et al., 2007</td>
</tr>
<tr>
<td>The Centre for Disease Control’s rape prevention and education programme</td>
<td>USA</td>
<td>Basile et al., 2005</td>
</tr>
<tr>
<td>A survivors of torture programme</td>
<td>Canada</td>
<td>Thurston and Ramaliu, 2005</td>
</tr>
<tr>
<td>Evidence-based criminal and juvenile justice programmes</td>
<td>USA</td>
<td>Chemers and Reed 2005</td>
</tr>
<tr>
<td>Three juvenile justice programmes</td>
<td>New Jersey, USA</td>
<td>Finckenaue, 2005</td>
</tr>
<tr>
<td>Community-based, multi-strategy approach to physical activity promotion</td>
<td>Toronto, Canada</td>
<td>Dwyer et al., 2003</td>
</tr>
<tr>
<td>Staff training in special care units for persons with dementia</td>
<td>Ontario, Canada</td>
<td>Johncox, 2000</td>
</tr>
<tr>
<td>Restaurant health promotion programme</td>
<td>Ontario, Canada</td>
<td>Macaskill et al., 2000</td>
</tr>
<tr>
<td>State technology development programmes</td>
<td>Georgia, USA</td>
<td>Youtie et al., 1999</td>
</tr>
<tr>
<td>Community-based prevention programmes to reduce minority over-representation in juvenile justice programmes</td>
<td>Pittsburgh, USA</td>
<td>Welsh et al., 1996; Piquero, 1998</td>
</tr>
<tr>
<td>A patient care and outcome programme</td>
<td>Canada</td>
<td>Casebeer and Thurston, 1995</td>
</tr>
</tbody>
</table>
Evaluability assessment: what are the core elements?

An EA seeks to determine whether an evaluation should be undertaken, and if so what questions it should seek to answer, and how the evaluation study should be designed. It is important to distinguish EA from evaluation itself. The goal of an EA is to inform evaluation decisions, not to determine whether a programme of intervention is effective or cost-effective.

The methods used in EAs vary, but there are a number of common core elements. They include:

- systematic engagement with stakeholders from the outset;
- elaboration, testing and refinement of an agreed theory of change;
- identification and review of existing data sources; and
- the making of recommendations for or against evaluation.

Engaging stakeholders: an important function of EA is to ensure that evaluation findings are useful for decision-makers. Involving stakeholders throughout the process means that key decisions about what form a subsequent evaluation should take are jointly owned, and reflect stakeholders’ priorities as well the practical and methodological constraints on evaluation study design. Who to involve will depend on the nature of the intervention, but typically will include both policy-makers and those responsible for delivering the intervention, and it is often useful to involve people involved in routine data gathering or monitoring of the intervention. Involving stakeholders directly, rather than relying on documentary information, should provide a more accurate, detailed and up-to-date characterisation of the goals and design of the intervention. It should also help to ensure a shared understanding and realistic expectations about what an evaluation can and cannot deliver.

Developing a theory of change: one of main motivations for developing EA approaches was to find a way of achieving clarity and a common understanding of what an intervention was intended to achieve. Setting out the goals and components of the intervention, and linking these to the intended outcomes in the form of a logical model or causal diagram is a good way of achieving such a shared understanding. A draft model can be sketched out by the researchers, based on documentary information, and then refined and elaborated either in interviews or workshop-style meetings with stakeholders. Getting stakeholders together, if it is practical, may be a more effective way of identifying and resolving uncertainties, and may be quicker and more efficient than a series of individual interviews.

Reviewing existing research literature and data sources: the focus of an evaluation will depend on what is already known about the intervention in question, and what are the most important remaining uncertainties. For a very novel or experimental intervention, the key issue may be which of a variety of models should be the focus of future development. For a well-established intervention, whose effectiveness has already been demonstrated in
some settings, the focus may instead be on refining delivery of the intervention, identifying implementation problems, or establishing whether it is effective in a new setting. Data sources will include published literature, including previous evaluations of similar interventions, policy or programme-specific documents, and routinely collected monitoring or outcome data. Access to administrative data, especially if information on exposure can be linked to information on outcomes, is often the key to an efficient, affordable evaluation design.

**Making recommendations:** an evaluability assessment is a decision-making tool, so it is important to provide a clear set of recommendations, based on the goals of the intervention identified through the theory of change work, the questions that stakeholders want to answer, what is already known about effectiveness, cost-effectiveness and implementation issues from previous research, and what data sources are available for future evaluation. Even if all these considerations support one particular approach, it is useful to present an appraisal of a range of options, including the option of not proceeding with an evaluation, so that the grounds for the recommendation are explicit and persuasive. Ideally, stakeholders should be involved in reviewing and agreeing a draft set of options before a final report is presented.

These elements are summarised in Figure 1. Although they are presented, for simplicity, as a series of sequential steps, in practice, the stages will overlap. The amount of time and effort to devote to each activity will vary from one EA to the next, depending on strength of the existing evidence base, the complexity or degree of development of the intervention, the number of key stakeholders and so on. But to be useful, an EA should be completed rapidly, ideally within three months of inception, so that the process does not hold up decisions about whether and how to proceed with an evaluation. A timetable should be agreed with stakeholders at the outset, with time built in for clearing the final report.
How might evaluability assessment be used in Scotland?

Two evaluability assessments have recently been conducted in Scotland by a consortium of researchers from the MRC/CSO Social and Public Health Sciences Unit at the University of Glasgow, the Scottish Collaboration for Public Health Research and Policy at the University of Edinburgh, and the Evaluation team at NHS Health Scotland.

**Evaluability assessment 1: implementation of free school meals**

The first EA, conducted on behalf of the Education Directorate of the Scottish Government looked at the implementation from January 2015 of free school meals for all children in the first three years of primary school in Scotland (Beaton et al 2014). It included a rapid review of the literature on free school meals, a workshop with stakeholders to inform the development of a theory of change, an assessment of the availability and quality of sources of data for monitoring and evaluating the impact of the policy, and a further meeting with stakeholders to discuss and agree the recommendations.

The theory of change workshop identified six key outcomes: increased school meal uptake, cash savings for families not already in receipt of free school meals, increased demand for food from local and sustainable sources, healthier diets, improved school behaviours and educational attainment. It also highlighted a number of unintended consequences, including impacts on other aspects of school life such as provision of PE, and impacts on school meal uptake by children in P4 to P7. The review of data identified a number of existing sources
which could potentially contribute to an evaluation, including data collected through HM Schools Inspectorate, the annual Scottish Government Healthy Living Survey and the Growing Up in Scotland study. New primary data collection was also considered, including surveys with children and school catering staff, and qualitative research with families.

Taking account of the constraints of time and cost, the EA recommended making best use of existing data. New data collection was also recommended to evaluate the implementation of the policy. In line with this recommendation, an evaluation study was commissioned to identify and measure: variations in implementation; factors contributing to differences in uptake across schools and local authorities; barriers and facilitators to implementation, and how these were overcome or used by local authorities and schools; unintended consequences of implementation, positive or negative, and whether and how schools/local authorities attempted to mitigate any negative consequences; and to suggest ways of improving the implementation of free school meals for all P1 to P3 pupils.

Evaluability assessment 2: implementation of the Family Nurse Partnership

The second EA (Wimbush et al., 2015) focused on the implementation of the Family Nurse Partnership (FNP) by Scottish NHS Boards. The FNP is an intervention developed in the US, offering intensive, structured home visiting support delivered by a specially trained nurse to teenage first-time mothers from early pregnancy until their child’s second birthday. Trials in the US and the Netherlands have shown positive impacts on a wide range of outcomes, and a large UK-based trial is due to report in 2015. FNP is implemented by NHS Boards, overseen by NHS Education for Scotland (NES). A feasibility study conducted in Lothian, where FNP was first introduced in Scotland, suggested that implementation was feasible, but no impact evaluation has yet been undertaken.

The EA was conducted on behalf of an Evaluation Research Advisory Group set up by the Scottish Government to advise on evaluation options. Three workshops were held, attended by stakeholders from the Scottish Government, NES and members of the research team, to develop an agreed theory of change for the implementation of FNP in Scotland, and to review existing data sources, including the monitoring data collected by NES, routinely collected NHS data on pregnancy and child health outcomes, survey data and previous research findings. A further meeting was held with Scottish Government and NES stakeholders to present and discuss evaluation options, following which a report was prepared for the Evaluation Research Advisory Group. The report recommended a natural experimental evaluation, using routinely collected National Health Service data to compare FNP participants with teenage first time mothers who gave birth during intervals between recruitment to FNP, coupled with a process evaluation to explore implementation issues, and an economic evaluation using a cost-consequence framework to compare outcomes with costs.
Both EAs were completed over a period of approximately three months, from inception to submission of a draft report. In each case, the intervention was well-defined, as was the existing evidence base, so the recommendations focused on evaluation options. Longer may be needed to complete an EA of a less well-developed intervention, or where the existing evidence is very sparse, and the recommendations may need to include proposals for further development work before an evaluation could usefully be undertaken.

**Conclusions**

Evaluability assessments have been used in a wide variety of settings, and applied to a wide range of interventions across a number of policy areas and levels of government and public administration. The two that have recently been completed on behalf of the Scottish Government demonstrate the potential of the method to inform evaluation planning in Scotland. So far, no EAs of CPP-level interventions have been conducted in Scotland, but it is here that there may be most scope to use the approach both to support decisions about what when and how to evaluate, and to clarify the aims and objectives of public service improvements. What Works Scotland is keen to work with CPPs to identify opportunities for EA.

EA is likely to be most useful when resources have been identified to support evaluation work, but there is uncertainty about whether an evaluation is feasible, or what methods would work best. Used well, EA can improve decision-making about whether to evaluate existing services or novel interventions and contribute to effective evaluation design. By engaging stakeholders in the process of developing and appraising evaluation options from the outset, it should also encourage good working relationships between policy makers, practitioners and researchers, and help to ensure that evaluation findings are relevant, timely and usable.
Appendix

1. Evaluability assessment: scoping review
A number of reviews of the use of evaluability assessment have been conducted, and there is a useful online bibliography at http://mande.co.uk/blog/wp-content/uploads/2013/02/Zotero-report.htm. This was last updated in 2013.

We conducted a rapid scoping review to identify key methodological papers and examples of completed EAs up to the present. We searched the following databases using the term ‘Evaluability assessment’ in title or topic: Web of Science, Scopus, Google Scholar, Social Services Abstracts, World Bank and the WHO. Searches were conducted on 6 March 2015.

The articles were screened by title and abstract to identify those likely to be relevant. Books, articles unavailable online, documents in languages other than English, or in which evaluability assessment was one of multiple methods, were excluded. After removal of duplicates we identified around 100 relevant publications. Searches of OpenDOAR and Google resulted in a very large number of additional hits, and further screening of these databases would identify a substantial number of further papers.
2. Bibliography

Theory and methods of evaluability assessment

Journal articles


**Reports**


**Uses of evaluability assessment**

**Journal articles**


Reports


