Annexes for the Building Capacity to Use Research Evidence (BCURE) Realist Evaluation: Stage 2 Synthesis Report

Table of contents

1. Terms of Reference and BCURE Logframe

ITT Volume 3
Terms of Reference for Evaluation of Approaches to Build Capacity for Use of Research Evidence

Title:	Evaluation of Approaches to Build Capacity for Use of Research Evidence

A. Introduction

1. DFID is committed to supporting research and its effective use by policy makers and practitioners. This commitment is driven by the assumption that making more effective use of evidence will enable countries to make better policy and programme decisions, ultimately enabling them to develop more rapidly and sustainably. In the past DFID has focused on the supply of high quality research, with less work done to ensure that there is a corresponding demand for research evidence in developing countries. However, emerging evidence suggests that there are significant gaps in capacity of decision makers in the south to use research effectively, which is hampering research uptake.

2. In response to these gaps, DFID has recently launched a programme called Building Capacity to Use Research Evidence (BCURE). This is a three-year £13 million programme aimed at increasing the ability of policy makers, practitioners and research intermediaries in the South to use research evidence for decision making. The overall goal of the BCURE programme is for 'Poverty reduction and improved quality of life', and its overall purpose is for 'Policy and practice to be informed by research evidence'.

3. Improving the use of research evidence in decision making is a relatively new area for donor support, meaning that the evidence base on what works is limited. Therefore, a significant component of the BCURE programme is an evaluation of both – the wider challenge of supporting evidence-based decision making and the value of the BCURE programme itself, drawing comparisons to other capacity-building programmes where appropriate. In doing so, the primary objective of the evaluation is to **help strengthen the global evidence base on whether capacity-building approaches to supporting evidence-informed policy making can be a cost effective way to reduce poverty and, if so, how can they be implemented to achieve the greatest impacts.**

4. The direct recipients of the services will be DFID's Research and Evidence Division and governance cadre. The published final report is expected to be of value to donors and practitioners in the Research Uptake community.

B. Building Capacity for the Use of Research Evidence (BCURE)

5. The BCURE programme was procured in 2012/2013 through open competition. A large number of initial proposals were received, of which twelve were selected to develop into full proposals, including theories of change, work plans and logical frameworks. Of these twelve proposals, five were selected for funding and have now progressed to the contracting stage. A sixth proposal is still under discussion.

6. Each of the five successful proposals will employ a different approach to capacity building. The five projects will begin between September 2013 and January 2014, last three years each and end between August and December 2016. Each project is worth between £1.3 and £3.4 million. Three of the projects have already been issued contracts, with the remaining two projects expected to receive contracts within the next month.

	Primary Provider	**Description**	**Focus countries**

A	Adam Smith International	Support African cabinets to implement evidence-based decision processes, focusing on post-conflict states.	Sierra Leone, Liberia and South Sudan
B	Finalising contract	African-led programme to strengthen use of research evidence for health policy making.	Kenya and Malawi
C	Finalising contract	Develop online training on use of evidence aimed at policy makers.	India, Pakistan and Afghanistan
D	INASP	Develop and implement courses on use of evidence, focusing on civil servants and parliamentarians.	Ghana, Zimbabwe and South Africa
E	University of Johannesburg	Develop and implement courses on evidence, focusing on civil servants.	South Africa and Malawi

7. A decision will be made on whether to progress with the sixth proposal shortly; further details on this proposal may therefore be shared with those bidders invited to progress to the ITT stage.

8. A short overview of each project is provided in Annex 1. The full project proposals will be will be shared with those invited to submit a full tender. The BCURE programme business case and intervention summary provides further background to the overall programme design, including the original theory of change. It can be accessed on the project pages of DFID website. This ToR should be considered as DFID's definitive thinking on this evaluation, rather than the BCURE business case.

C. Purpose, Scope and Evaluation Questions

9. The primary purpose of this evaluation is to 'strengthen the evidence base to support evidence-informed policy making in developing countries'. This assessment will help DFID and others make better choices in the future, when deciding whether and how to support and implement capacity-building programmes on evidence use. In order to make this assessment, the evaluation is expected to draw on both the BCURE programmes and the existing body of evidence related to building capacity to use evidence for decision making.

10. The secondary purpose of this evaluation is to 'evaluate the success and value for money of the BCURE projects in building capacity to use research evidence for decision making'. This assessment will help inform DFID decisions about whether to provide additional funding to these projects beyond the original three-year contract.

11. The provisional evaluation questions are:

i) What different factors influence the extent to which policy-making organisations in developing countries use research evidence for decision making?

- What organisational structures, processes and systems help or inhibit the use of evidence by policy making institutions?
- What characteristics help or inhibit the use of evidence by individuals within those organisations? Including (but not limited to):
  - Educational history (including subject focus, level of attainment, location of education, predominant pedagogical approach, etc.)
  - Existing skills or knowledge
  - Cultural or attitudinal behaviour

- What wider institutional factors support or inhibit the use of evidence by policy-making institutions, including the role of civil society?

ii) How effective are the BCURE projects in achieving their stated outcome of increasing the use of research evidence in decision making?

- In each project, what were the observable changes in …
  - organisational policies, systems or process;
  - individuals' knowledge and skills;
  - the wider institutional environment (including civil society);

… and how effective were these in increasing the use of research evidence in decision making processes?

- To what extent were these changes driven through local leadership/ownership (i.e. how endogenous was the process) and what effect did this have on the projects' effectiveness?
- What is the relative quality of support provided by the project when designing and implementing changes to organisational policies, systems and processes? Including (but not limited to):
  - How well did this support and the final changes meet organisational needs? (i.e. to what extent did the projects implement a 'best fit' approach?)
  - What is the likely medium and long-term sustainability of these changes?
- What is the relative quality of training and pedagogy in the capacity-building approach adopted by each project? Including (but not limited to):
  - To what extent to the pedagogical approaches used match with 'best practice' for supporting adult and organisational learning?
  - How well does this support meet individual learning needs? (i.e. to what extent did the projects implement a 'best fit' approach?)
- What approaches are most effective in building the capacity of local civil society organisations? Including (but not limited to):
  - How effectively did the projects increase the capacity of local civil society organisations to use effective pedagogical approaches in training?
  - How effective were multi-country networks in increasing the local capacity of civil society organisations?
- Overall, how does each project's model of capacity building relate to other models of capacity building – both within and outside of the BCURE programme – in terms of value for money?

iii) Drawing on the lessons from the BCURE programmes and other relevant interventions, what factors influence the effectiveness of capacity-building interventions in increasing the use of research evidence?

- What organisational-level changes introduced by capacity-building interventions are most effective at increasing the use of research evidence in a policy-making institution?
- What programmatic factors help or inhibit the uptake of these changes? Including (but not limited to):
  - Which roles in an organisation should capacity-building interventions target, in order to maximise the uptake of evidence in decision making?
  - How should senior decision makers be involved in designing and/or overseeing capacity-building interventions?
  - How can organisational-level changes best help support efforts to increase individual capacity to use research evidence and vice versa?
- What programmatic factors influence how effective capacity-building interventions are at increasing an individual's ability to use research evidence effectively? Including (but not limited to)
  - What pedagogical approaches to increasing individual capacity to access, appraise and use research evidence are most effective in increasing objectively measured capacity?
  - Looking at different types of capacity building (e.g. training, mentoring, secondments etc.) what features predict success in increasing individual capacity to use research?

- To what extent can a capacity-building programme influence the wider institutional environment, in order to help support the greater uptake of research evidence in decision making? Including (but not limited to)
  - How effective are efforts to strengthen civil society networks in supporting greater uptake of research evidence?
- What factors are important for the long-term sustainability of changes implemented by capacity-building interventions? Including (but not limited to)
  - To what extent do changes in individual capacity affect the overall culture of evidence use in a policy making institution?

iv) What impacts do capacity-building interventions that are specifically aimed at increasing the use of research evidence have on …

- Increasing the use of research evidence in actual policy and programme decision making?
- Improving the relative quality of policies and programmes, in comparison with other technical assistance programmes aimed at improving policy making and/or supply side research evidence interventions?[1]

12. In order to answer these questions, it is expected that the evaluation will develop a methodology or framework for measuring the degree to which research evidence has been used in policy-making process.

13. There is some scope to amend or add to evaluation questions. Short-listed bidders will be invited to suggest what (if any) changes that they would make to the evaluation questions, as part of the ITT. Further guidance on this may be provided in the ITT pack.

D. Design and Methodology

14. Those tenderers invited to submit a full tender are invited to propose an evaluation design and methodology that best delivers the purpose and required outputs. This should also cover the potential risks and challenges for the evaluation and how these will be managed. DFID has not endorsed particular methodology(ies) for the conduct of research on capacity-building programmes. We would expect a design that takes a mixed methods approach, combining primary data collection from the BCURE projects and secondary evidence synthesis and analysis from existing sources. Primary data collection in non-BCURE countries and/or interventions may be proposed.

15. Tenderers should spell out with the approach and methods which they will use. It would be helpful if bidders explain why they selected the options they propose to use and briefly outline what other options they considered, if any. Please note that we are committed to quality and rigour in line with international good practice in evaluation.

16. The successful tenderer will refine their proposal within the first six months of the contract, in consultation with DFID, the BCURE project providers and other relevant stakeholders.

17. Proposed designs should clearly show how they will address well-known challenges with evaluating the impact of capacity-building programmes aimed at long-term cultural and institutional changes. These challenges will include:
    - Complexity and time lag: The pathway from increased beneficiary skills/knowledge to embedded changes in practice can be long and complex. In addition, the duration between 1) beneficiaries

[1] Technical Assistance programmes could include sector or organisation specific support aimed at improving the relative quality and/or effectiveness of programmes or policies. Supply side research evidence interventions refer to support to online research portals and other research uptake activities.

acquiring new skills and/or knowledge, 2) the application of these skills when designing policies and programmes, and 3) benefits to poor people from improved policies can be long and variable, and may be outside the span of this evaluation. While these two challenges affect all evaluations of capacity-building programmes, they are particularly relevant to this evaluation because the BCURE projects are being implemented simultaneously with (rather than preceding) the evaluation. This means that the proposed designs should acknowledge the degree to which they expect to be able to answer the evaluation questions within the timeframe.

- Contribution/attribution: the BCURE capacity-building support may well not be the only factor impacting on the changes observed.
- Context: the evaluation will need to draw lessons from across a wide range of countries and contexts.

18. The evaluation is expected to focus on the use of research evidence in a broad sense, i.e. published academic research papers; statistical databases; 'established' (i.e. widely debated and accepted) policy papers and positions; and evaluation findings. It does not include experiential evidence (i.e. evidence based on professional insight, skills or experience) or all types of contextual evidence (i.e. evidence based on likely uptake or impact within a given community), though some type of contextual evidence may be usefully included. Tenderers are welcome to include a definition of research evidence in their proposals, where they feel this may be helpful to clarify their proposed research design and approach.

Specific requirements: evaluation design

19. The evaluation must include the development of a programme-level Theory of Change (ToC) during the inception phase. While we have not taken a view on the whether this ToC should or should not have a central role in the evaluation approach and analysis, this will be a valuable tool for DFID and other organisations considering designing or funding similar types of capacity-building programmes. At a minimum, this ToC should draw upon the initial Theories of Change presented in the BCURE business case and the five BCURE project proposals.

20. The evaluation should include at least one case study per BCURE project.

21. Secondary evidence synthesis and analysis should be conducted in line with DFID's guidance on 'Assessing the Strength of Evidence' (2013). The Literature Review should include an examination of the different analytical frameworks used to evaluate capacity for use of research evidence.

Section 1. Sources

22. Sources of data that will be used in the evaluation would, at a minimum, include:
    - **Background documentation:** BCURE business case and project proposals
    - **Secondary data and literature:** a document review and analysis of existing evidence. This should include research evidence on interventions to build capacity to use evidence. Research/evaluations carried out in low income contexts will be particularly relevant, though tenderers should also consider what lessons can be drawn from research carried out in other contexts. The analysis may also draw relevant lessons from research on related themes – for example research into effective approaches to supporting adult learning or research into organisational learning and change.
    - **Primary data gathered by the Evaluation team:** e.g. interviews with key partners and users – including face-to-face meetings – surveys or other data collection methods with beneficiaries and stakeholders.
    - **Primary data gathered by the BCURE project providers:** e.g. data from the projects' monitoring frameworks, progress reporting etc.

In choosing an approach and methods, the tenderer should as far as possible, set out the different data sources they expect to use – including types of primary data – and what weighting they would expect to attribute to data when forming their evaluation conclusions.

23. The BCURE projects will be an important source of data. The evaluation is therefore expected to work closely with BCURE project providers, in order to:
    - Support providers to suggest amendments to their draft monitoring frameworks, in order to maximise alignment with the evaluation objectives;
    - Comment on monitoring tools developed by providers, such as training assessment forms, and the information gathered from those tools; and
    - Participate in annual BCURE lesson learning meetings.

24. BCURE projects were made aware in advance of DFID's plans for independent external evaluation; good levels of cooperation can be anticipated with regard to reasonable requests to support the evaluation. Input from projects does not need to be costed.

25. Noting the volume and quality of applications to the BCURE programme, tenderers invited to submit an ITT may wish to suggest a role within the evaluation for certain unsuccessful applicants (of full proposals and/or concept notes). Further information on this will be included in the ITT information pack.

Ethics

26. The evaluation should ensure that it adheres to the ethical evaluation policies of DFID and the evaluation principles of accuracy and credibility.

E. Timing and Scope

27. The evaluation should start as soon as possible, in order to facilitate early engagement with BCURE projects. Taking into consideration logistical and procurement requirements, our anticipated start date is around April 2014. The evaluation will last approximately three years and three months (39 months), ending mid-2017. However, bidders may suggest a later completion date in 2017, where they believe that this will significantly strengthen the evaluation findings, given their research design. There is the option of a one-year extension in case of unforeseen circumstances, though DFID's strong preference is for the evaluation to conclude no later than December 2017.

28. DFID also reserves the right to scale up/scale back the evaluation programme depending on the requirements.

29. The evaluation is expected to include some assessment of project activities in all 11 of the BCURE beneficiary countries. We do not have a view as to what level of engagement in each country would be most appropriate, nor whether engagement should be split equally between all countries or focus on particular countries. The successful provider will be responsible for arranging their own logistical arrangements. However, the BCURE project providers will provide some support with identifying and contacting key contacts.

30. The primary focus of this evaluation is approaches to increase the systematic use of research evidence to inform policy making. Efforts to *influence* particular policies with a given piece of research are not the focus of this evaluation. Tenderers are welcome to include a definition of 'policies' in their proposals, where they feel this may be helpful to clarify their proposed research design and approach.

31. Capacity building/development refers to the capacity of individuals, organisations and the broader institutional framework within which individuals and organisations operate to deliver specific tasks and mandates.
32. The evaluation is expected to focus on Lower-Income Countries and those Middle-Income Countries with a high poverty burden. However, the evaluation may consider evidence from other countries where this is helpful.

F. Outputs

33. The Evaluation team will produce the following outputs:

- **Inception Report and initial literature assessment** within six months. This should include refinements/amendments of evaluation questions and full methodology; overarching theory of change; suggested amendments to the monitoring frameworks for the BCURE projects; identified sources of data and risk management strategy; communications strategy; work plan and any proposed budget revisions (within the agreed total contract value).
- **Stage 1 of the evaluation** within twelve months**,** comprising findings from secondary data and initial collection of primary data. This report should focus on evaluation question 1, though may helpfully include findings for the other evaluation questions, as available.
- **Stage 2 of the evaluation** by April 2016, comprising an initial report on evaluation question 2, in order to inform decisions on future DFID support under the BCURE programme. The exact format for stage 2 will be agreed during the inception phase. As the projects will have only completed between 28 and 32 months of their 36 month contracts, this will impose some constraint on the strength of conclusions possible at this stage.
- **Draft stage 3 of the evaluation** within 36 months (approximately December 2016), comprising a draft report of all the evaluation questions. This report will be commented on by DFID, with areas for revision and further research highlighted.
- **Final stage 3 of the evaluation** within 39 months, comprising the full report (maximum of 150 pages with a maximum six-page Executive Summary) that incorporates feedback obtained on the draft report. This report will be externally peer reviewed, to be organised by DFID.
- Appendices with details on the methodology, informants, etc.

34. DFID's intention is for the evaluation findings to be available and shared widely within the international community, in order to strengthen the evidence base in this area. This means that publication of the evaluation findings – in particular, stages 1 and 3 – will be required to comply with DFID's Enhanced and Open Access Policy. In addition, tenderers are invited to suggest how they would share findings through peer reviewed publications and other communication outputs and channels, as part of the ITT.

G . Management, Reporting and Financial arrangements

Management arrangements

35. The evaluation will be overseen by a Steering Group, who will be responsible for approving the evaluation outputs and commenting on draft reports. The steering group shall comprise:
    - Jessica Prout and Nathanael Bevan from DFID's Evidence into Action team, who are managing the BCURE programme
    - A DFID evaluation adviser and/or governance specialist not directly involved in BCURE
    - One or two external representatives

36. Day-to-day management of the study will be undertaken by Jessica Prout and the deputy programme manager of the Evidence into Action team.

Financial and Reporting arrangements

37. Bidders are invited to explain how they would link payment to results, as part of the ITT. DFID's preference would be for payment to be made against achievement of quarterly or bi-annual milestones, as a form of output-based contract. Payments must be accompanied by short technical reports, detailing progress against the milestones, work plan and budget.

38. In addition to technical reports, the successful bidder is expected to meet bi-annually with the steering group. As part of these meetings, they will be expected to deliver up to four presentations to the

steering group (one in presenting the inception report; one in presenting stage 1; one in presenting stage 2; and one in presenting the draft stage 3 report). Meetings at which the successful bidder is presenting will take place in London; other meetings will take place either in London or via telephone, depending on logistics.

39. Mandatory financial reports include an annual forecast of expenditures (the budget) disaggregated monthly for the financial year April to March. This should be updated either quarterly or bi-annually, in line with the agreed payment schedule, alongside a report of actual expenditure over the period. The successful bidder must also submit yearly external audit reports on their annual financial statements.

40. Key Performance Indicators (KPIs) will be agreed with the successful bidder during the inception phase.

Inception phase

41. The evaluation will have an inception phase of up to eight months, during which the inception report and initial literature will be finalised, submitted to and agreed by DFID. There will be a formal contract break at the end of the inception phase and DFID reserves the right to terminate the contract at that point if the work undertaken during the inception phase is unsatisfactory or agreement cannot be reached on the remainder of the evaluation (budget / detailed methodology and work plan).

H. The Evaluation team

42. Pre-Qualification Questionnaires (PQQ) from suitably qualified organisations and consortia are equally welcome. Lead organisations for the consortiums contracted to deliver the BCURE projects are not eligible to apply (as set out in 41. in the BCURE terms of reference). Other BCURE consortium members are eligible to apply, but must fully explain in an Annex to their PQQ how they would manage any conflict of interest that may potentially arise. The proposed evaluation team may not include any individual who is contracted as part of a BCURE project.

43. The supplier will design, co-ordinate and draw together the evaluation findings in a final report. They will quality assure the outputs and validate the data collected.

44. The BCURE project providers will also seek to facilitate access to stakeholders who have direct links with the programme, but the evaluation team will have to make direct approaches to other stakeholders and beneficiaries who are in scope of their evaluation design.

45. DFID welcomes proposals that:
    - Where the evaluation is being conducted by one organisation from a high income country, includes plans in the PQQ for helping to build local capacity to conduct high quality evaluations.
    - Where the evaluation is being conducted by a consortia, that this either includes member organisations from low or middle-income countries (preference), or includes plans in the PQQ for helping build local capacity to conduct high quality evaluations.

Skills and qualifications

46. As outlined in the PQQ, the essential competencies and experience that the contractor will need to deliver the work are:
    - Extensive knowledge and application of evaluation methods and techniques, preferably with experience in implementing evaluations of a similar scope and size to this ToR
    - Strong qualitative and quantitative research skills
    - A good understanding of capacity building
    - Strong analysis, report writing and communication skills, preferably with experience in publishing evaluation and/or research findings in peer reviewed publications

- Experience of engaging with Southern partners

47. Desirable competencies and experience are:
    - Experience in evaluating, research or delivering capacity-building interventions
    - A good understanding of research uptake
    - Expertise in assessing value for money

Further advice

48. Enquiries regarding these Terms of Reference can be submitted as dialogue questions via the DFID supplier portal. Where appropriate, answers to these questions will be posted and will be visible to all potential suppliers.

Duty of Care

49. The Supplier will be responsible for the safety and well-being of their personnel and Third Parties affected by their activities, including appropriate security arrangements. They will also be responsible for the provision of suitable security arrangements for their domestic and business property. The Supplier is responsible for ensuring that appropriate arrangements, processes and procedures are in place for their personnel, taking into account the environment they will be working in and the level of risk involved in delivery of the Contract (such as working in dangerous, fragile and hostile environments, etc.). The Supplier must ensure their personnel receive the required level of training and where appropriate complete a UK government approved hostile environment or safety in the field training prior to deployment.

50. Tenderers must develop their PQQ Response and Tender (if Invited to Tender) on the basis of being fully responsible for Duty of Care. They must confirm in their PQQ Response that:
    - They fully accept responsibility for Security and Duty of Care.
    - They understand the potential risks and have the knowledge and experience to develop an effective risk plan.
    - They have the capability to manage their Duty of Care responsibilities throughout the life of the contract.

If you are unwilling or unable to accept responsibility for Security and Duty of Care as detailed above, your PQQ will be viewed as non-compliant and excluded from further evaluation.

51. Acceptance of responsibility must be supported with evidence of Duty of Care capability and DFID reserves the right to clarify any aspect of this evidence. In providing evidence, interested Suppliers should respond in line with the Duty of Care section in Form E of the PQQ.

52. DFID will provide risk assessments for the relevant countries when issuing the ITT pack. Bidders will be expected to prepare Duty of Care plans as part of their technical response.

I. Budget

The budgeted expenditure for this work over a three-year period is between £700,000 and £950,000.[2] Value for money will be a key criterion in selection and the final budget will be agreed with the successful provider.

[2] The BCURE business case budgeted for up to £2 million to be split between three evaluations on research capacity building and uptake.

BCURE Joint Logframe

PROJECT NAME	Building Capacity to Use Research Evidence (BCURE) programme						
IMPACT	**Impact Indicator 1**		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	
Better design and implementation of government programmes and policies leads to reduced poverty	Worldwide governance indicator on government effectiveness	**Planned**	From 2012 dataset, listing by rank: South Sudan: 3 Afghanistan: 7 Zimbabwe: 11 Sierra Leone: 11 Liberia: 12 Bangladesh: 22 Pakistan: 23 Kenya: 35 Malawi: 38 India: 47 Ghana: 52 South Africa: 64				
		Achieved					
			Source				
	Impact Indicator 2		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	
	Inequality-adjusted Human Development Index (IHDI)	**Planned**	From 2012 dataset, listing by IDHI score South Africa: 0.629 Ghana: 0.558 India: 0.554 Kenya: 0.519 Bangladesh: 0.515 Pakistan: 0.515 Malawi: 0.418 Zimbabwe: 0.397 Liberia: 0.388 Afghanistan: 0.374 Sierra Leone: 0.359				

			South Sudan: unranked				
		Achieved					
			Source				

OUTCOME	**Outcome Indicator 1**		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	**Assumptions**
Strengthened and embedded in-country capacity (skills, systems and culture) to access, appraise and apply research evidence and data, which influences international best practice.	Changed skills and/or processes in partners have led to an increased use of evidence in policy and programme decision making, as detailed in case studies (cumulative)	**Planned**	No data available	Six case studies (one per project)	12 case studies (two per project)	18 case studies (three per project)	Evidence-informed policy leads to better decision making and greater poverty reduction.
		Achieved					
			Source				
			Project reports, verified by DFID technical leads				
INPUTS (£)	**DFID (£)**		**Govt (£)**	**Other (£)**	**Total (£)**	**DFID SHARE (%)**	
						100%	
INPUTS (HR)	**DFID (FTEs)**						
	1.5						

OUTPUT 1	Output Indicator 1.1		Baseline	Milestone 1	Milestone 2	Target (date)	Assumption
Greater use of evidence in cabinet decision making in Africa, with a focus on Sierra Leone, Liberia and South Sudan (see nested logframe 1)	Cabinet secretaries have improved ability to oversee revised Cabinet processes, as measured by: - Revised Cabinet manuals are developed and used - Tracking systems developed and used to oversee implementation of Cabinet decisions - Number of trained policy analysts (or equivalent) in Cabinet Secretariats that are able to review evidence use - Proportion of strategic* proposals that are reviewed for quality by the Secretariats	**Planned**	* Cabinet manuals out of date * No effective process for tracking implementation * No policy analysts * No proposals reviewed by Cabinet Secretariat	* Revised cabinet manuals in Sierra Leone and Liberia * New tracking systems developed for monitoring cabinet proposals * At least 3 trained policy analysts in place over 3 countries * 15% of strategic proposals are reviewed	* Revised cabinet manual in South Sudan and support in place in Sierra Leone and Liberia * New tracking system approved and being used in all 3 countries * At least 6 trained policy analysts over 3 countries * 50% of strategic proposals are reviewed	* Cabinet secretariat processes conducted in line with revised manuals * Cabinets have accurate data on implementation progress * At least nine trained policy analysts over 3 countries * 75% of strategic proposals are reviewed	1) Cabinet Secretariats have sufficient budgets and political backing to implement project activities 2) High-level support from Presidents and Ministers to agree and implement reforms, including providing the necessary staff time and resources from ministries 3) Cabinet Secretaries and other senior officials are available to participate in international workshops on given dates. 4) That political or other external events does not prevent programme implementation; in particular, that South Sudan remains stable enough to engage meaningfully in project
		Achieved					
		Source					
		Quarterly reports; Cabinet Secretariat monitoring tools and data; training records; discussions with beneficiaries					
	Output Indicator 1.2		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	
	Ministers have greater ability to interrogate the quality of proposals submitted to Cabinet, as measured by: - Proportion of strategic* Cabinet proposals that are circulated to Ministers prior to Cabinet - Cabinet committee structures implemented - Proportion of relevant Cabinet items considered by Cabinet committee - Percentage of all Ministers who participate in workshops and describe it as 'good' or 'excellent' (cumulative)	**Planned**	* Between 0 and 15% compliance with proposals circulated to cabinet members * No sub-committees of cabinet * No Ministers trained	* 15% compliance with country target for circulating cabinet proposals * Committee structures approved * 10% of Ministers attend training and rate it good or excellent	* 30% compliance with country target for circulating cabinet proposals * Committees interrogate proposals * 30% of cabinet agenda items considered by committees * 25% of Ministers attended	* 50% compliance with country target for circulating cabinet proposals * Committees functioning without external support * 40% of items considered by committees * 40% of Ministers attended	
		Achieved					

		Source					
		Quarterly reports; Cabinet Secretariat monitoring tools and data; training records; discussions with beneficiaries					
	Output Indicator 1.3		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	
	Line ministries are better able to develop evidence-informed proposals, as measured by: - Network of Cabinet Focal Persons (CFPs) in Ministries established and functioning - Percentage of Ministries with trained CFPs (cumulative) - Number of training days delivered to CFPs	**Planned**	*No cabinet focal persons (CFPs) in Sierra Leone and Liberia * 7.6% of ministries with trained CFPs * No training	* CFPS nominated * Purpose of CFPS agreed by Ministers * Training strategies agreed	* CPFs in place and supported * 60% of ministries with a trained CFP * 1,000 person training days delivered	* CFP network self-sufficient * 75% of Ministries with trained CPFs * 2,500 person training days	
		Achieved					
		Source					
		Quarterly reports; training records; discussions with beneficiaries					
IMPACT WEIGHTING (%)	**Output Indicator 1.4**		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	
20%	Project guidelines, advice and training materials are shared effectively with others, particularly African Cabinet Secretaries, as measured by: - Participants in African Cabinet Development (ACD) network who assess international activities as 'good' or 'excellent' - Number of high-level workshops held - ACD Evidence-based Policy Toolkit is developed and disseminated - Number of media articles covering programme activities (cumulative)	**Planned**	* No materials	* 35 participants in ACD network who rate as good or excellent * 1 high-level workshop * proto-type toolkit * 9 articles on programme activities, of which 6 are in beneficiary countries	* 70 (culm.) participants in ACD * 2 high-level workshops * toolkit developed * 18 news articles, 12 in beneficiary countries	110 (culm.) participants * 3 high-level workshops * toolkit upgraded and subject to at least 40 requests * 25 news articles (18 in beneficiary countries)	
		Achieved					
		Source					**RISK RATING**
		ACD reports and feedback; newspaper or electronic articles					High, given instable operating environment (South Sudan) and high levels of political buy-in required.

OUTPUT 2	Output Indicator 2.1		Baseline	Milestone 1	Milestone 2	Target (date)	Assumptions
Greater use of evidence to inform policy decisions in India and Pakistan (see nested logframe 2)	High quality assessment report completed, as measured by: - Survey and data instruments developed - Data collected and analysed	**Planned**	No available assessment	* Assessment instrument draft, piloted and refined (February 2014) * At least 250 observations * Analysis of training needs of initial training cohorts completed	* Instruments rolled out and further refined * Additional 150 observations * Preliminary data analysis from other instruments	* Instruments made public * End data set of 500 observations * End-line data analysed and assessment report complete	1) Partner organisations willingly participate in data collection and training activities 2) That training participants return to an environment that allows them to use their learning 3) Increased capacity to understand and produce evidence-based policy proposals leads to increased number of evidence-based policy proposals.
		Achieved					
		Source					
		Assessment instrument developed for the project					
	Output Indicator 2.2		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	
	Curriculum materials developed, as measured by: - Number of online modules developed and tested - Number of civil servants trained in full set of modules - Level of proficiency in technical skills - Attitudes towards use of evidence in decision making	**Planned**	No materials developed for the country contexts	* 2 modules developed (1 day training) * At least 80 civil servants * Specific measures for learning rubric developed to assess changes in trainees' technical skills and attitudes * Baseline data collected among initial training cohorts in all focus countries	* At least 120 civil servants * 6-8 modules developed (3 to 4 training days)	* At least 300 civil servants	
		Achieved					
		Source					
		Course materials developed					
	Output Indicator 2.3		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	

	Pilot projects successfully implemented, as measured by: - Number of demonstration and pilot projects selected for funding and completed due diligence process (cumulative) - Number of case studies developed, based on demonstration / pilot projects	**Planned**	No pilot projects	* At least 5 demonstration projects	* 3 pilot projects selected	* 6 pilot projects selected * 6+ case studies	
		Achieved					
		Source					
		Data and reporting on demonstration projects and pilot projects					
	Output Indicator 2.4		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	
	Policy dialogues held, as measured by: - Number of policy workshops held - Number of people attending workshops, including number of female presenters (cumulative) - Number of policy dialogue reports	**Planned**	None	* 2 policy workshops held by December 2014 * 60 attendees to workshops with 4 female presenters by December 2014 * 2 policy dialogue reports by July 2014	* 4 policy workshops held by December 2015 * 120 attendees to workshops, with 8 female presenters by December 2015 * 4 policy dialogue reports by July 2015	* 6 policy workshops held by July 2016 * 180 attendees to workshops with 12 female presenters by July 2016 * 6 policy dialogue reports by July 2016	
		Achieved					
IMPACT WEIGHTING (%)		**Source**					**RISK RATING**
15%		Records of policy dialogue workshops through quarterly reports and beneficiary feedback					Medium
INPUTS (£)	**DFID (£)**		**Govt (£)**	**Other (£)**	**Total (£)**	**DFID SHARE (%)**	
INPUTS (HR)	**DFID (FTEs)**						

OUTPUT 3	Output Indicator 3.1		Baseline	Milestone 1	Milestone 2	Target (date)	Assumptions
Improving the skills, systems and environments to use evidence within the governments and parliaments in Ghana, South Africa and Zimbabwe (see nested logframe 3)	Policy making staff from selected countries have improved skills for and understanding of Evidence-Informed Policy Making (EPIM), as measured by: - Tailored course for Civil Service Training College (CSTC) in Ghana developed and implemented - Number of public institutions participating in training in Zimbabwe - Changes to South African Government processes to increase the use of evidence - Support provided to Ghanaian and South African parliaments - Number of policy dialogues and knowledge cafés held in Zimbabwe	**Planned**	*No existing courses that support the skills for EPIM *Facilitators do not receive pedagogy training or refresher courses on a regular basis *Facilitators have not worked on courses for EIPM in the past	*MOUs signed with CSTC in Ghana and departments (where appropriate) *EIPM course content developed or adapted from existing *Trainers in civil service colleges identified	*Trainers at the CSTC receive pedagogy and EIPM training *EIPM course/modules trialled with 1 cohort	*EIPM course/modules trialled with 2 cohorts and adopted by CSTC in Ghana	1) Elections in three target countries and other external events do not result in a change of political or high-level support 2) That participants on the course return to an environment that allows them to use their new skills 3) That there is sufficient public appetite for discussions around EIPM in Zimbabwe 4) That consortium partners have sufficient skills to deliver project activities effectively
			Needs assessment demonstrates: - Lack of awareness of benefits of EIPM - Demand from policy makers for support for their staff - Lack of expertise & skills to use & manage research - Poor communication of research	* Agreement reached with 3 institutions in Zimbabwe * EIPM course content developed	* EIPM course content trialled with 3 cohorts * EIPM champions identified (at least 2 per institution) * Mentoring programme designed	* 6 EIPM champions mentored in how to improve use of evidence in their departments * EIPM course delivered to 3 institutions in Zimbabwe	

			Current state of evidence use in South African ministries to be determined through baseline survey	* Collaborating departments selected, with project engagement starting in at least one department * Improved capacity of Human and Social Research Council (HSRC) in South Africa to facilitate processes	* Approaches to improve management of the evidence base developed and reviewed * Second government department identified * HSRC share process of supporting govt departments with other consortium partners	*Lesson learning documents for work with government departments articulating the benefits of using evidence management approaches/tools *HSRC capacity developed to be able to handle future demand
			Baseline to be set following review of parliamentary research structure in year 2 (Ghana) and engagement with portfolio committee (South Africa)	Familiarisation meetings with parliament and parliamentary research directorate in Ghana	* Review of parliamentary research structure in Ghana * EIPM awareness for MPs in Ghana * Parliamentary staff trial EIPM course in Ghana * Engagement with relevant portfolio committee to explore how to scrutinise the use of evidence in the policy making process in SA	*Increased capacity of staff to use evidence + further demand for capacity building from GH parliament *Parliamentary committees engage to explore how to better scrutinise policy and the use of evidence in SA

		Zimbabwe: 2 knowledge cafés in 2012	1 Policy dialogue and 1 knowledge café in Zimbabwe	3 Policy dialogues and 1 knowledge café in Zimbabwe	*6 policy dialogues and 3 knowledge cafés, with 50% focused on issues that disproportionately impact women. *Media coverage of policy dialogues *Café and dialogues routinely attended by a wide range of stakeholders
	Achieved				
	Source				
	Annual project reports; end of project evaluation; civil service school course list; formal and informal media reports				
Output Indicator 3.2		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**
Number of case studies and other communication outputs from the small grants programme and project consortium on building capacity for research use.	**Planned**	0	4 small grant projects identified and funded	3 case studies published from small grant projects 8 projects identified and funded since start of programme	6 case studies published (cumulative)
		N/A	3 communication outputs	6 communication outputs (cumulative)	*12 communication outputs (cumulative) * Consortium symposium and learning conference held
	Achieved				
	Source				
	Blogs; case studies; annual reports				
Output Indicator 3.3		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**

	Consortium partners are better able to deliver capacity-building activities, as measured by: - Improvements in partners' systems, processes and/or staff kill levels - Demand from others for support (outside of project beneficiaries)	**Planned**	Organisational assessment demonstrates: - Partners have limited capacity (skills and experience) implementing M&E plans and strategies (Ghana and Zimbabwe) - Partners have some capacity (skills and experience) using project & financial management systems - Partners have sufficient pedagogical skills, capacity and knowledge of EIPM	* All partners have a M&E plan in place * All consortium staff who will be directly responsible for delivering training refresh their training skills.	* Partners use collaborative project management tools * Partners use M&E tools and templates to collect data	* Partners improve their capacity to develop and implement an M&E plan * Partners show clear improvement in financial and project management * Partners show improvement in their pedagogical skills and knowledge on EIPM	
IMPACT WEIGHTING (%) 20%			* Partners have limited capacity (skills and experience) designing and implementing communication plans and strategies (Ghana and Zimbabwe) * Partners have limited capacity (skills and experience) to develop and use some communications tools	* South Africa: Identification of appropriate personnel in HSRC and training by ODI in application of demand-side toolkit * Communications strategy work plan developed	* HSRC team leads on application of the toolkit in at least one Ministry * Zimbabwe partner identifies champions in key ministries for mentoring support * Ghana partner works with parliamentary resource department to	* Partners show capacity to develop and implement a communication plan * Request to support capacity building from at least one non-project department or committee in all consortium partner countries	**Risk rating** Medium: Elections are expected in all partner countries. The range (types, location and organisations) of consortium activities is spread out which should go some way to mitigating this risk. The potential impact of the risk in a specific area is high e.g. elections may

					develop training plan		impact on the feasibility of policy dialogues in Zimbabwe or change the priorities of the civil service in any one country
		Achieved					
		Source					
		Consortium inception phase capacity assessment report; members post-consortium work plan; end of project evaluation					

IMPACT WEIGHTING (%)	**DFID (£)**	**Govt (£)**	**Other (£)**	**Total (£)**	**DFID SHARE (%)**
20%					
INPUTS (HR)	**DFID (FTEs)**				

OUTPUT 4	**Output Indicator 4.1**		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	**Assumptions**
Civil servants in South Africa and Malawi have improved capacity and support to use evidence to inform policy (see nested logframe 4)	Project governance and the Africa Evidence Network, as measured by: - Number of needs assessments and partnerships with public policy and delivery partners - Core resources on capacity building developed, including new mentoring and secondment functions	**Planned**	No governance arrangements in place	* Landscape reviews and needs assessments completed * Existing resources (training materials) on capacity building and mentoring systems published * 150 members of Africa Evidence Network, participation at colloquium & use of website	To be agreed once baseline is set: number of secondments for South Africa and Malawi To be agreed once baseline is set: number of partnerships with institutions to deliver capacity-building activities	To be agreed once baseline is set	1) That mentored personnel at government levels will go on to mentor others 2) Sufficient senior-level buy-in to gain traction for reforms with ministries. 3) That participants return to an environment that allows them to use their new skills, following training/mentoring etc.
		Achieved					
		Source					

		Data collected from landscape reviews, needs assessments and other fieldwork.				
	Output Indicator 4.2		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**
	Project raises awareness of evidence-informed policy making and enhancing capacity in research use among civil servants, as measured by: - Number of training workshop places - Examples of increased use of evidence in policy documents - Improved ability of workshop participants to assessment and synthesise research	**Planned**		* Pilot workshops delivered in South Africa for 40 people (min 30% female) and learning integrated into year 2 plans * At least 1 policy paper reviewed or developed using BCURE support using research evidence in conjunction with partner agency * Engagement with senior personnel	To be agreed once baseline is set: percentage able to assess and synthesise research	To be agreed once baseline is set: number of examples of use of evidence in policy documents
		Achieved				
		Source				
		Pre- and post-training surveys, Follow-up surveys, Stakeholder interviews, Policy documents				
IMPACT WEIGHTING (%)	**Output Indicator 4.3**		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**
15%	Further support mechanisms established that enhance the application of learning among civil servants, as measured by: - Number of male and female civil servants mentored - Number of male and female civil servants seconded on experiential work placements - Case studies of good practice developed and shared	**Planned**	* 0 mentoring relationships * 0 secondments * Invited to present at review of the 2-year national policy-research-nexus meeting (4/14); Invited to contribute to annual reflection meeting of National	Five pilot mentoring relationships complete Two secondments complete Invitations to one key national-level meeting per quarter; membership of one strategic steering group	To be agreed during inception phase	* 20 women and 20 men mentored * Other targets to be agreed during the inception phase

			Evaluation Strategy (4/14); Invited to strategic review of PSPPD (5/14).				
		Achieved					
		Source					**RISK RATING**
		Mentorship reports; follow-up surveys; email records					Medium
INPUTS (£)	**DFID (£)**	**Govt (£)**	**Other (£)**	**Total (£)**	**DFID SHARE (%)**		
INPUTS (HR)	**DFID (FTEs)**						

OUTPUT 5	Output Indicator 5.1		Baseline	Milestone 1	Milestone 2	Target (date)	Assumption
Improved use of evidence for health policy in Kenya and Malawi (see nested logframe 5)	Optimised institutional leadership and capacity to enhance evidence use: - Number of leaders in MoH and parliament and evidence champions engaged to advocate for their active role in addressing bottlenecks to evidence use - Number of research evidence use sessions held at high-level symposia/meetings in MoH and parliament and health research conference/seminar - Number of sessions held at existing regional fora to promote research prioritisation - Number of activities linking policy institutions, research institutions, policy makers and researchers	**Planned**	* 0 * 0 * 0 * 0	* 22 leaders in MoH engaged (9 & 13 in Kenya & Malawi, respectively); 18 leaders in Parliament respectively (11 & 7 in Kenya and Malawi, respectively); recruited 20 evidence champions (12 & 15 in Kenya and Malawi, respectively) * 1 research evidence meeting held in Kenya; 0 held in Malawi * 1 sessions held at Directors' Joint Consultative Committee (DJCC) * 4 policy science cafés held (3 in Kenya and 1 in Malawi)	* 20 leaders in MoH engaged (10 in each country); 14 leaders in Parliament engaged (7 in each country); 20 evidence champions engaged (10 in each country) * 2 meetings held (1 health research conference in each country) * 2 sessions held (1 session at DJCC & 1 session with Health Ministers) * 4 policy science café (2 in each country); at least 80% participants giving positive assessment of the policy science cafés	* 20 leaders in MoH engaged (10 in each country); 14 leaders in Parliament engaged (7 in each country); 20 evidence champions engaged (10 in each country) * 4 meetings held (2 health research conference in each country) *5 sessions held (2 sessions with DJCC & 2 sessions with Health Ministers and 1 Best Practices forum) * 12 policy science cafés held (7 in Kenya and 5 in Malawi); at least 80% participants giving positive assessment of the policy science cafés	1) Enhanced evidence use in decision making will result in an increase in evidence-informed health policies 2) Increased capacity of mid-level policy makers to use research evidence/data in decision making will result in an increase in evidence-informed health policies 3) Effectively managing and coordinating the programme will result in its effectiveness in improving the capacity of policy makers to use or consider research evidence in their decision making processes
		Achieved					
		Source					
		To be agreed in inception phase					
	Output Indicator 5.2		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	

Enhanced capacity of mid-level policy makers in MoH and Parliament in use of research evidence, as measured by: - Number of mid-level policy makers from MoH and parliament trained in use of research evidence - % trainees reporting that the training workshop improved their knowledge and skills immediately after the training workshop and 1 year after workshop - Number of parliamentary clerks participating in UK POST internship program	**Planned**	* 0 * 0 * 0	* 40 mid-level policy makers trained (20 in each country consisting 15 from the MoH and 5 from parliament) * 80% * 2 parliamentary clerks/research officers (1 in each country); 2 briefs generated by interns; 2 workshops facilitated by interns	* 30 mid-level policy makers trained in both Kenya and Malawi * 80% * 2 parliamentary clerks/research officers (1 in each country); 2 briefs generated by interns; 2 workshops facilitated by interns	* 40 mid-level policy makers trained (20 in each country consisting 15 from the MoH and 5 from parliament) * 80% *4 parliamentary clerks/research officers (1 in each country); 4 briefs generated by interns; 4 workshops facilitated by interns
	Achieved				
	Source				
	To be agreed in inception phase				
Output Indicator 5.3		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**
Effective Programme Management and Coordination: - Number of Consortium planning meetings and DFID BCURE Partners Planning meetings held to assess progress and plan for the coming year - Number of meetings of the Programme Advisory Committee (PAC) and mid-term review of the programme in each country - Introduction of a robust financial and programme management systems	**Planned**	*0 *0 *0	*2 meetings held (1 SECURE Health Program Partners Planning meeting & 1 DFID BCURE meeting); record of programme enhancements as a result of attendance of BCURE meeting. * 6 meetings held (2 meetings for PAC (1 in each country); 4 steering committee meetings) *Financial and	*2 meetings held (1 SECURE Health Program Partners Planning meeting & 1 DFID BCURE meeting); record of programme enhancements as a result of attendance of BCURE meeting. * 6 meetings held (2 meetings for PAC (1 in each country); 4 steering	*6 meetings held (3 SECURE Health Program Partners Planning meeting & 3 DFID BCURE meeting) * 19 meetings held (3 in each country for PAC and 12 Steering committee meetings; 1 mid-term review meeting) * Efficient financial and programme management systems in place

				programme management systems procured and operationalised	committee meetings; 1 mid-term review meeting) * Financial and programme management systems monitored and evaluated		
		Achieved					
IMPACT WEIGHTING (%)		**Source**					**RISK RATING**
15%		To be agreed in inception phase					Medium
INPUTS (£)	**DFID (£)**		**Govt (£)**	**Other (£)**	**Total (£)**	**DFID SHARE (%)**	
INPUTS (HR)	**DFID (FTEs)**						

OUTPUT 6	**Output Indicator 6.1**		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	**Assumption**
Improved use of evidence in government decision making in Bangladesh (see nested logframe 6)	Government Policy formulation procedures are evidence based, as measured by: - Policy development procedures produced centrally which mandate the use of evidence - Methodologies, guidelines and templates to support the evidence-based policy development procedures are produced	**Planned**	Current procedures do not mandate this and documents do not support evidence-based approach	To be confirmed during inception phase	To be confirmed during inception phase	Target ministries adopted procedures and guidance	1) There is sufficient senior-level buy-in to gain traction with Ministries for training 2) Local research organisations are able and willing to work with government ministries 3) Senior-level buy-in from Cabinet Secretary and
		Achieved					
		Source					
		To be agreed in inception phase					
	Output Indicator 6.2		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	

	Improved ability in line ministries to follow evidence-based policy formulation process, as measured by: - Number of policy proposals produced in target line ministries which incorporate evidence in their development - Scores of Line Ministry officials on pre- and post-training tests for training on ex-ante assessments and evidence literacy	**Planned**	0 officials achieving a 25% increase	Milestones on policy proposals to be agreed during inception phase 30 officials achieve 25% increase	Milestones on policy proposals to be agreed during inception phase 60 officials achieve 25% increase	Milestones on policy proposals to be agreed during inception phase 90 officials achieve 25% improvement on their capacity to use evidence	Ministers to agree and implement government-wide processes and systems to increase use of evidence
		Achieved					
		Source					
		To be agreed in inception phase					
	Output Indicator 6.3		**Baseline**	**Milestone 1**	**Milestone 2**	**Target (date)**	
	Greater collaboration between line ministries and local research providers, as measured by: - Number of policy proposals in target line ministries which featured evidence or inputs from local research providers - MoUs signed between target line ministries and local research providers	**Planned**	To be confirmed - Based on number of proposals in target line ministries that include evidence or inputs from local researchers	Baseline +5 MOU milestones to be agreed during inception phase	Baseline +8 MOU milestones to be agreed during inception phase	Baseline +10 MOU milestones to be agreed during inception phase	
		Achieved					
		Source					
		To be agreed in inception phase					
IMPACT WEIGHTING (%)	**Output indictor 6.4**	**Planned**					
15%	Research is made available on factors which influence the uptake of evidence-based policy making within each of the line ministries, as measured by: - Assessment frameworks are developed for each target line ministry - Assessment frameworks are applied at mid-point and end point of support to target line ministry	**Planned**	To be confirmed - based on assessment frameworks which will be developed for each ministry	3 frameworks	6 frameworks developed	6 frameworks and assessments undertaken	
							RISK RATING
		To be agreed in inception phase					Medium

2. BCURE programme-level management and learning

The BCURE programme is managed through an overarching logical framework that aggregates the component programmes (see Annex 1). The individual BCURE programmes each have their own logframes and programme managers (from DFID's Evidence into Action team). The portfolio is not expected to work as a 'sum of the parts' programme. However, all the implementing partners and their DFID programme managers share learning from their programmes on strategies and approaches (e.g. training curricula) and collaborate if appropriate.

Programme teams participate in an annual learning event facilitated by DFID, supported by an online communications platform, managed by DFID.[3] The BCURE evaluation also feeds into the cross-programme learning by sharing findings at the learning event. DFID staff lead and facilitate the internal learning and knowledge exchange aspects of the programme. The evaluation team leads on communicating the evaluation findings with a wider audience to promote uptake and use.

3. Evaluation design and methodology

3.1 Introduction

The BCURE interventions work in complex government contexts, with myriad contextual conditions influencing potential outcomes: diverse historical institutional trajectories; variety in political and economic conditions, government systems and organisational cultures; and a wide range of participant characteristics (individuals' identities, gender and ethnicities). A realist approach was selected for the three-year BCURE impact evaluation because DFID was interested in understanding not just *whether* BCURE worked, but *how and why* capacity building can contribute to increased use of evidence in policy making in these different types of contexts. The primary aim of the evaluation is to strengthen the evidence base on how capacity building can promote EIPM, to inform decisions within and beyond DFID about whether to fund and how to design this type of programme in the future (see Annex 1).

Realist evaluation works through opening up the 'black box' between interventions and outcomes, through developing and testing *programme theory* (an explanation of how, why and in what contexts interventions lead to particular outcomes). Programme theory consists of linked sets of hypotheses about the mechanisms that cause an intervention to work or not work in particular contexts, to lead to specific outcomes. These hypotheses are known as 'context–mechanism–outcome' (CMO) configurations (see Box 1) – the core analytical units of realist evaluation (Pawson & Tilley, 1997; Wong et al., 2013).

Box 1: Context, mechanism and outcome

Mechanisms are the causal forces, powers, processes or interactions that generate change within an intervention – including the choices, reasoning and decisions that people make as a result of the resources provided by the programme. An intervention such as a training course is not a mechanism. The mechanism is the 'thing' that explains *why* training changes behaviour (or does not) in a particular setting.

Mechanisms are triggered only in certain **contexts**. Contextual factors may include *individual* characteristics that affect how people respond to opportunities (e.g. gender, ethnicity, education); *interpersonal* factors that affect trust and buy-in (relationships between stakeholders and programme implementers); *institutional* factors (the rules, norms and culture of the organisation in which the intervention is implemented); and *infrastructural* factors – the wider social, economic, political and cultural setting of the programme (Pawson & Tilley, 2004).

Outcomes refer to intended and unintended short-, medium- and long-term changes resulting from an intervention.

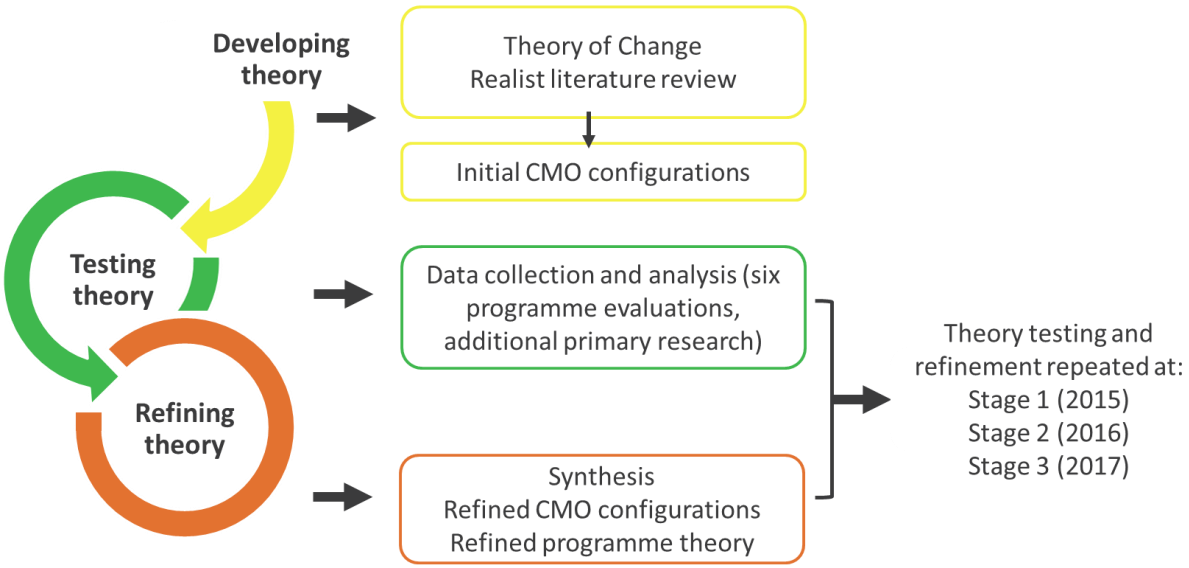
(Source: Pawson & Tilley, 1997; Westhorp, 2014; Punton et al., 2016b)

[3] For the BCURE blog, please see https://bcureglobal.wordpress.com/

Central to realist evaluation is recognition of the fundamental importance of context in shaping how and why programmes succeed or fail to lead to change. Programmes cannot simply be replicated in diverse contexts and expected to result in the same change, because contextual differences will cause the resources provided to trigger different mechanisms and lead to different outcomes. By providing insights into how and why programmes work (or do not work) in different contexts, realist evaluation can help implementers learn how best to scale up or roll out a programme (Westhorp, 2014).[4]

Realist evaluation encompasses three broad stages: developing theory, testing theory and refining theory. These are iterative rather than linear; theory is developed, tested, refined and tested again as knowledge accumulates. Figure 1 provides an overview of the evaluation design.

Figure 1. BCURE evaluation design



The evaluation began by designing a common theory of change (CToC). This was later developed into a realist programme theory, and used to develop initial CMO configurations to test during data collection. The CToC, programme theory and CMOs are discussed in Section 3.2.

The CToC was used to define the evaluation questions, outlined in Section 3.3. It also shaped the design of the main evaluation components:[5]

1. Six programme evaluations of BCURE-funded projects, incorporating primary data collection within one country (the 'country case study'), and analysis of monitoring and implementation documents from all country contexts (see Section 3.4).
2. A realist literature review, synthesising published papers and grey literature related to capacity building for EIPM (Section 3.5).
3. An 'impact case study', consisting of additional primary research on a similar intervention to BCURE that had been running for a longer period and therefore closer to seeing 'impact', in order to provide evidence on how capacity building for EIPM contributes to improvements in policy quality (the ultimate goal of the BCURE programme) (Section 3.6).

[4] This explanation of realist evaluation is adapted from a CDI paper published by the BCURE evaluation team: Punton et al. (2016b).

[5] During the initial stages of the evaluation an additional component was proposed: a series of 'non-BCURE case studies', examining other interventions that were either comparable with or complementary to the BCURE projects, in order to help strengthen the evidence base around how different capacity building interventions affect different people in different settings. However, there were a number of challenges in identifying and conducting meaningful non-BCURE case studies. Interventions selected as case study subjects needed to have relatively similar aims and approaches to BCURE in order to help test our theory; and the BCURE team also required sufficient access to stakeholders and to outcome data in order to draw meaningful conclusions about what happened and why. A pilot case study was conducted in 2015, and it was decided that the value added was insufficient to justify further investment in additional cases. In 2016, the evaluation steering committee agreed that in Stage 2 the resources would be re-allocated to the impact case study.

4. A synthesis of findings from the above components, investigating how and why capacity building for evidence use works or does not work in different contexts (Section 3.7).

Data collection and synthesis is repeated each year for three years to enable the evaluation to track programme results over time, and iteratively test and refine our theories about how and why particular outcomes have occurred in different contexts.

3.2 Developing and refining theory

The BCURE evaluation began by articulating an overarching CToC for the programme. The first iteration of the CToC drew on the evaluation team's existing knowledge – and professional hunches – about the nature of capacity building, and how capacity building can contribute to EIPM. The CToC followed a logic-model approach that helped bring the BCURE programmes into a single framework. It describes a set of propositions about building capacity for EIPM that sketch out the short- to long-term process of change that the BCURE programmes are seeking to influence. The full Stage 1 CTOC and diagram are presented in Annex 5. In summary our Stage 1 CToC was as follows:

Box 2: BCURE Common Theory of Change

Developing the capacity of decision makers to use research evidence (by building knowledge, skills, commitment, relationships and systems at four levels: individual, interpersonal, organisational and institutional) will allow them to access, appraise and apply good-quality evidence more effectively when forming policy. This will improve the quality of policies, ultimately benefiting more poor people.

The four levels of capacity change outlined in Box 2 provide a central framework for the evaluation. They convey the concept of capacity development as multidimensional, and capacity as a function of different factors and processes working together and reinforcing each other at:[6]

1. *Individual level:* individual behaviour (decisions and actions) in relation to EIPM, and the skills, knowledge, motivation, attitudes, commitment, values and personal incentives that affect this.
2. *Interpersonal/network level:* the relationships between individuals and groups that affect evidence interpretation and use, including formal and informal communities (or networks) of individuals or organisations.
3. *Organisational/government* level: an organisation's systems, policies and procedures, practices, culture or norms, which incentivise or inhibit evidence access, appraisal and application in decision making. This includes 'system-level' factors within government that affect EIPM, such as national or subnational laws, policies, regulations, governance systems and 'institutional rules of the game'. Our definition of 'government' includes government administration and parliamentary scrutiny functions (including elected opposition politicians).
4. *Institutional level:* the broader enabling environment for evidence use *outside* of government, including the role of external actors such as international donors, civil society and the media, and the influence of external factors such as crises, global events, socioeconomic change, as well as broader societal factors that influence EIPM, such as culture, norms, collective beliefs, attitudes and values. This includes the institutional role of the BCURE partners themselves within their national contexts.

Our CToC states that multidimensional change across these four domains will contribute to change in the *quality of policy development processes.* The BCURE literature review highlights an inherent tension between

[6] There are many definitions used in the literature to describe levels of capacity change. We have adapted DFID's definitions from the 2010 'How to Note on Capacity Building in Research' (DFID, 2010). This document uses 'institutional' to denote 'changes in the rules of the game'. Other readers may interpret 'institutional' to mean 'systemic' or 'environmental' change. We have opted to consider the government system as falling within a broadly conceived organisational change category because organisations within the government system are bound by common, cross-cutting rules, incentives and procedures. This means 'institutional' change then encompasses all non-governmental influences within the wider environment. However, we recognise that the boundaries between the levels of change are fuzzy and dynamic, and we consider the implications of these dynamics in our analysis.

approaching EIPM as a complex system that is infused with power and politics (which is difficult to reconcile with 'rational' concepts such as 'policy quality'), and the basic premise of the BCURE programme that better and more routine use of evidence leads to better quality policy development. This suggested the value of adopting an iterative approach to the measurement of 'policy quality;' and engaging critically with this concept over the course of the evaluation. We have drawn on Newman et al.'s (2012) definition of policy quality, along with insights from the theoretical literature on EIPM discussed in the literature review (e.g. Nutley et al., 2002; Sutcliffe & Court, 2005; Jones, 2009; Broadbent, 2012; du Toit, 2012) to develop the working definition in Box 3.

Box 3: Working definition of 'policy quality'

A policy development process can be considered to be 'good quality' if:

- Multiple types of evidence were considered in the process – including but not limited to research evidence (e.g. also including public opinion, process and practice knowledge, critical and reflective knowledge).
- The quality of evidence was seriously considered (in a way that took into account standards of evidence, while also accepting the limitations of evidence hierarchies).
- The process of decision making involved engagement with evidence (accessing it, appraising it, discussing it)…
    - …at multiple points …
    - …with multiple stakeholders with different viewpoints and perspectives…
    - …in a way that enabled real debate and discussion on the issues raised by evidence…
    - …and where evidence had a demonstrable influence on the decisions made (thinking beyond 'instrumental' influence to also consider less direct pathways of influence, for example on how people conceptualise issues).

The CToC was used to shape the research questions for the BCURE literature review, from which initial CMO configurations were developed. The literature review (discussed in Section 3.5) identified theories in the wider literature about how capacity building can contribute to EIPM. These were used to develop our first iteration of CMO configurations – hypotheses about how and why BCURE interventions might lead to different outcomes in the CToC, and how these outcomes might link to, catalyse and reinforce one another. Stage 1 of the evaluation began to test and refine these CMOs (presented in Annex 4). The refined theories were then tested and further refined during Stage 2, and are presented in Section 4 of the main report.

When operationalising a realist design, realist evaluators have identified a recurring conceptual challenge in differentiating between the *mechanism* and the *intervention* (Dalkin et al., 2015). To clarify this difference, we decided to incorporate features of the intervention as an additional element to our CMO configurations for BCURE in order to separate out features that are inherent in or under the control of the programme (such as training design or length) from contextual factors that are not (such as professional incentives to participate in the training) when considering what might 'spark' a particular mechanism. This gives us the formulation C+I+M=O (CIMOs), used throughout the synthesis report.[7]

At Stage 2, we have developed our CToC into an explicitly realist programme theory. A realist programme theory explains '(some of) how and why, in the 'real world', a programme 'works', for whom, to what extent and in which contexts' (Wong et al., 2016). This moves beyond the 'logic model' approach of the CToC to focus explicitly on the underlying mechanisms providing the active ingredients to 'spark' change (Blamey & Mackenzie, 2007). In order to move from CToC to programme theory, we developed our thinking around a) the resources provided through BCURE activities and b) the 'entry points' of the BCURE programmes.

Through seemingly different interventions, all the BCURE programmes provide seven main types of 'resources'. A common way of conceptualising mechanisms within realist evaluation is to consider 'resources and reasoning' – how different actors respond to the resources, opportunities and constraints provided by the

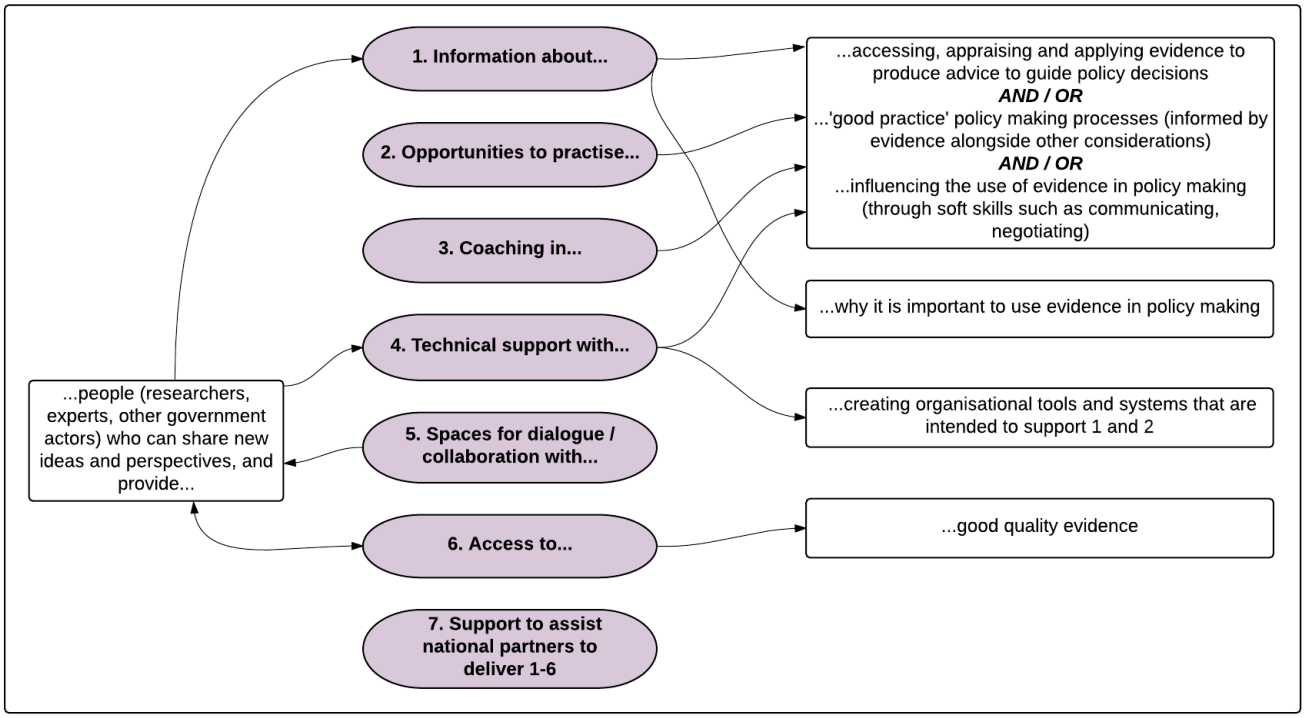
[7] At Stage 1 and during Stage 2 data collection we phrased these components in a slightly different order: I+C+M+O. We have amended this for conceptual reasons: without the right contextual factors, the intervention will not spark the mechanism (even if 'well designed'), so we decided to put context first. This decision is purely presentational.

programme (Westhorp, 2014). Thinking about resources helped us unpack generic terms such as 'training' and 'mentoring', which mask very different approaches undertaken by different partners. The main resources provided by BCURE are as follows (and depicted in Figure 2):

1. **Information:** this includes both technical information about how to access, appraise and apply evidence and normative information about the role that evidence *should* play in policy making. Information is often provided within training courses, alongside…
2. **Opportunities to practice skills** – in particular technical skills through practical exercises within training courses. In some cases, interventions provide opportunities for participants to 'learn by doing' by applying evidence appraisal or policy development skills within specific policy making processes.
3. **Coaching:**[8] personalised, tailored, ongoing and hands on support to an individual or a group from either an individual 'mentor' or an organisation, in order to help build technical EIPM skills, soft skills that help mentees use evidence more effectively in their work or skills to make decisions using evidence within 'good practice' policy making processes.
4. **Technical support:** provision of advice, consultancy, expertise or an 'extra pair of hands' to help produce a specific evidence-informed process or product, or to help design/facilitate tools or systems to promote EIPM. There is a somewhat blurry line between coaching and technical support, but one defining feature of coaching is its interpersonal element: it involves a personal relationship between the coach and the recipient or 'mentee.'
5. **Spaces for dialogue and collaboration:** either formal or informal 'spaces' (in the form of events, courses, online platforms, etc.) that bring different actors together to discuss and debate issues relating to EIPM. This can lead to further resources being provided as a result of collaborations developing out of these interactions, such as access to good-quality evidence, and further information provision and/or technical support from fellow participants.
6. **Access both to evidence and to people who can support EIPM.** This resource is usually embedded in activities that provide resources 1-6 – for example technical support and spaces for dialogue often provide access to experts or researchers who can help provide relevant evidence to inform policy decisions; training courses that primarily provide information and opportunities to practice skills may also signpost trainees towards where to find good quality evidence as well as providing access to experts in the form of speakers or facilitators.
7. Some BCURE partners provide a further resource: **support to national partners** (e.g. as part of the implementing consortium) to build their organisational capacity to deliver 1–6.

[8] The term 'coaching' is used because several BCURE partners are running activities under the label of 'mentoring', but this comprises a variety of quite different types of intervention, some of which are more akin to technical support.

Figure 2. Resources provided by BCURE



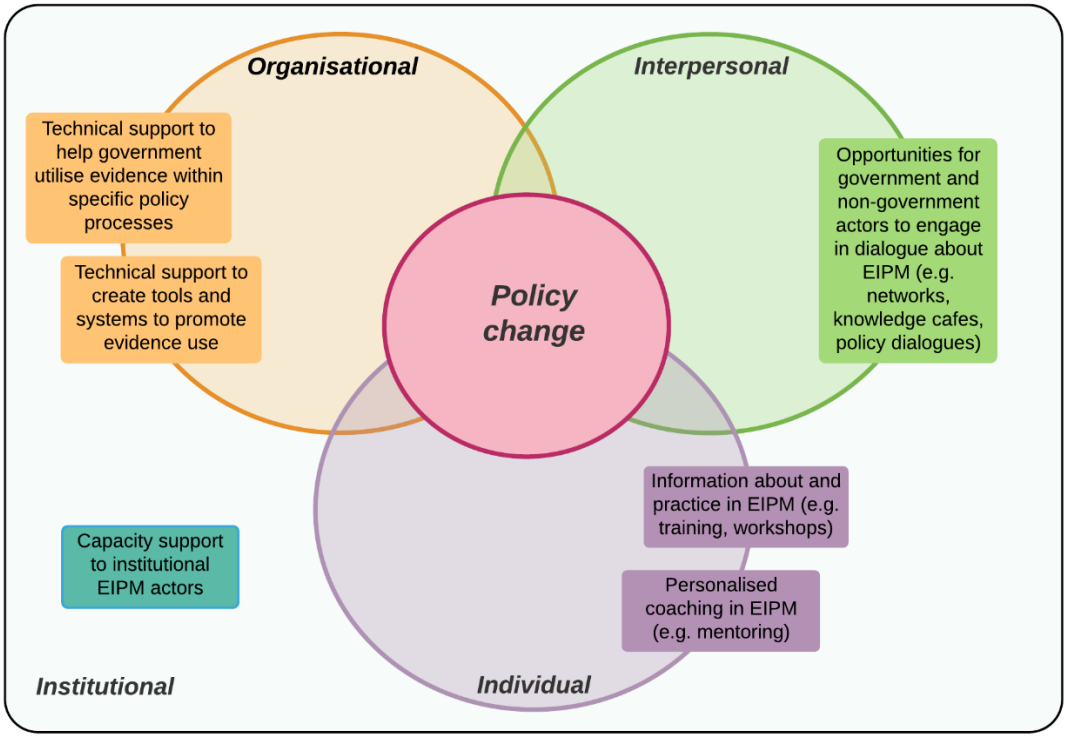
The BCURE interventions introduce resources through different 'entry points' at different levels. They might initially target *individuals* with information and opportunities to practise skills, provide spaces for dialogue between different *groups* of stakeholders, deliver technical support to *organisational* systems and processes and/or develop the capacity of *institutional* actors to promote EIPM (summarised in Figure 3).

Interventions at one level begin to influence outcomes in other domains of change, which then start to combine and reinforce each other.[9] Our findings at Stage 2 suggest that the timing and sequencing of entry points is important and different approaches may be more suitable in different contexts – discussed further in Section 5 of the main report.

To reflect this dynamic evolution, we changed the visual representation of our theory. Abandoning the conventional left-to-right logical progression of our CToC, we instead developed a diagram depicting intersecting domains of change (Figure 3). This shows how different resources introduced by BCURE at entry points contribute to outcomes at different levels, and allows us to visualise how change at each level catalyses change at other levels.

[9] For example, VakaYiko began work in Zimbabwe by developing and delivering an EIPM training course to selected ministries and parliament (individual level), alongside various networking events that brought together government stakeholders with external actors (interpersonal level). This led onto the development of a mentoring scheme, which involves technical support to tools and systems within the targeted organisations to promote evidence use, by working with selected mentees who attended the EIPM training (organisational level). Conversely, the ACD programme in Sierra Leone began at organisational level, by developing tools and systems to promote better policy making with evidence use at its heart. As this process developed, training and networking activities were designed to support individuals to use the new tools effectively. The consequences of different entry points are discussed further in Section 4 of the main report.

Figure 3. Different entry points of the BCURE interventions



Our Stage 1 programme theory brings together the evidence on resources and entry points and knits the Stage 1 CIMOs into a coherent story. This is the theory we started with at the beginning of Stage 2, based on evidence from Stage 1 data collection. It was further tested and refined at Stage 2, as detailed in Sections 4 and 5 of the main report.

Annex 4 presents the full programme theory at Stage 1, the detailed CIMOs that underpin it and how these were refined at Stage 2 based on the evidence collected through the country case studies. Annex 4 also presents the refined programme theory at the end of Stage 2 of the evaluation.

3.3 Evaluation questions

The BCURE evaluation addresses two overarching evaluation questions (EQs). These are based on the questions posed in the Terms of Reference (Annex 1), revised in the inception phase following discussions with DFID.

1. How effective are the BCURE projects in achieving their stated outcome of increasing the use of evidence in public sector decision making, and influencing longer-term changes in policy quality?
2. How and why does capacity building for evidence use work/not work, for whom, to what extent, in what respects and in what circumstances?

The initial evaluation framework identified ten evaluation questions underlying the two overarching EQs, which were designed to test different parts of the CToC. This proved to be unwieldy, and the framework was streamlined for Stage 2. It was decided to focus on five questions, built around the four domains of capacity change (individual, interpersonal, organisational and institutional) within our programme theory, with a fifth question relating to policy quality as follows:

- Five evaluation questions that relate to the five outcome domains in the CToC (individual, interpersonal, organisational, institutional and policy quality change).
- Exploring causal explanations for the specific outcomes identified within each outcome domain identified through Stage 1, as well as other outcomes: positive/negative, intended/unintended.
- Additional cross-cutting themes, including:
  - Gender and inclusion issues.

- Changes in government and political contexts and implications for programmes.
- Perceptions of role of EIPM in BCURE contexts.
- Value for money issues.
- Lessons for EIPM capacity building programme commissioners, implementers and evaluators.

Our Stage 1 CIMOs were then aligned with the EQs, based on which domain of capacity they helped explain.

Table 1 presents the evaluation framework for Stage 2.

Table 1. BCURE Stage 2 evaluation framework

BCURE Stage 2 evaluation framework			
Evaluation questions	**Outcomes to be explored**	**Data and sources**	**Analysis**
EQ 1. How and why did BCURE contribute to individual-level change? 1.1 What outcomes were achieved? 1.2 How did the interventions lead to outcomes? *(Testing Stage 1 CIMOs 1, 2, 3, 4, 5)* 1.3 How sustainable were the outcomes? 1.4 What was BCURE's contribution to the outcomes?	Increased awareness and enthusiasm. Increased knowledge and skills. Behaviour change: people using evidence more and more effectively in their work. Other individual-level outcome(s) – positive or negative, intended or unintended.	Primary data from country case study: - Individual semi-structured interviews with intervention participants. - Key informant interviews with individuals in different roles in government and outside, involved and not involved. - Document, data and content reviews, e.g. policy products, government organisational procedures. Secondary data from BCURE partner: Pre- and post-tests or surveys, follow-up surveys, case studies, interviews, written examples/evidence of behaviour change (correspondence, memos, policy briefs), results trackers.	Assessment against rubrics for: - Extent of change. - Programme's contribution to change. - Quality and strength of evidence. CIMO analysis Contribution analysis Analysis of cross-cutting themes
EQ 2. How and why did BCURE contribute to interpersonal-level change? 2.1 What outcomes were achieved? 2.2 How did the interventions lead to outcomes? *(Testing Stage 1 CIMOs 7, 8)* 2.3 How sustainable were the outcomes? 2.4 What was BCURE's contribution to the outcomes?	New champions for EIPM. Improved relationships to promote EIPM. Other interpersonal-level outcome(s) – positive or negative, intended or unintended.	Primary data: as above. Secondary data from partner: - Follow-up surveys, case studies, interviews, written evidence of 'championing' in action (correspondence, examples of initiatives started or groups convened). - Networking event write-ups, post-event surveys, documents demonstrating new initiatives or improved relationships.	Assessment against rubrics for: - Extent of change; - Programme's contribution to change; - Quality and strength of evidence. CIMO analysis Contribution analysis Analysis of cross-cutting themes

EQ 3. How and why did BCURE contribute to organisational-level change? 3.1 What outcomes were achieved? 3.2 How did the interventions lead to outcomes? *(Testing Stage 1 CIMOs 6, 9, 10, 11, 12 13, 14)* 3.3 How sustainable were the outcomes? 3.4 What was BCURE's contribution to the outcomes?	New/improved organisational tools and systems. High-level buy-in and support. Other organisational/ institutional-level outcome(s) – positive or negative, intended or unintended.	Primary data: as above. Secondary data from partner: - Documents describing design/development/ evolution of new tools and systems. - Written evidence of high-level buy-in, e.g. correspondence, minutes of meetings, examples of new initiatives with high-level support, media reports.	Assessment against rubrics for: - Extent of change. - Programme's contribution to change. - Quality and strength of evidence. CIMO analysis Contribution analysis Analysis of cross-cutting themes
EQ 4. How and why did BCURE contribute to institutional-/system-level change? 4.1 What outcomes were achieved? 4.2 How did the interventions lead to outcomes? *(No CIMOs identified in Stage 1; to be developed at Stage 2)* 4.3 How sustainable were the outcomes? 4.4 What was BCURE's contribution to the outcomes?	To be determined	Primary data: as above. Secondary data from partner: - Documents describing design/development/ evolution of new tools and systems. -Written evidence of high-level buy-in, e.g. correspondence, minutes of meetings, examples of new initiatives with high-level support, media reports.	Assessment against rubrics for: - Extent of change. - Programme's contribution to change. - Quality and strength of evidence. CIMO analysis Contribution analysis Analysis of cross-cutting themes
EQ 5. How and why did BCURE (and similar EIPM capacity building interventions) contribute to changes in policy quality? 5.1 What outcomes were achieved? 5.2 How did the interventions lead to outcomes? *(No CIMOs identified in Stage 1; to be developed at Stage 2)* 5.3 How sustainable were the outcomes? 5.4 What was BCURE's contribution to the outcomes?	Better use of evidence in policy process/policy documents. Other policy-level outcome(s), positive or negative, intended or unintended	Primary data: As above and from impact case study country. Secondary data: formal and informal documentation which demonstrate use of evidence - e.g. policy briefs, correspondence, media reports, public speeches or statements; as well as draft and final policy documents.	Assessment against rubrics for: - Extent of change. - Programme's contribution to change. - Quality and strength of evidence. CIMO analysis Contribution analysis Analysis of cross-cutting themes

3.4 Programme evaluations

The central pillar of the BCURE evaluation is the series of annual programme evaluations, examining each of the six BCURE projects.

3.4.1 Approach

Each year, six programme evaluations are conducted. Each evaluation consists of:

- An independent review of secondary monitoring data and implementation/strategy documents produced by the project team.
- A country case study, involving primary data collection by the evaluation team within one of the countries targeted by the project.

The programme evaluations perform two functions:

1. Producing an internal management report for each project, intended to verify outcomes identified by the BCURE programme monitoring data (and identify additional outcomes), capture key lessons and recommendations including regarding sustainability and generate an assessment on programme effectiveness and contribution that can inform decision making about the programme.
2. Collecting data on how and why different BCURE activities have contributed to different patterns of outcomes, in order to contribute to identifying, testing and refining theories about how and why BCURE interventions were able/unable to lead to change.

3.4.2 Selection of country case studies

The BCURE programmes are working in 12 countries.[10] The evaluation is only able to cover six. The country case studies were selected using case replication logic (Yin, 2003). Country contexts were grouped into three broad case types based on a typology of anticipated contextual conditions:

1. Typical: where the contextual conditions are mixed but could offer some degree of political stability and established institutions to support EIPM.
2. Challenging: where the contextual conditions could, according to preconceived assumptions, create difficulties for introducing EIPM.
3. Favourable: offering, on first viewing, the most favourable conditions for EIPM – for example a high degree of stability, ordered institutional practices, a good degree of political openness.

Pragmatic considerations of security and access also informed the final selection. Table 2 gives an overview of the countries and the reason for their selection.

Table 2: Country case study selections

BCURE country case study	Case replication logic
Harvard BCURE: Pakistan	The Stage 1 case study focused on India: 'favourable' case (*literal replication*). However, in 2016, activities ceased in India as a result of a refocusing of the UK government's relationship with the country. Pakistan was selected as a replacement as it is the main alternative focus of the Harvard programme. Pakistan is a 'challenging' case (*theoretical replication).*
UJ-BCURE: South Africa Impact case: South Africa	'Favourable' case (*literal replication).*
SECURE Health: Kenya	'Typical' case (*literal and theoretical; both similar and contrasting results possible).*

[10] Uganda was added in Stage 2 through an extension granted to the VakaYiko programme.

ACD: Sierra Leone (though Stage 1 evaluation data collection will be difficult)	'Challenging' case (*theoretical replication).*
ECORYS: Bangladesh	'Typical' case *(both similar and contrasting results possible).*
VakaYiko: Zimbabwe	'Challenging' case (*theoretical replication*).

3.4.3 Selection of informants for the case studies

In a realist evaluation, decisions about sampling are driven by a consideration of who the researchers need to talk to in order to test their theory. Our Stage 2 sampling approach was therefore purposive, built around the programme theory and the CIMOs generated at Stage 1. Where possible, we aimed to speak to respondents interviewed at Stage 1 again, in order to help build up a picture of emerging outcomes over time – but only if these respondents had continued to play a role in the project and were likely to be able to provide further insights into the programme theory.

An initial sample was generated by the programme evaluation lead following a review of programme documents and stakeholder lists provided by the programme. The sample was purposive, with stakeholders selected based on their involvement in the Stage 1 fieldwork (and continued involvement in the programme since), relationship to the BCURE programmes, role in the government system and relationships to each other. Where possible, samples included trainees and their line managers or colleagues, or mentees and their mentors, in order to triangulate insights. Programme evaluation leads were provided with sampling guidance and a sampling matrix to ensure a systematic approach was taken across the six evaluations– see Supplementary Annex. Each country case study consulted between 25 and 30 stakeholders.

The initial sample was then shared with BCURE partners and revised with their input (e.g. where they suggested important new stakeholders to speak to or indicated certain stakeholders were no longer involved). The final selection of participants was made by the programme evaluation team and, where relevant, involved some random selection in order to mitigate any potential bias that may have arisen from BCURE partners recommending stakeholders (e.g. where there were large cohorts of trainees some were randomly selected).

Selected respondents were then emailed, and the purpose of the interview was explained. In some cases, BCURE partners facilitated introductions and helped follow up to secure appointments.

The sampling process was very much an iterative one. Many respondents initially included in the sample proved unavailable during the fieldwork period, and had to be replaced by other respondents with similar characteristics. In addition, during interviews, respondents occasionally suggested additional stakeholders they felt it was important to speak to in order to gain an understanding of particular outcomes, and this was accommodated where possible.

Table 3 provides an overview of the categories of respondent.

Table 3: Programme evaluation respondents

Category of respondent	Description	Purpose
BCURE programme staff and direct implementing/ consortium partners, including BCURE facilitators, mentors and trainers.	Individuals managing the programme, in country and in the UK, including consortium partners. Also, individuals hired by the BCURE partner to deliver training and mentoring, facilitate sessions, etc.	To understand their views on how and why their interventions influence change, in different ways for different people and in different contexts.
BCURE programme participants	Individuals directly participating in BCURE interventions (training, mentoring, workshops, knowledge cafés,	To understand their experiences of participating in the BCURE programme, and their

	policy dialogues, discussions around organisational systems development, etc.).	perspectives on how and why change has/has not happened.
High-level stakeholders, e.g. senior leaders in national government; national research community; others	Individuals not necessarily directly participating in BCURE interventions (although they may be) but who have a high-level insight into how the government system operates, the role of EIPM and the wider influence of BCURE and other influencing factors, e.g. director of a ministry, an MP. Also potentially including line managers of programme participants who can comment on emerging outcomes at organisational level.	To understand their views on how and why the BCURE interventions are influencing change, in different ways for different people and in different contexts. To understand their views on macro-contextual factors that influence EIPM, including political/administrative context for policy making; benefits/disadvantages of EIPM; enablers/barriers to EIPM; political and bureaucratic incentives/disincentives for EIPM; how evidence is used/not used in policy making currently and why.
Civil society/other external stakeholders	Individuals who can give an insight into the wider role of EIPM in the system, and the influence of BCURE and other influencing factors on EIPM, e.g. director of a research institute.	

The total number of stakeholders consulted for the Stage 2 programme evaluations is summarised in Table 4.[11] Full lists of respondents were included in the programme evaluation reports (which are confidential between DFID and the programme teams), but are not included in this Annex in order to ensure anonymity.

[11] Note this does not include interviews and conversations with DFID managers.

Table 4: Total number of stakeholders interviewed

Category of respondent	Total stakeholders consulted for Stage 2 programme evaluations
BCURE programme staff and direct implementing/consortium partners, including BCURE facilitators, mentors and trainers	46
BCURE programme participants	85
High-level stakeholders, e.g. senior leaders in national government; national research community; others	26
Civil society/other external stakeholders	31
Total	**188**

3.4.4 Data collection sources and methods

The programme evaluations involved the following data collection sources and methods.

- Exploratory workshops with BCURE implementing partner staff. These aimed to help the evaluator understand fully the nature of the interventions implemented, and observed changes with different groups. They also aimed to explore the team's perceptions on how and why the interventions contributed to change, and blockages to change. The workshop guide for evaluators is presented in the Supplementary Annex.
- Semi-structured interviews with key stakeholders (as detailed in Table 4). Generic topic guides were developed for each stakeholder category (see Supplementary Annex). These were customised by programme evaluation leads for each interview. At Stage 1, the evaluation attempted to use a Delphi Panel methodology to consult high-level and civil society stakeholders. However, this proved difficult to implement and was not thought to add significant value to the evaluation, and so was discontinued at Stage 2 in favour of semi-structured interviews with these stakeholders.
- Compilation of programme monitoring and implementation documents, e.g. government organisational procedures; or training materials, participant data, attendance data, targeting of invitations and selection criteria for individuals. A list of relevant documents was shared with the BCURE partners, who compiled them for review by the programme evaluation leads.

It was hoped the evaluation would also have access to relevant government documentation, such as policy documents, but this was not possible.

During data collection, evidence underpinning particular findings was triangulated in three ways:

- Internally, within interviews – claims about change were triangulated through asking for examples and further detail from the respondent.
- Between different interview respondents (different categories of respondent, different individuals within the same department, line managers and line managers).
- Between primary and secondary data sources.

3.4.5 Data analysis methods

Primary data from workshops and interviews was written up using a template structured according to the evaluation questions (see Supplementary Annex).

The programme evaluation leads then extracted evidence into a Microsoft Excel analysis spreadsheet, as follows:

- Transcripts were reviewed for insights on the *outcomes* mentioned by respondents. Each outcome mentioned was entered into a new row in the spreadsheet. Each outcome was coded according to whether it was observed (the interview respondent stated that it had already happened), anticipated (it had not happened yet but the respondent expected it to) or implied (no explicit mention of the outcome was made but the interview data enabled the evaluation team to infer, tentatively, that the respondent had observed or anticipated it). Where a respondent had also been interviewed in Stage 1,

programme leads reviewed the transcript from the previous year, to gain a sense of whether outcomes had been furthered or deepened.

- The transcript was then reviewed for any evidence of how and why the outcome came about: the *mechanisms* respondents felt had contributed to the outcome and the *contextual and intervention factors* respondents felt had enabled (or prevented) the mechanism 'firing'. This process was an interpretive rather than mechanical one, requiring skill and judgement on the part of the researcher to decide how best to categorise the data. This information was entered (in summary form, along with verbatim quotes) alongside the outcome data.
- Where a source provided evidence of only part of a CIMO (e.g. suggesting a particular mechanism was important without providing any insights into the contextual or intervention factors that sparked it), cells were simply left blank.
- Each row of data was also coded against our EQs, according to whether it related to outcomes at individual, interpersonal, organisational, institutional or policy level.

This created a catalogue of evidence from the interview data, in order to help the programme evaluation leads systematically and transparently assess the strength of evidence behind particular changes and identify how and why these changes were thought to have come about.

Secondary data: Documents were reviewed by the programme evaluation leads with the help of a research assistant. Programme leads compiled summary notes in Word. Evidence relating to *outcomes* was extracted into a second Microsoft Excel document review spreadsheet, as follows:

- Documents were reviewed for insights on the outcomes generated by the programme.
- This information was entered in summary form into the spreadsheet, coded according to which EQ the data related to.

Data from primary and secondary sources was then synthesised by the programme evaluation lead to draw conclusions on:

- Any evolution in the contextual challenges faced by the programme in the case study country context.
- Progress against programme milestones since 2015, including any adaptations to plans.
- Summary of evidence on the outcomes achieved against each of the EQs.
- Insights into BCURE's contribution to the outcomes, including preliminary analysis on how and why the outcomes were achieved. However, a full realist analysis was not conducted at programme level; instead, the data was fed into the overall synthesis.

3.4.6 Rubrics for assessing extent of change, strength of evidence and contribution

To aid the analysis and ensure consistency in judgement across the programme evaluations, the programme evaluation leads applied rubrics to assess the **extent of change**, the **strength of evidence underpinning the assessment of change** and a qualitative judgement on the programme's **contribution** to change in relation to each EQ. The ratings assigned by each lead evaluator were moderated by the team leader and/or programme director to ensure consistency across the evaluations. This involved reviewing the checking the ratings given against the evidence discussed in the text to identify any discrepancies between the ratings given, and the evidence presented. Discrepancies were discussed with the lead evaluator, who then either revised the rating or strengthened the discussion of the evidence.

Evaluative judgement: extent of change
Substantial change: evidence that change has scale, depth and sustainability.
Established change: evidence of change and/or improvement that is consolidated and widespread, and has potential to be sustainable.
Moderate change: evidence of change and/or improvement but not widespread.
Early change: some evidence of scattered change, but not consolidated.
No evidence of change or improvement.

Strength of evidence
An indication of the strength of evidence underpinning the assessment about the extent of change is made using two criteria**: 1) reliability of data sources and 2) extent of triangulation between data sources.**

These are then brought together into a rubric that enables us to rate the strength of evidence in a systematic way.

1) Reliability of data sources
We have grouped data sources for the programme evaluation into four categories. The sources have been **ranked** in increasing order of robustness, with 1 representing the least robust – anecdotal evidence from programme staff that may not be independent and exhibit various biases – and 4 the most independent and robust primary data from a range of programme stakeholders, which, although it may also exhibit biases, has been collected by the evaluation team using robust methods to mitigate biases.

Monitoring and evaluation (M&E) data is included in category 3 **only** if the evaluation team has judged the data to be robust. Where M&E data is less reliable, we have categorised it as 2.

Data sources (1 = least robust, 4 = most robust)

1. **Verbal reports from programme staff** (through interviews conducted as part of the evaluation).
2. **Strategy and implementation documents** (e.g. proposal, inception report, training materials, online resources, workshop reports, quarterly reports). This category also includes less reliable M&E data.
3. **M&E data** collected by the programme (reviewed by the evaluation team and judged to be reliable, e.g. pre/post-testing data; needs assessments; baseline reviews).
4. **Primary evaluation data** (largely consisting of interview data with programme stakeholders and evaluators' observation, not including programme staff interviews).

2) Extent of triangulation
During data collection and analysis, we have used **triangulation** as a technique to assess the strength of evidence underpinning a finding. Triangulation is applied internally to the primary data set (4) and externally between the four respective data sources. The extent of triangulation between data sources allows us to make a judgement about the strength of the evidence underpinning evaluation findings.

Other factors affecting the strength of evidence, for example **prevalence**, are discussed in the narrative within Section 4 of the main report in order to further nuance findings.

There are different configurations of triangulation between sources of varying reliability in our dataset. For example, evidence for a finding may derive from one of the most reliable sources (4) and one of the least reliable (1), or from three of the least reliable sources (1, 2 and 3) or from primary data only (4). In some cases there may be triangulation between all four sources, which gives us the strongest evidence.

Strength of evidence rating
In order to communicate the strength of evidence in a transparent and systematic way, we have brought the two criteria of **robustness** and **extent of triangulation** together into a rubric to enable rating of the strength of evidence underpinning our evaluative judgements. Our aim is to give the reader an indication of our judgement of the strength of evidence behind our findings. This approach provides a transparent basis for the evaluation findings and judgements and enables systematic rating across the different programme evaluations. However, the approach is limited, as it is only an approximation and there are combinations of sources that do not fit neatly within the rubric.

'Strong evidence' consists of evidence from strategy and implementation documents, confirmed by reliable M&E data, verified by primary evaluation data. 'Reliable' and 'partial' evidence consists of various combinations of less reliable or less well-triangulated data sources.

Strength of evidence rating

Strong evidence: evidence exists from data sources 2 + 3 + 4 – strategy and implementation documents, confirmed by reliable M&E data, verified by primary evaluation data.

Reliable evidence: evidence exists from data sources 1 + 2 + 3 **OR** from 4 + 1 + 2 **OR** from 4 + 3 (e.g. verbal team reports, supported by strategy and implementation documents, confirmed by reliable M&E data or primary evaluation data or primary evaluation data confirmed by reliable M&E data).

Partial evidence: evidence exists from data sources 1 + 2 only **OR** 1 + 3 **OR** 2 + 3 **OR** from 4 only (e.g. verbal team reports, supported by strategy and implementation documents).

Evidence from only one source: evidence exists from only one of data sources 1, 2 or 3 **OR** evidence exists from only one stakeholder in 4.

Insufficient evidence: there is insufficient evidence to make a judgement.

Contribution

Finally, a judgement was made regarding the significance of the programme's contribution to change. This represents a qualitative judgement on the part of the lead evaluator, based on a consideration of evidence collected relating to other factors that may have contributed to change.

Contribution rating

+++: Evidence that programme made a crucial contribution to observed change/observed change is directly attributable to the programme.
++: Evidence that programme made an important contribution.
+: Evidence that programme made some contribution.
-: Evidence of no contribution, or no improvement evident.
0: Insufficient evidence to make an assessment.

3.5 Literature review

A realist literature review (Punton et al., 2016a) was conducted during the early stages of the evaluation, in 2014–2015.[12] The aim of the review was to provide a practical summary of recent evidence on what works to promote EIPM, in order to both contribute to the wider evidence base and refine our CToC and begin developing CIMO configurations. The findings informed the CToC and the development of the first iteration of CIMOs tested in Stage 1. Insights from the literature review have been drawn out in Section 4 of the main report. However, the literature proved less useful at Stage 2 than at Stage 1, because our theories have evolved beyond the boundaries of the evidence considered during the early stages of the evaluation. The literature review will be updated in 2017, prior to the Stage 3 evaluation, in order to further test and refine the programme theory and incorporate evidence that was missed or not yet available in 2015.

[12] Available from http://www.itad.com/knowledge-products/bcure-literature-review/

3.6 Impact case study

The impact case study aims to generate evidence on how capacity building for EIPM can lead to improvements in the quality of policy processes, the hoped-for ultimate impact of the BCURE programmes. It is designed to complement the BCURE programme evaluations.

This component was developed because the BCURE evaluation Terms of Reference requires the evaluation to gather evidence on how and why capacity building for EIPM can influence system-wide shifts in government institutions – including changes in how policy is made and enhancing the policy process (see Annex 1). It was recognised that it may be difficult to demonstrate these shifts as a result of specific BCURE projects, within the three-year life of the project and within the resources available for the evaluation. The impact case provides an opportunity to examine a non-BCURE capacity building intervention that has been operating for a longer period of time, to look for evidence on how and why they influence evidence use and the quality of policy processes.

The impact case study was the focus of an evaluability assessment and scoping process during the inception phase, detailed in the Inception Report. South Africa was selected as the country that most closely met the criteria. The study focuses on the Department for Planning, Monitoring and Evaluation (DPME), exploring the National Evaluation System (NES) as an example of a capacity support initiative that intervenes at organisational level to enhance evidence use in policy making and has been established for some time (since 2011), providing an opportunity to investigate how capacity building can promote change in the longer term.

The core research question for the impact case is: **How has DPME's support to the NES influenced evidence use and contributed to changes in the quality of policy processes?**

To answer this, the case study looks specifically at two experiences with line ministries. The first is the updating of the government of South Africa's early childhood development (ECD) policy following a DPME-facilitated Diagnostic Review in which the Department of Basic Education (DBE) had a leading role. The second experience is the evaluation of the Department of Trade and Industry's Business Process Services (BPS) programme and changes in the programme design arising from the evaluation.

There are three main analytical strands to the impact case study: developing and testing CIMOs at the level of organisational change; researching the policy development process in order to provide insights into the concept of 'policy quality'; and exploring the interrelationships and dynamics between CIMOs and how they influence policy processes.

The case study seeks to explain how and why evidence produced through the evaluation/review of these policies/programmes has been used in decision making. The case study also examines connections between evidence use and enhancement of policy processes in the two departments concerned.

The Stage 2 exercise involved review of relevant documentation as well as interviews in South Africa with DPME staff members, intervention participants, high-level stakeholders, civil society or other external stakeholders and service providers. This built on data collection from Stage 1: in total 39 interviews were conducted in Stage 1 and 2, involving 32 unique interviewees. Data was analysed in the same way as programme evaluation interview data, as detailed in Section 3.4.5 above.

Table 5: Number of stakeholders consulted in impact case study

Category of respondent	Total stakeholders consulted for impact case study
DPME staff	8
Intervention participants	11
High-level stakeholders, e.g. senior leaders in national government; national research community; others	8
Civil society/other external stakeholders	5
Total	**32**

3.7 Overall synthesis

The synthesis brings together the findings from the programme evaluations, literature review and impact case study in order to draw generalisable conclusions about how and why different BCURE interventions have contributed to different patterns of outcomes in different contexts. The purpose of the synthesis at Stage 2 is to produce an evidence-based set of refined CIMOs and a refined programme theory. The synthesis does not focus on *performance judgements* of the individual BCURE programmes; these are contained within the BCURE programme evaluations and are confidential between DFID and the programme implementation teams.

Overview of sources for the synthesis

A CIMO database was developed to combine the coded CIMO data from the 220 semi-structured interviews conducted for the programme evaluations and the impact case.[13] This data was 'cleaned', and then a series of systematic analytical steps were followed to identify patterns in the data, for use in testing and refining our Stage 1 CIMOs. The analysis also draws on the full Stage 2 programme evaluation reports, and the BCURE Literature Review.

Synthesis approach

The synthesis provides a realist analysis on how and why different BCURE interventions have contributed to different patterns of outcomes in different contexts. The purpose of the synthesis at Stage 2 is to produce an evidence-based set of refined CIMOs and a refined programme theory. We found two of the steps of meta-ethnography helpful to provide a clear and transparent structure for the synthesis of findings from the six programme evaluations and the impact case study (see Box 4).

Box 4: Meta-ethnography

Meta-ethnography is an interpretive synthesis method, involving the transfer and translation of ideas, concepts and meanings across different sources (Noblit & Hare, 1988). While we did not apply the method in full, we found two of its steps helpful to structure the synthesis of findings from the six programme evaluations and impact case: determining how the evidence was related and 'translating' the sources into one another.

- Determining how the evidence is related: identifying points of comparison (reciprocal translations) or opposition within the data, and identifying 'lines of argument' – developing inferences about the 'whole' through 'comparing and sorting interpretations, examining similarities and differences, and then integrating or framing these within a new interpretation that can be applied across all the studies' (Pope et al., 2007).
- 'Translation': attempting to 'translate' the new interpretations into each of individual evaluation datasets, checking to see how far new concepts accurately reflect each of the evaluation findings and scrutinising conceptual differences.

The steps in the synthesis were as follows:

- A two-day participatory evaluation team workshop was held, where the data was examined and the two steps in meta-ethnography were initially applied.
  - The team read through the coded CIMO data to identify common concepts, themes or metaphors that applied across the sources. These were identified by asking, 'Is this an example of something we have seen elsewhere? Is there a common concept we can use to explain these things?'
  - The team also attempted to isolate more abstract explanations or models that explained groupings of findings across the cases (lines of argument). Lines of argument were identified by asking, 'Can this concept, theme or metaphor be explained using a more abstract concept, theme or metaphor that encompasses and goes beyond the more specific explanation?' This analysis was used to start constructing tentative new and revised CIMO configurations.

[13] Note that not all interviews were included in the CIMO database, as some did not provide insights into how and why outcomes came about (or failed to).

  - An initial translation step was applied, in which team members were asked 'Does this apply in your BCURE context? Are there any nuances from interviews with respondents in your setting?' This enabled scrutiny of differences within the data, which were used to adjust, refine and caveat our CIMO configurations

- Following this exercise, two team members reviewed the synthesis database in a systematic way. This involved the following process:
  - 'Cleaning' the data. This represented an early analysis step, involving ensuring CIMO data was consistently entered in the appropriate columns, and in particular paying attention to the 'M' column to ensure this contained information about the 'causal spark.' This frequently involved returning to interview transcripts in order to check interpretations and add further detail.
  - Applying colour coding to outcome data, according to the extent or type of change observed.
  - Applying 'codes' to each row to group together similar I, C, M and O factors, as well as to tag outcomes associated with particular resources provided by BCURE. Codes were developed iteratively: the data was examined, initial promising codes were identified and codes were applied and then refined as further data suggested the need to nuance, combine or split them.
  - Applying filters to the database to isolate patterns in the coding, and provide an insight into the strength of evidence behind the patterns. Filters allowed us to identify how many respondents associated particular outcomes with particular resources provided by BCURE, and particular mechanism, context and intervention factors.
  - Writing up this analysis into the report narrative and using this to develop new and refined CIMO configurations.

  This was a highly analytical and creative process. It was undertaken by two members of the core team, which enabled cross-checking of coding and analytical decisions, and constant communication via Skype and email to help clarify, refine and challenge the analysis.

- Following the analysis of the synthesis database, the analytical narrative and refined CIMO configurations were translated across the original sources, by re-examining the full programme evaluation reports to consider how well they reflected and encompassed the ideas as originally expressed by respondents and analysed by the programme evaluation leads. The findings were then nuanced and caveated accordingly.

At the end of the synthesis process, we had a revised set of CIMO configurations representing our 'best guesses' at the end of Stage 2 about how BCURE interventions are leading to change. These provided new insights into how elements of our programme theory lead to and reinforce other elements, and were used to refine our programme by nuancing expected outcomes and adjusting the anticipated links between them. The CIMOs and programme theory will be revisited, tested and refined for a final time at Stage 3 of the evaluation.

Strength of evidence behind CIMOs: Throughout the synthesis narrative, we discuss the prevalence of perspectives underpinning the CIMOs. 'Prevalence' refers to the number of interviews in which respondents expressed a particular theory (or part of a theory) about how and why change happened or is expected to happen, which was developed at synthesis stage into a coherent CIMO.

This approach does not provide an exact 'count' of the number of times particular CIMOs 'happened.' Our programme theory is broad, and it was not possible to explore all elements of it with all interviewees – discussed in Section 3.11 below. Even if a particular element of the theory was discussed with a respondent, they may have observed something (an outcome, mechanism, feature of context, etc.) but did not mention it for any number of reasons – for example they did not think of it, did not understand the question, thought something else was more interesting or did not feel comfortable discussing it. The prevalence data therefore simply reports how often outcome *x* was mentioned in the same interview as context, intervention factor or mechanism *y*, which is a broad indicator of the frequency of co-occurrence and provides a way of being systematic about the patterns in the data.

The process of developing CIMOs is a creative and interpretative one. In some cases, the analysis suggested certain factors might be important to enable particular mechanisms even though they were not explicitly correlated with particular outcomes by respondents – for example certain features of the national context. These have been flagged in the text and in our revised CIMO configurations as areas where there is limited evidence to date, but which may be important to explore at Stage 3.

3.8 Ethics

The key ethical issue faced in the evaluation is protecting and managing the confidentiality of government documentation and stakeholder views at the local level. A number of the BCURE partners are operating at a high level within government and as such have access to government policy processes as they unfold. Access to these processes and the actors involved has been navigated with the close collaboration of the BCURE partners, in order to avoid the evaluation negatively impacting the relationships that BCURE providers have worked hard to develop.

We ensured appropriate permissions were obtained from individuals before commencing data activities, with consent obtained at the beginning of interviews to record the discussion and to use the insights gained in our reports (see Supplementary Annex). Unique anonymous interview codes have been attached to each transcript and referenced in the text. Where the content of quotes may identify an individual, this information has been removed.

The BCURE programme evaluations are confidential reports viewed only by DFID and by the programme teams. The synthesis report aims to reflect on findings at a higher level of abstraction, allowing us to avoid detailed reporting on sensitive issues relating to particular government processes.

3.9 Evaluation team

The evaluation was undertaken by a team from Itad, in partnership with Stellenbosch University. The evaluation team are independent from DFID, and there are no conflicts of interest within the team in relation to any of the BCURE programmes. The team included lead evaluators for each programme evaluation, working in partnership with national consultants to conduct the country case studies. The full core team and their roles are detailed below:

- Isabel Vogel (Itad associate): team leader, lead on SECURE Health (Kenya) programme evaluation and the synthesis
- Rob Lloyd (Itad staff member): project director and quality assurance.
- Melanie Punton (Itad staff member): lead on VakaYiko (Zimbabwe) programme evaluation, support to UJ-BCURE programme evaluation, support to synthesis, lead on Literature Review
- Joe Bolger (independent consultant): lead on ECORYS (Bangladesh) programme evaluation; lead on impact case study
- Babette Rabie (Stellenbosch): lead on UJ-BCURE (South Africa) programme evaluation; support to impact case study
- Gregory Gleed (Itad staff member): lead on Harvard (Pakistan) programme evaluation
- Teresa Hanley (independent consultant): lead on ACD (Sierra Leone) programme evaluation

The programme evaluations were supported by national consultants Alfred Odour (Kenya), Munhamo Chisvo (Zimbabwe), Andrew Lavalei (Sierra Leone), Maheen Sultan (Bangladesh) and Rafiq Jaffer (Pakistan). Research assistance support was provided by Louise Horton and Verdiana Biagioni.

3.10 Intended users of the evaluation

The evidence base on capacity development for EIPM is small, largely derived from the health field, and weighted towards studies examining the impact of training on individual capacity. There are significant evidence gaps around the role of interpersonal and organisational interventions in promoting change, and regarding the influence of EIPM capacity development on policy change and improved quality of policy

development processes. There is a particular lack of evidence on capacity development for EIPM in developing countries. Operational insights into how to design and implement this type of intervention in developing country contexts are also lacking.

To strengthen this evidence base, the BCURE evaluation provides robust evidence on how and why different approaches to capacity building for EIPM work, for whom and in which contexts, in developing countries. These lessons are intended to be directly applicable to the commissioning, design, implementation and adaptation of EIPM capacity building programmes in developing countries to improve results.

Therefore, the intended users of the synthesis report are, in the first instance, BCURE's managing team at DFID's Research and Evidence Division and the BCURE partners responsible for delivering BCURE programmes, to inform improvements within the current portfolio of programmes.

The findings are also intended to be of use to a wider audience of donors, funders, commissioners and implementers who are considering future EIPM capacity development programmes. These evaluation users may be in numerous fields, such as governance, public management and administration, and research and evidence utilisation. For these audiences, the evaluation findings provide evidence on:

- How and why different interventions lead to change, and contextual factors that affect outcomes.
- How interventions can be combined in multi-level capacity development strategies.
- How and why capacity development interventions can contribute to organisational and institutional shifts to embed EIPM behaviours and systems, ultimately enhancing policy development processes.

3.11 Limitations to the synthesis

There are some key challenges and limitations to the synthesis, in terms of timing, the dataset and methodological challenges.

- **Partial dataset:** primary data comes only from the selected case study countries, not from all programme sites. It is, therefore, limited in what it can say about how the BCURE programmes work in all their settings.

- **Ensuring consistency of data collection and analysis across a diverse team:** six different programme evaluation leads collected data, with the support of six national consultants. In addition, several new team members joined at Stage 2. There was limited time and budget to train the team comprehensively on the principles of conducting realist interviews, or on coding CIMO data. We attempted to mitigate this through a two-day team workshop prior to data collection, involving a full introduction to the programme theory and basic training on realist interviewing and analysis. Programme leads then provided training in-country to national evaluators prior to data collection. In addition, the CIMO dataset was cleaned at synthesis stage, and additional data incorporated that may have been missed during the initial coding process. Further training will be provided at Stage 3 to continue building the capacity of the team.

- **Granularity of data:** it has been challenging to reach an appropriate level of abstraction when analysing CIMO data. It is easy to over-partition these configurations down to very micro sets of factors. During the analysis we have attempted to reach a useful level of generalisability in the data analysis that can facilitate the application of the findings in planning and implementation.

- **Time demand for synthesis:** a key challenge arises from the time and resource investment required for achieving a good-quality qualitative synthesis of the enablers/barriers and CIMO data. This affects all stages, from requiring more time for interviews and data processing as well as reporting. We have mitigated this by undertaking as rigorous a process as resources allow for Stage 2 and being pragmatic.

As well as the general limitations above, the Stage 2 evaluation process had some specific data limitations which have influenced what has been possible in the synthesis.

- **Limited access to monitoring and other documentary sources in order to triangulate interview data:** many of the outcomes relating to changes in behaviour, relationships and organisational norms are

intangible and emergent, and the BCURE programmes have not systematically monitored them. Therefore, we have necessarily relied more strongly on interviews (see earlier point relating to strength of evidence). With regard to changes in policy decisions, it has proved difficult to obtain documentary evidence from government partners, for reasons related to confidentiality and access limitations. In addition, in most cases BCURE is not aiming to influence specific policies and so it is not possible to know in advance which documents might be useful to support claims made in interviews about organisational or policy change. We have mitigated this through the triangulation approaches described above and in Annex 3.4.

- **Prioritising outcomes and theories to assess within the limited time available for interviews:** the evaluation examined a wide range of outcomes at individual, interpersonal, organisational, institutional and policy level; and a wide range of theories about how and why BCURE was thought to contribute to these outcomes. It was necessary to prioritise which outcomes and elements of the programme theory to test with different stakeholders. This was not always easy, particularly when respondents were involved in a range of different interventions, theorised to work in different ways. We attempted to mitigate this limitation by designing unique interview guides for each respondent that aimed to test the most relevant theories for each respondent, and using later interviews to plug gaps in earlier ones. However, it proved difficult to ensure such a wide range of theories were systematically examined and insights fully triangulated. We plan to conduct a prioritisation exercise with DFID in advance of Stage 3 to select the most important CIMOs, in order to address this limitation.

- **Positive (confirmation) bias of respondents:** there is a very real possibility of confirmation bias in the primary data arising from the power dynamics of interviewing in developing country government settings. Evaluators can be seen as representing the international funder, and positive messages about programme outcomes may be given in an attempt to continue funding for the programme. We have mitigated this in three ways: in the interview process, by approaching the same topic from different angles with various interviewees and by asking for concrete examples to corroborate any claims of change; in the sample, by interviewing a range of participants including stakeholders external to the project, and cross-checking claims of change made by civil servants with their managers and peers; and in the analysis, by triangulating between data sources (i.e. different interview respondents, and where possible, secondary data) within the same case. However, the challenge of accessing documentary sources of evidence, and the challenge of investigating a wide range of theories and outcomes across a relatively small number of interviews, have both limited how far it has been possible to mitigate this limitation.

The data limitations described above have affected what was possible in the Stage 2 synthesis. Primarily, while the Stage 2 analysis has built on Stage 1 to *identify and further develop* a wide range of theories about how BCURE appears to be contributing to change at different levels, it has been unable to confidently *verify* these theories and conclude that change happened in the ways theorised rather than in some other way. The Stage 3 evaluation process will be designed to robustly test a narrower range of outcomes, and enable theories to be systematically tested against alternative explanations of change. This is discussed in Section 5.5.

4. Programme theory and CIMO refinement: from Stage 1 to Stage 2

Section 3.2 describes the approach used to develop and refine programme theory within the evaluation. This section details our Stage 1 programme theory and CIMOs, documents the changes made and the rationale behind these changes at Stage 2 and presents our refined programme theory at the end of Stage 2 of the evaluation.

Programme theory at the end of Stage 1[14]

When the programme 'entry point' is through interventions at individual level…

- Providing civil servants and senior government decision makers with information about the importance of evidence in decision making, alongside information about and/or opportunities to practice accessing, appraising and applying evidence in policy making processes, can *crystallise* existing knowledge or awareness of the concept of EIPM, leading to increased enthusiasm for it (CIMO 1). When participants see that new knowledge and skills are immediately applicable to their work, these resources can spark *eye-openers*, leading to behaviour change in the way they use evidence in their day to day work (CIMO 2). When participants are actively involved in a policy process, these resources can spark *game* changers, in which behaviour change influences the way evidence is used within these policy processes (CIMO 3). Following up training interventions with coaching can help embed new skills and enable knowledge to translate into behaviour change (CIMO 4).
- Providing coaching in the form of one-to-one mentoring can lead to *peer learning* as mentors and mentees learn together through applying different skills, technical knowledge and experience 'on the job' (CIMO 5) – resulting in mentees using evidence more or more effectively in their work.

When individuals began using evidence more in their day-to-day work, this can catalyse organisational change through…

- Enabling people who lack overt decision making power but who have opportunities to model EIPM behaviours in their job (when they are committed to or passionate about EIPM and have good interpersonal skills) to act as *junior champions*, demonstrating the value EIPM can bring to build organisational buy in 'from below' (CIMO 11).
- Developing a *'critical mass'* of people whose behaviour change can diffuse throughout the organisation (when a sufficient number have been reached, at different levels of seniority), increasing organisational commitment and buy-in to the concept of EIPM (CIMO 6).

When the 'entry point' is through interventions at interpersonal level…

- Providing networking opportunities for government and non-government actors (researchers, civil society, the media, the general public) to engage in dialogue about issues relating to EIPM promotes awareness of the importance of using evidence to inform decisions (CIMO 7), and enables participants to *learn from each other* about different policy issues, in an evidence-informed way (CIMO 8). Bringing people together also provides participants with access to researchers, experts and government actors, enabling *new relationships* to develop, potentially translating into new collaborations that facilitate EIPM, and/or provide civil servants with better access to good quality evidence.

When the 'entry point' is through interventions at organisational level…

- Providing technical support to assist government ministries, parliaments and cabinets to use evidence within specific policy processes builds organisational capacity to use evidence through '*learning by doing'*, resulting in new co-produced policy products or processes that are informed by evidence (CIMO 9). Supporting senior stakeholders to promote EIPM within their organisations (who have seniority, commitment to the issue and good interpersonal skills) also enables them to act as *transformational*

[14] Note that the CIMO numbers mentioned here relate to the *Stage 1 CIMOs.*

leaders, who can push change 'from above' to support EIPM and initiate reforms, resulting in high-level buy-in for EIPM and potentially new organisational tools and systems to promote it (CIMO 10).

- Where there are structural capacity gaps, providing technical support to help establish structures for policy making with evidence use at their heart can create a *focal point* for EIPM (CIMO 12). New tools and systems to promote evidence use can also *facilitate* staff members to use evidence within their jobs better or more easily (CIMO 13) and/or provide positive or negative incentives to individuals, which *reinforce* the use the evidence within policy processes (CIMO 15).
- New evidence-informed policies and products, and success stories of evidence use having 'good results', can have a *demonstration effect*, showcasing the positive results that evidence can bring to policy processes. This can lead to increased organisational commitment and buy-in to (and potentially increased organisational kudos and resources for tools or systems that promote) EIPM (CIMO 14).

Stage 1 CIMOs explaining individual-level change

	Intervention	Context	Mechanism	Outcome
1: the 'crystalliser' *Based on seven interviews from four countries: India, Zimbabwe, South Africa, Kenya*	Where interventions are less directly relevant to participants but still offers practical knowledge of and insights into EIPM…	… and/or (perhaps) where participants start with lower capacity / awareness…	…this **crystallises** awareness of EIPM, and /or allows application of EIPM labels to current practices **…**	…leading to increased awareness of / enthusiasm for EIPM (but not behaviour change yet)
2: the 'eye opener' *Based on 14 interviews from two countries: Zimbabwe and Kenya; plus ACD regional conference*	Where training interventions are practical, interactive, needs-focussed, offer practical skills, and target people who can directly apply learning…	… and where there are external pressures or motivations to apply training … and participants already have internal motivation for EIPM…	…this sparks an **eye-opener**, in which participants see that training is immediately applicable to their own work, and put it into practice**…**	…leading to immediate behavior change in which individuals apply EIPM principles in their own work.
3: the 'game changer' *Based on four interviews from two countries (Kenya, Zimbabwe)*	Where training interventions are directly linked to a policy process or relevant to processes participants are directly involved in, and courses offer practical learning about EIPM…	… and where there is direction / permission / support from senior management… and participants already have internal motivation for EIPM…	…this sparks a **game-changer**, in which participants see that training is immediately applicable and use new knowledge to inform the process they are involved in…	…leading to immediate behavior change around EIPM feeding into instrumental policy and process change.
4: embedding capacity *Based on four interviews from three countries: Kenya, Sierra Leone, South Africa; plus ACD report*	Following up training interventions with mentoring or similar (e.g. mentoring, training-of-trainer approaches)…	…and where there has been previous participation in training…	…helps to embed new skills and enable new capacities to translate into behavior change…	…resulting in participants applying new skills in practice (behaviour change).

5: peer learning on the job *Based on seven interviews from two countries: South Africa and South Sudan*	Mentorship structured around 'on the job' needs, mentors having appropriate skills to meet these needs plus interpersonal skills, match in seniority, appropriate modality and length of mentorship…	…and where there is organisational support for mentorship, practicing of EIPM skills, mentors and mentees have time and commitment to engage, and there is a 'click' between mentors and mentees…	…mentoring sparks **peer learning** as mentors and mentees learn together through applying different skills, technical knowledge and experience 'on the job'…	…simultaneously increasing EIPM capacity and developing new EIPM practices and behaviours.
6: 'sleeping beauties' *Based on three interviews from three countries: India, Kenya, South Africa*	Where training has reached a 'sufficient' number of the 'right' people (perhaps mid-level staff as well as senior managers) in an organisation, including through combining training and other interventions…	…and (perhaps) where there is limited history and culture of EIPM in the organisation…	…this creates a **critical mass** of people (sleeping beauties) whose new knowledge and behaviour change can diffuse…	…leading to (possibly slow) dissemination of buy-in and commitment to EIPM within and beyond the organisation, and (perhaps) creating the conditions for champions to emerge (ICMOs 10-11)

How were these Stage 1 CIMOs refined at Stage 2?

- The 'eye opener' and the 'game changer' theories were merged into Stage 2 **CIMO 1**. Both theories still appear valid based on additional data collected at Stage 2, but more work has been done to unpack the mechanism at work and link it with learning theory – drawing on the Kirkpatrick framework (Kirkpatrick Partners, n.d.) and the concept of 'self-efficacy' (Bandura, 1977). Analysis suggested that a common mechanism linked the 'eye opener' and 'game changer' theories and that they could be framed as one. Stage 2 data provided further evidence of the importance of the intervention and contextual factors identified at Stage 1.
- The 'game changer' theory also informed **CIMO 6** (because it is about how individual change can filter up to influence organisational change, through feeding better quality evidence into decision making processes).
- There was little evidence on the 'crystalliser' theory at Stage 2. While there was some evidence of training interventions that had increased knowledge but not led to behaviour change, this seemed to be more a result of a *blocking context* relating to the training having limited relevance to participants' day-to-day work (e.g. in Pakistan), rather than a separate mechanism. The theory has not been *disproved* and may still have explanatory value, but is not considered at Stage 2 owing to lack of evidence.
- 'Peer learning on the job' and 'embedding capacity' were merged into **CIMO 2.** The Kirkpatrick framework and Pawson's categories of mentoring resources (Pawson, 2004) were used to further unpack the mechanism and frame it at a higher level of abstraction that encompassed both Stage 1 theories. Stage 2 data provided further evidence of the importance of the intervention and contextual factors identified at Stage 1.
- The 'sleeping beauties' theory was substantially revised and developed into **CIMO 6**, relating to how individual-level practice change can 'filter up' to organisational change. Stage 2 data suggested the mechanism was not in fact 'critical mass' – this proved quite a blurry concept that different stakeholders defined in different ways. Rather, training a 'sufficient number' of people seemed a necessary contextual factor for individual change to 'filter up' and lead to organisational-level outcomes.

Stage 1 CIMOs explaining interpersonal-level change

	Intervention	Context	Mechanism	Outcome
7: awareness through networking *Based on six interviews in two countries: Zimbabwe and South Africa*	Where participants from different sectors engage in open dialogue in an informal setting, and small groups are facilitated to exchange perspectives using evidence…	… and where there is a perceived need to build relationships between policy and research (and other stakeholders) to tackle and issue…	…this **raises awareness of the importance of EIPM** through open dialogue between stakeholders…	…resulting in improved awareness of EIPM and improved relationships with relevant stakeholders, including between policy and research.
8: collaborative learning *Based on seven interviews in four countries: Kenya, Zimbabwe, South Africa impact case, Sierra Leone*	Where a practical, informal, participatory and collaborative format is used, involving people with diverse, relevant expertise, and senior figures are independently facilitated to have a structured dialogue using evidence…		…this enables **collaborative learning from others** using evidence…	…resulting in increased EIPM capacities, translating into EIPM commitments and behaviour change.
9: 'learning by doing' through co-production *Based on seven interviews in two countries: South Africa and South Africa impact case, India*	Where BCURE staff provide direct support within a specific policy process, ensure that targeted staff play a key role and feel ownership of the process, and can act as independent mediators/ facilitatiors…		…this enables **learning by doing through co-production**, building EIPM capacity through active engagement of government and BCURE actors in an EIPM policy process…	…leading to improved capacity for EIPM, increased support for EIPM, new (evidence-inormed) policy products and processes, and/or 'demonstration' effects that further catalyse EIPM (ICMO 14)

How were these Stage 1 CIMOs refined at Stage 2?

- The 'awareness through networking' theory was refined into **CIMO 4**. Pawson's typology of mentoring resources resonated with the data relating to networks, and was used to articulate the underlying mechanisms more clearly. Stage 2 data provided further evidence of the importance of the intervention and contextual factors identified at Stage 1.
- The 'collaborative learning' theory was refined into **CIMO 3**. The underlying mechanism was unpacked using insights from the Kirkpatrick model and also Pawson's typology of mentoring resources. Stage 2 data provided further evidence of the importance of the intervention factors identified at Stage 1.
- 'Learning by doing through co-production' was refined and reframed as 'accompaniment' (**CIMO 9**), drawing on literature from the health policy and governance fields (Faustino & Booth, 2014; IDRC, 2014). Stage 2 data provided further evidence of the importance of the intervention factors identified at Stage 1, but generated additional detail and allowed greater nuancing of findings.

Stage 1 CIMOs on champions

	Intervention	Context	Mechanism	Outcome
10: transformational leaders *Based on 14 interviews in five countries: India, Sierra Leone, Zimbabwe, South Africa impact case, Kenya; plus Transform Nutrition case study*	Senior stakeholders being identified and supported informally to promote EIPM…	…where individuals have seniority within the system; commitment and passion; and interpersonal skills, good political relationships, credibility and respect…	…enables individuals to act as **transformational leaders**, exercising high level influence on other senior government figures to push change from above to support EIPM, and initiate reforms…	…resulting in high level buy-in and support for EIPM and/or new organisational tools and systems for EIPM
11: junior champions *Based on six interviews in three countries: South Africa, Kenya, India*	Interventions developing capacity for EIPM among individuals in an organisation…	…where indivduals lack overt decision making power but hold positions that provide opportunities for modelling EIPM behaviours; have good interpersonal skills; and are committed to EIPM…	… enables individuals to act as **junior champions,** pushing change from below through modelling and diffusing EIPM practices…	…building organisational buy-in through demonstrating the value of EIPM and (potentially) becoming or creating future transformational leaders.

How were these Stage 1 CIMOs refined at Stage 2?

- The Stage 2 data provided further evidence on the role of champions in catalysing change and promoting EIPM. However, few of the BCURE programmes are specifically working with champions through targeted interventions (rather, BCURE partners acknowledge that their interventions can potentially create or inspire existing champions, and recognise the importance of champions in providing an entry point for them to work within government). Because activities involving champions are not a core part of the programme, it has therefore been difficult to work out how they fit in to our programme theory.
- Rather than retain stand-alone theories on champions, it was therefore decided to incorporate the data into other CIMOs, in order to develop theories on the role that particular committed or charismatic individuals might play within BCURE programmes. For example, champions are part of the context that helps enable **CIMO 7** and are a potential outcome of **CIMO 2** and **CIMO 9.** These theories on champions have not been disproved and are therefore not being discarded, but rather are being 'parked' at Stage 2, and the insights incorporated into other theories.

Stage 1 CIMOs explaining organisational-level change

	Intervention	Context	Mechanism	Outcome
12: focal point *Based on one interview in Kenya; plus secondary reports from the ACD programme*	Where manuals, tools and processes for EIPM are used as an 'entry point', designed around an important, high profile policy process or issue; and the process is participatory…	…and where there is a structural capacity gap (often broader than EIPM), and a demand for greater structure in a context of instability or fragility…	…new systems, tools and processes create a **focal point** for EIPM in establishing broader structures for policy making with EIPM at their heart…	…resulting in improved policy products or processes, generating buy in and resources for EIPM, and/or resulting in learning-by-doing (ICMO 9) / demonstration effects (ICMO 14).
13: facilitation *Based on one interview in Sierra Leone, and secondary reports from the ACD programme*	Where EIPM tools, systems or policies are directly relevant to work that is already being done…	…and where processes are supported by high level champions (ICMO 7)…	…this leads to the mechanism of **facilitation**, in which tools or systems provide practical assistance, enabling people to do their jobs better or more easily…	…resulting in the EIPM system or tool being used, and (potentially) increasing the value of evidence through demonstrating the benefits it can bring (ICMO 15)
14: demonstration effect *Based on seven interviews in four countries: India, Zimbabwe, Kenya, South Africa impact case*	Interventions resulting in concrete products or actions that demonstrate EIPM…	…where structures are in place to allow diffusion and replication…	… success stories of EIPM having 'good results' can have a **demonstration effect**, influencing others through demonstrating the positive potential of EIPM…	…leading to increased buy-in / demand for EIPM, and potentially increased organisational kudos and resources for EIPM
15: reinforcement *Based on fifteen interviews in three countries: South Africa and South Africa impact case, Kenya, Sierra Leone*	Organisational tools or systems involving positive or negative incentives to adopt EIPM behaviours…	…where the tool, system or agency is strategically positioned or has legislative backing, lending it authority…	… this enables the mechanism of **reinforcement**, in which external positive or negative incentives influence behaviour…	…resulting in EIPM systems or tools being used, and manifested in behaviour and/or policy change.

How were these Stage 1 CIMOs refined at Stage 2?

- There was little evidence on the 'focal point' theory at Stage 2, which was very tentative last year. This theory has therefore been dropped. It should be noted that the theory has not been *disproved* and may still have explanatory value, but is not considered at Stage 2 owing to lack of evidence.
- The 'demonstration effect' was refined and became **CIMO 8.** The mechanism was reframed as a 'showcase' in order to increase the clarity of the concept, but is essentially the same. Stage 2 data provided more insight into the intervention and contextual factors necessary to spark the mechanism.

- The theories on 'facilitation' and 'reinforcement' have been retained at Stage 2 (as **CIMOs 11** and **12**), but there is still relatively little evidence on them.

Table 6. Summary of refinements to Stage 1 CIMOs

Stage 1 CIMOs	Status at Stage 2	Corresponding Stage 2 CIMOs
1: the crystalliser	Dropped	Outcome explained through 'blocking' contextual factors in CIMO 1
2: the eye opener	Refined and merged with 3	CIMO 1
3: the game changer	Merged with 2, but also informed CIMO 6	CIMO 1 (elements in CIMO 6)
4: embedding capacity	Refined and merged with 5	CIMO 2
5: peer learning on the job	Refined and merged with 4	CIMO 2
6: sleeping beauties	Substantially refined	CIMO 6
7: awareness through networking	Refined	CIMO 4
8: collaborative learning	Refined	CIMO 3
9: learning by doing	Refined	CIMO 9
10: transformational leaders	Parked	Insights incorporated into other CIMOs (e.g. CIMO 2, 7, 9)
11: junior champions	Parked	Insights incorporated into other CIMOs (e.g. CIMO 2, 7, 9)
12: focal point	Dropped	CIMO 9
13: facilitation	Retained	CIMO 11
14: demonstration effect	Refined	CIMO 8
15: reinforcement	Retained	CIMO 12

New CIMOs that have emerged from the data at Stage 2, and articulated in Section 4 of the main report, are:

- CIMOs 5 and 7 (relating to how individual behavour change catalyses change at an organisational level).
- CIMO 10 (relating to organisational decisions to formally adopt new tools or comprehensive systems to embed EIPM).
- CIMOs 13 and 14 (relating to the role of national BCURE partners as institutional EIPM actors)

The refined CIMOs at Stage 2 were used to revise the overall programme theory. The programme theory narrative and summary diagram are detailed below.

Stage 2 Programme Theory

When the programme 'entry point' is through interventions at individual level…

- Providing information about EIPM (its importance, and how to access, appraise and apply evidence in decision making), alongside opportunities to practise skills, generate self-efficacy (a feeling of 'now I

know how') and lead to behaviour change when training is directly relevant, there is management support and training comes at the 'right time' for the organisation (CIMO 1).

- Coaching provides *encouragement*, which generates or embeds a feeling of self-efficacy ('now I know how'); *contacts and sponsorship* that give access to useful networks; and *advice and a guiding hand* that promote understanding and builds confidence. This can result in participants changing their behaviour in relation to EIPM where they have either personal motivation or organisational incentives to do so. Success depends on coaching being driven by clear objectives based on participants' needs, and the coach having the right interpersonal and professional qualities to provide for these needs (CIMO 2).
- Facilitated spaces for dialogue and collaboration can enable advice and sharing of perspectives to generate knowledge and influence attitudes about EIPM, including learning about what others have done when facing similar challenges. This is made possible where interventions bring together diverse groups of people with relevant interests, and provide space to share challenges, in a context of a positive wider discourse in support of EIPM. However, this learning may be put into only use if there are existing direct opportunities to do so, although spaces for dialogue potentially create a conducive context for other interventions to stimulate behaviour change at a later stage (CIMO 3).
- Providing individual-level support (such as training or coaching) in a sensitive, collaborative way can provide a 'foot in the door' for BCURE partners, generating permission and buy in for them to begin implementing organisational reforms – this could be a particularly important 'way in' in contexts where it is not possible to start working directly at organisational level, for example where access to government is difficult to secure (CIMO 5).

When individuals began using evidence more in their day-to-day work, this can catalyse organisational change as follows:

- When a sufficient number of individuals (including some with leadership roles) begin accessing, appraising and applying evidence more in their work, this can 'filter up' and lead to higher-level recognition of the value of an evidence-informed approach – through senior staff seeing and being impressed by good-quality evidence products and through these products feeding into senior decision making processes and improving them (CIMO 6).
- When individual support influences individuals in mid-level roles, who are committed and passionate and who have supportive senior management, they can formally cascade their learning through introducing new ways or working and new structures and processes within their organisations (CIMO 7).

When the 'entry point' is through interventions at interpersonal level…

- Facilitated spaces for dialogue (e.g. between policy makers, researchers, civil society and citizens) can create and strengthen connections or generate a sense of closeness and trust, resulting in new and improved relationships. This is more likely where open, informal dialogue is enabled, where the 'right' composition of people are in the room, and in contexts where existing networks are weak or dysfunctional but there is a positive wider discourse in support of EIPM. Where participants have the motivation or opportunity to utilise new relationships, they can be used to share information or advice, or can lead to new organisational collaborations (CIMO 4).

When the 'entry point' is through interventions at organisational level…

- Providing technical support to co-produce tools or systems that facilitate staff to use evidence more effectively, where this is done in a collaborative and innovative way, can generate good examples that 'showcase' the value of evidence for quality, performance and delivery. These 'showcases' provide user-friendly decision support tools that help individuals use evidence, but also build understanding and buy-in among senior staff about the value of evidence for decision making, resulting in examples 'diffusing' out to inspire new reforms elsewhere (CIMO 8).
- Where there is pressure to improve performance from senior levels and where an external partner has established trust through previous activities, this can enable an 'accompaniment' mechanism: high-level stakeholders give partners the permission to provide ongoing, tailored support to help them embed

EIPM. This can lead to uptake of recommendations from processes facilitated by the partner, adoption of tools or systems, and possibly the emergence of an internal unit to 'own' and 'champion' EIPM (CIMO 9).

- Providing technical support to co-produce tools or systems that facilitate staff to use evidence more effectively can spark a high-level decision to formally adopt the tools or systems to help standardise EIPM within the organisation. This is more likely when they link to other government procedures and are backed by sufficient authority. Adoption can be on a small scale (e.g. adopting templates), but, in a context where there are high-level government 'owners' of EIPM, adoption can also be large scale (e.g. adopting a comprehensive policy and planning system to promote, embed and monitor the quality of evidence use throughout the policy cycle and into the future) (CIMO 10).

Organisational level change can then filter down to influence individual behaviour through:

- Tools or systems to promote EIPM sparking a *facilitation* mechanism – providing practical assistance enabling people to do their jobs better / more easily. This results in the system or tool being used, and (potentially) increasing the value of evidence through demonstrating the benefits it can bring (CIMO 11).
- Tools or systems that involve positive or negative incentives to adopt EIPM behaviours sparking a *reinforcement* mechanism, in which positive incentives or risk of negative consequences influence behaviour, and lead to individuals deciding to change the way they access, appraise or apply evidence in decision making (CIMO 12).

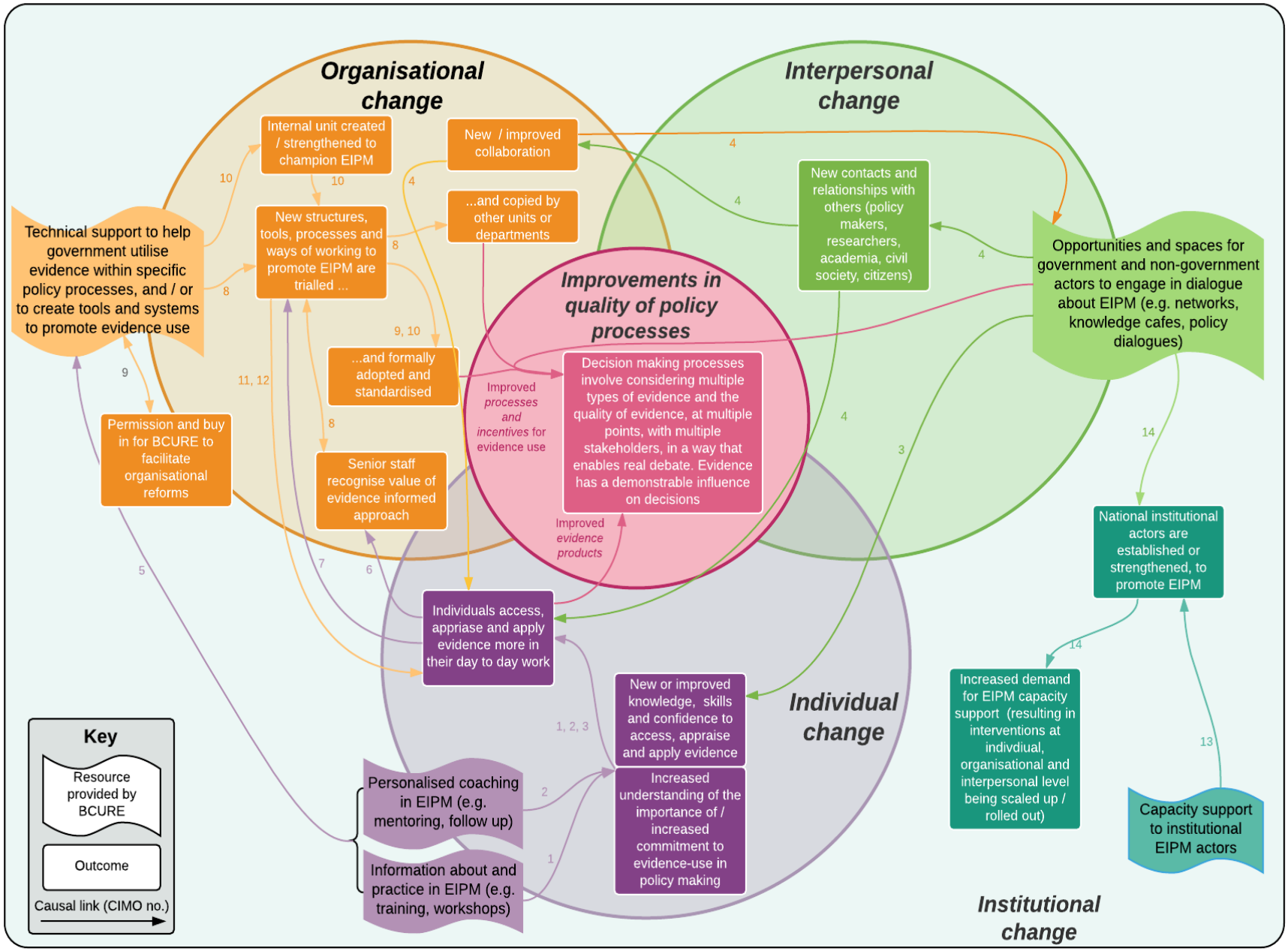
When the 'entry point' is through interventions at institutional level…

- Supporting local organisations to deliver EIPM capacity building activities (directly through organisational capacity support, and/or indirectly through providing opportunities for national partners to 'learn on the job'), can strengthen organisational capabilities through 'learning by doing.' This can result in the establishment or strengthening of national institutional actors, which can act as a 'hub' for EIPM, are capable of running successful programmes to promote it and are potentially able to continue supporting it once the programme has ended (CIMO 13).
- Where local organisations successfully deliver programme activities and/or explicitly aim to build relationships with government departments and other EIPM actors, this enables partners to 'relate and attract' – providing exposure to new collaborators. This can lead to increased demand for partners to provide capacity building support for EIPM from new actors not originally targeted by the programme – which can provide a crucial entry point where there are sensitivities around influencing government decisions, and hence where it is difficult for 'outsiders' to gain entry to government organisations (CIMO 14).

Capacity change at individual, interpersonal, organisational and institutional level combines to contribute to improvements in quality of policy processes through:

- Improving *evidence products* (i.e. how evidence is prioritised, analysed, visualised and presented in briefing notes, evaluations etc), which feed better quality or additional types of evidence into decision making processes.
- Improving *processes and incentives* for evidence use – facilitating and incentivising decision makers to participate in policy development processes that involve explicit consideration of evidence.

Programme theory at the end of Stage 2



5. Stage 1 Common Theory of Change

BCURE evaluation: CToC narrative[15]

The BCURE CToC gives the evaluation a consistent and robust overarching framework to frame the BCURE interventions and the evaluation activities. This initial version of the CToC (December 2014) is preliminary and will be refined at stages as the evaluation progresses. The evaluation team developed the current version of the CToC, following a review of the BCURE partners' proposals and documentation. We also drew on the Evidence Review, as well as on the team's expertise and knowledge of the field.

The Theory of Change depicts the **activities** (interventions and outputs) of BCURE providers. These involve **individual-level interventions** (such as training); **interpersonal-level interventions** (such the use of 'evidence champions' in organisations, and the development of policy and evidence networks); and **organisational interventions** (including the development of policies, systems and procedures for evidence use). These activities predominantly target high-level government policy makers (such as ministerial staff) and mid-level government policy makers (such as mid-level civil servants).[16]

These interventions are anticipated to lead to change at **individual, interpersonal, organisational and institutional** levels. Change at each of these four levels is expected to influence changes in others, in non-linear ways.

At **individual level**, BCURE activities will improve the skills and knowledge of targeted stakeholders, increasing their capacity for EIPM. Activities will also result in increased positive intention among and commitment of individuals to use evidence, and in individuals placing greater value on evidence in their work. At **interpersonal level**, organisational 'champions' will endorse EIPM and help move the agenda forward in their institutions; and networks will be developed and strengthened between national and international institutions, providing an environment for learning and engagement.

Both direct interventions and short-term changes are expected to contribute to **organisational-level change**, including the development of systems and procedures, policies and guidelines, and professional development opportunities, which together will support and incentivise EIPM. Individual, interpersonal and organisational-level change will also contribute to change at **institutional level**, including increased interest in EIPM within civil society, the media and the public, facilitating these actors to more effectively engage with EIPM.

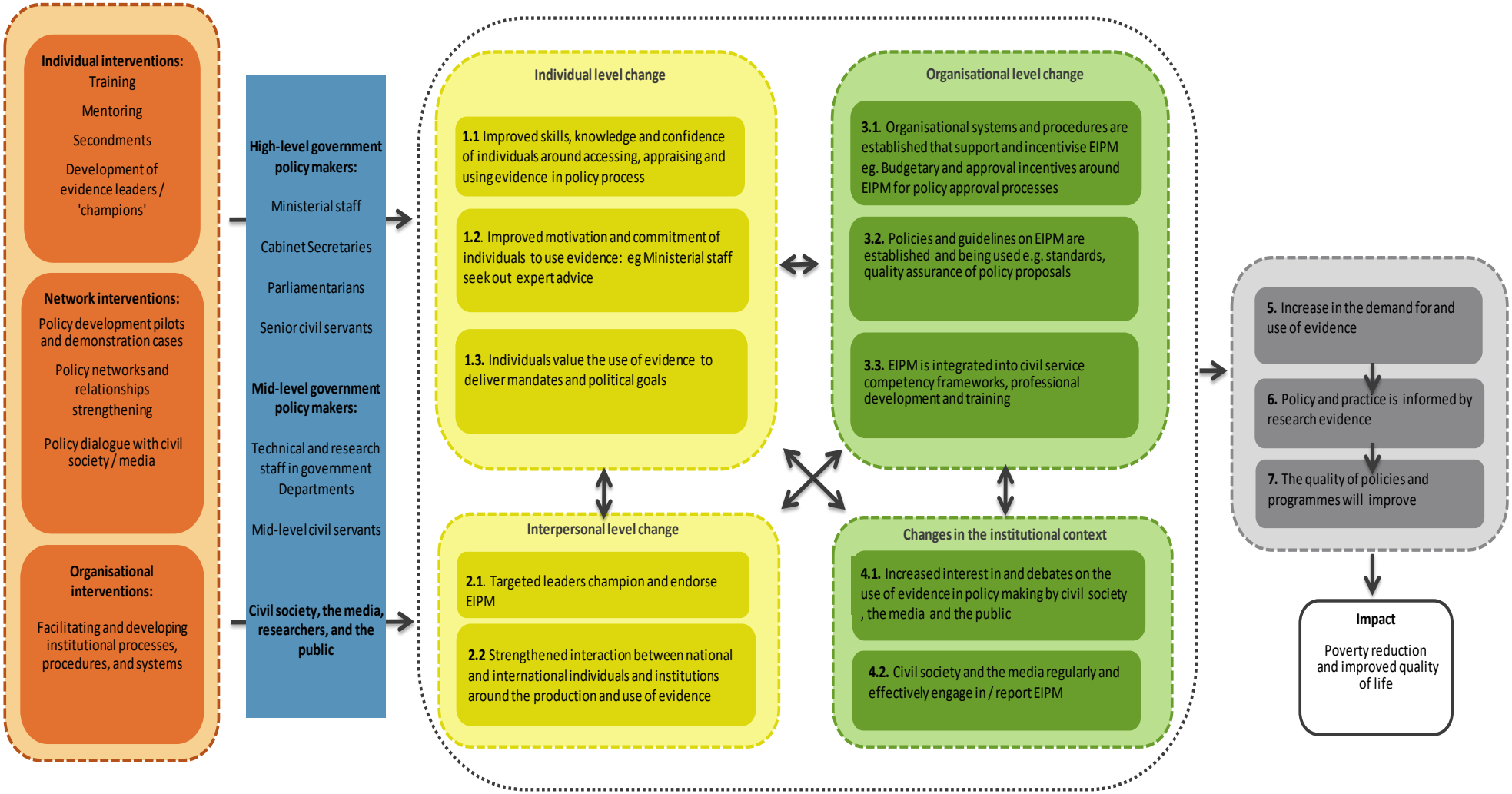
Finally, the combination of individual, organisational, network and institutional change will **increase demand for and use of evidence** among targeted stakeholders, which will result in **policy and practice being increasingly informed by evidence**. This in turn will lead to **improved quality of policies and programmes**. These long-term changes will lead to the programme impact: **poverty reduction and improved quality of life.**

The Theory of Change can be summarised in two sentences as follows:

Developing the capacity of decision makers to use research evidence (through building knowledge, skills, commitment, relationships and systems) will allow them to access, appraise and apply good-quality evidence more effectively when forming policy. This will improve the quality of policies, ultimately benefiting more poor people.

[15] As described in Section 3.2, this CTOC was developed during the early stages of the evaluation and has since been developed into a realist programme theory.

[16] Although we have presented these at the left-hand side of the diagram for ease of reading, BCURE partners are planning interventions at different entry points across the Theory of Change.


Individual interventions:
Training
Mentoring
Secondments
Development of evidence leaders / 'champions'
Network interventions:
Policy development pilots and demonstration cases
Policy networks and relationships strengthening
Policy dialogue with civil society / media
Organisational interventions:
Facilitating and developing institutional processes, procedures, and systems
High-level government policy makers:
Ministerial staff
Cabinet Secretaries
Parliamentarians
Senior civil servants
Mid-level government policy makers:
Technical and research staff in government Departments
Mid-level civil servants
Civil society, the media, researchers, and the public
Individual level change
1.1 Improved skills, knowledge and confidence of individuals around accessing, appraising and using evidence in policy process
1.2. Improved motivation and commitment of individuals to use evidence: eg Ministerial staff seek out expert advice
1.3. Individuals value the use of evidence to deliver mandates and political goals
Interpersonal level change
2.1. Targeted leaders champion and endorse EIPM
2.2 Strengthened interaction between national and international individuals and institutions around the production and use of evidence
Organisational level change
3.1. Organisational systems and procedures are established that support and incentivise EIPM eg. Budgetary and approval incentives around EIPM for policy approval processes
3.2. Policies and guidelines on EIPM are established and being used e.g. standards, quality assurance of policy proposals
3.3. EIPM is integrated into civil service competency frameworks, professional development and training
Changes in the institutional context
4.1. Increased interest in and debates on the use of evidence in policy making by civil society , the media and the public
4.2. Civil society and the media regularly and effectively engage in / report EIPM
5. Increase in the demand for and use of evidence
6. Policy and practice is informed by research evidence
7. The quality of policies and programmes will improve
Impact
Poverty reduction and improved quality of life

6. BCURE programmes' implementation experience 2016

Intervention name	Lead implementing partner	Summary of implementation experience as of May/June 2016:[17]
Strengthening Capacity to Use Research Evidence in Health Policy (SECURE Health) *Budget: £2,279,177*	**African Institute for Development Policy (AFIDEP)**	**Activities undertaken:** SECURE has completed the development of guidelines for evidence use for the MoHs and parliaments of Kenya and Malawi. In Kenya, the guidelines include an additional module on systematic policy development procedures, in response to the gap identified by the evaluation last year. In both Kenya and Malawi, the guidelines were given Cabinet level sign-off in May 2016. Four interns, from Kenya and Malawi, completed a one-month internship programme at UK POST from October to November 2015 and from February to March 2016. Two science policy cafés have been held, one on cholera in Kenya, and in Malawi on the health system's ability to respond to outbreaks, with a focus on Ebola. Work on the national health research framework has been completed in Malawi, but the Research for Health (R4H) Policy Framework in Kenya has been severely delayed owing to difficulties in gaining the required consensus from national and county-level administrations. **Ongoing or new activities:** Completion of the prospective support to policy development processes and identification of recommendations to improve this; completion of the mentorships; dissemination of lessons. **Adaptation to plans in response to contextual challenges:** Follow-up training and mentoring have been less successful through 2015–16 owing to pressures of workload for participants, and some issues with the design of the follow-up activities. As a result, trainees were organised into two groups to work on collective policy briefs. The science policy café strategy has been redesigned in both countries. The internal mid-term review of the programme recommended that the strategy be shifted towards supporting actions arising from the Kenyan café on free maternal health care and the Malawian café on research to policy linkages. **Progress against milestones:** SECURE programme has met or exceeded the majority of its milestones in both Malawi and Kenya, with some key exceptions – in particular around the development of the Kenya Research for Health Policy Framework. **Expected end date:** October 2016
African Cabinet Decision-Making Programme (ACD) *Budget: £3,118,031*	**Adam Smith International (ASI)**	**Activities undertaken:** In Sierra Leone, the project has built on successes achieved in the first phase of the programme, including the Cabinet adoption of a new manual for Cabinet procedures, including an emphasis on requests for evidence in proposals to Cabinet. As of the Stage 1 evaluation, new and revised structures had been established to support organisational change, including a Cabinet Policy Review Unit (CPRU), a Cabinet Implementation Monitoring and Strategy Unit (CIMSU) and a Cabinet Focal Point (CFP) network to be the point for liaison between the Secretariat and each ministry. The Stage 2 evaluation found that the establishment of the CFP network has been reinforced through regular meetings and training, the CPRU has reviewed the majority of strategic memos to the Cabinet in line with its targets and the ACD team has provided ongoing advice to support the Secretariat through the national staff member embedded in the Secretariat and the team leader. In Liberia, there has been training for CFPs provided in cooperation with the Liberia Institute of Public Administration (LIPA); discussion of Cabinet procedures at the Cabinet retreat; the creation of a Cabinet database of memos and Cabinet decisions; provision to the president of an advisory paper on transition issues; and a process review undertaken by the ASI/ACGN national advisor. At the international level, the ACGN network has run meetings for Cabinet secretaries and provided training for policy analysts. A draft toolkit for Evidence Informed Cabinet Decision

[17] Source: BCURE Stage 2 Programme Evaluation reports, 2016.

Intervention name	Lead implementing partner	Summary of implementation experience as of May/June 2016:[17]
		Making was developed and submitted for review to the ACGN members, and two training workshops for policy analysts were held. **Ongoing or new activities:** In Sierra Leone, training for line ministries took place after the Stage 2 data collection had taken place, scheduled for July 2016. **Adaptation to plans in response to contextual challenges:** While there have been some delays in the programme because of various contextual factors including the Ebola epidemic constraining the ability of the government to undertake 'business as normal', with the exception of in South Sudan the overall plan of action remains unchanged since 2015. In South Sudan, activities remained extremely constrained in 2015 and 2015 owing to continued conflict, economic crisis and major infrastructure problems, as well as delays in the formation of the transitional government that affected the humanitarian and development community's ability to implement plans. These conditions led to ACD programme activities being halted to a large extent in South Sudan for a significant proportion of 2015/16. **Progress against milestones:** ACD is partially meeting its output milestones in relation to the original logframe, with progress most pronounced in Sierra Leone where the programme has made significant progress towards its final milestones. However, a number of key outputs had not yet been achieved, including (1) launch of the Cabinet manual at a donors' meeting; (2) the establishment of standing committees; and (3) training for line ministry staff delayed until July 2016. In Liberia there have been a number of significant steps towards the programme outcome but a major constraint is the staffing of the Cabinet Secretariat – two key staff resigned and replacement has been slow. The main activities in South Sudan in the 12 months from June 2015 have been the participation of two key Sudanese Cabinet Ministry staff in the international policy analyst workshops and the ACGN roundtable. In addition ASI/ACGN have worked closely with DFID to determine if and when to resume programme activities. **Expected end date:** A three-month no-cost extension has been agreed, so the milestones originally scheduled for August 2016 are now for November 2016.
Building Capacity for the Use of Research Evidence *Budget: £1,628,000*	**ECORYS**	**Activities undertaken:** Implementation of this project did not begin until October 2015, owing to a number of delays following the shift of the project from Zambia to Bangladesh (a result of Zambian stakeholders deciding not to proceed), and then further delays to the approval of the project. To date, draft EIPM Guidelines have been produced and training on the guidelines delivered to staff within the three pilot ministries (Commerce); Health and Family Welfare; and Environment and Forests. In addition, a training plan and Needs Analysis competed, a Training of Trainers course delivered and a draft EIPM training course developed and signed off by Cabinet Division. **Ongoing or new activities:** Over the next year, the EIPM training course will be delivered to Cabinet Division staff, and the draft EIPM Guidelines updated based on pilot results at end of year 1, to be signed off by Cabinet Division. **Adaptation to plans in response to contextual challenges:** There have not been any substantive changes in plans for BCURE in Bangladesh since the inception phase. Implementation was delayed pending approval of the government's Technical Assistance Project Proposal **Progress against milestones:** At this early stage, the programme is substantially on track against planned activities as per the workplans. **Expected end date:** March 2017
Towards a Culture of Evidence: Building Capacity for Evidence-Based Policy *Budget: £3,232,462*	**Harvard University**	**Activities undertaken:** Training needs assessments and individual and organisational constraints assessments were conducted in Pakistan. The latest figures accessed for the Stage 2 evaluation showed 632 civil servants have been trained to date – 305 in Pakistan and 327 in India. Six pilot projects are underway or completed. Five policy dialogues have been held (three in India and two in Pakistan), with a total of 196 participants. **Ongoing or new activities:** The full training course (six modules) was expected to be completed in Pakistan, Bangladesh and Nepal by August 2016. An Open Enrolment Course is due to be rolled out and facilitated by the end of 2016 in Pakistan. Eight case studies are underway, summarising lessons from the pilot projects. One further policy dialogue will be held in Pakistan.

Intervention name	Lead implementing partner	Summary of implementation experience as of May/June 2016:[17]
		Adaptation to plans in response to contextual challenges: As part of the UK government's refocus of their relationship with India, the BCURE project in India came to an end. A sustainability plan for India was developed to reflect on short to long-term programming implications. In the short term, BCURE Harvard will be able being able to continue the implementation of training without additional support. For longer-term sustainability, and dependent on the funds available, BCURE Harvard is exploring the possibility of employing a regional training officer who could oversee activities from Delhi. This person would have a regional focus but would also work with EPoD India, and EPoD headquarters. With the wind-down of activities in India there has been a reorientation of BCURE activities regionally, with a shift in emphasis towards working in Nepal and Bangladesh, which are still in the early stages programming. In February 2016, a pool of local trainers attended a training of trainers at Harvard with the aim of delivering training in Bangladesh in and Nepal from mid to late 2016. **Progress against milestones:** The Stage 2 evaluation found that the majority of milestones had been met. In Pakistan, five pilot projects have been delivered, exceeding the milestones set out in the logframe. Some delays to the final policy dialogue and the Open Enrolment Course may require a no-cost extension to be negotiated with DFID, in order to deliver activities by the end of 2016. **Expected end date:** August 2016, although there may be a no-cost extension to accommodate delays to some remaining activities, described above.
VakaYiko Consortium *Budget: £3,397,924*	**INASP**	**Activities undertaken:** In Zimbabwe, consortium partner ZeipNET completed EIPM training with participants from Parliament, the Ministry of Industry and Commerce and the Ministry of Youth, Indigenisation and Economic Empowerment. A mentorship scheme has been launched, implementing Action Plans with each of these three organisations, through two mentees from each. ZeipNET has also held two policy dialogues and two knowledge cafés this year. In Ghana, three sets of training were completed with participants from ministries and parliament, and three policy dialogues were completed by consortium partner GINKS. In both countries, VakaYiko has provided ongoing capacity support to ZeipNet and GINKS through a combination of formal training and informal support to ongoing activities. In South Africa, ongoing support has been provided to the process of developing an evidence strategy within the DEA. This has involved workshops with different teams within the DEA, and the submission of an improvement plan and monitoring and learning framework. Several documents summarising lessons from the work in South Africa have been drafted. A new workstream has been launched in Uganda, following a request from the director of the Uganda Department of Research Services for training. Three training modules have been delivered and a training policy for the Institute of Parliamentary Studies has been drafted and reviewed. At a global level, nine EIPM capacity building projects have been funded through the small grants workstream (all either complete or underway), and three case studies have been published to report on lessons learned. An EIPM toolkit has been developed and published, drawing on learning from the EIPM training courses. **Ongoing or new activities:** In Zimbabwe, the mentoring programme is still ongoing and learning and exchange visits are planned by the end of 2016. The new workstream in Uganda is underway, with various upcoming activities including sensitisation activities for MPs during 'Research Week', a paring scheme between Department of Research Services Science and Technology Section and the Uganda National Academy of Sciences, support to increase skills for science communication, and external events aimed at building Department of Research Services networks with local research institutes. Under the extension, ZeipNET will also work with the Ministry of Youth, Indigenisation and Economic Empowerment to roll out an evidence strategy, drawing on lessons from the South Africa work with DEA. The evidence strategy work will also be trialled in Ghana. There are also plans to pilot the VakaYiko toolkit with the Ghana Institute of Management and Public Administration and to facilitate a meeting of research departments within the parliaments of Uganda, Zimbabwe and Ghana, to enable reflection and learning exchange

Intervention name	Lead implementing partner	Summary of implementation experience as of May/June 2016:[17]
		Adaptation to plans in response to contextual challenges: There have been few changes to the programme's plans since Stage 1. A project extension has resulted in new activities being developed, described above. In addition, the mentorship scheme in Zimbabwe experienced a setback when a learning exchange visit to the UK (scheduled for the six Zimbabwean mentees in May 2016) was called off by DFID in April. DFID explained that the event was cancelled owing to concerns from the High Commission in Harare regarding how the visit might be viewed. The programme team are planning to replace the event with exchanges to Accra later in the year, tying in with the Consortium meeting. **Progress against milestones:** The Stage 2 evaluation found that the VakaYiko programme has achieved the vast majority of its milestones this year, and has implemented some additional activities over and above its milestones (e.g. policy dialogues in Ghana). **Expected end date:** DFID has granted an extension to the programme to **March 2017.** This encompasses the new strand of work in Uganda as well as various extension activities detailed above.

7. Programme evaluation assessments of change against each EQ

This Annex presents each Stage 2 programme evaluation assessment of change as a result of the six BCURE programmes, at individual, interpersonal, organisational, institutional and policy level.

EQ 1: Individual change 1.2 What were the observable changes in individuals' knowledge and skills?		
BCURE Programme	**Extent of observed change**	**Summary of evidence on outcome patterns**
Harvard (Pakistan)	Reliable evidence of moderate change Evidence that BCURE made a crucial contribution to this change (+++)	There is reliable evidence that suggests the blended BCURE training has made a crucial contribution to the acquisition of knowledge and skills among trainees. However, there is more limited evidence of trainees putting their skills into practice. There is some indication that this is hindered when participants do not have the opportunity to apply their new learning in their professional capacity.
SECURE Health (Kenya)	Strong evidence of moderate change Evidence that BCURE made a crucial contribution to this change (+++)	There was strong evidence of increased awareness and enthusiasm for EIPM, increased confidence, knowledge and skills and applications of EIPM skills (behaviour change). Other changes included improvements to respondents' ability to do their work, perceived improvements in policy briefing products and adaptation of curriculum and tools to train others, especially at the county level. Although it has been 12 months since the training and participants had received mentoring and follow-up, most respondents felt it was still early days and a critical mass of trained individuals had not yet been achieved. Therefore, the extent of change has been rated as moderate. However, as no other contributing factors have been identified apart from SECURE's interventions, the programme's contribution to observed changes has been rated as crucial.
ECORYS (Bangladesh)	Reliable evidence of early change Evidence that BCURE made some contribution to this change (+)	There have been some early improvements in EIPM knowledge. The initial activities undertaken in support of enhancing individual capacity, e.g. Training Needs Assessments, development of training materials, are seen as well aligned to BCURE's objectives. In addition, respondents are positive about key aspects of implementation strategies that might impact individual-level results, e.g. reliance on national consultants for training, including follow-up support by consultants at ministry level, as well as reliance on case materials relevant to Bangladeshi actors. Nevertheless, cautions were offered regarding the approach as well as overall prospects for sustainability, e.g. regular transfers of government staff, with risks of diminished impacts at ministry or division level; 'missing foundations' in government, which could limit enhanced and/or more systematic reliance on EIPM by trained individuals;[18] weaknesses in national training institutions as risks to the ongoing provision of adequately trained personnel; and changes in senior leadership in government partners), which could potentially undermine gains realised, including application of skills gained through training or mentoring.
VakaYiko (Zimbabwe)	Reliable evidence of established change Evidence that BCURE made a crucial	Primary data from Zimbabwe and monitoring data from Zimbabwe and Ghana demonstrates that training has been successful in building trainees' knowledge and skills, and trainees have been able to put their knowledge into practice in the months following the training. Many of these changes are directly attributable to the VakaYiko course. This success seems strongly linked to the fact that training was directly tailored to the day-to-day work trainees were already doing. Management buy-in has proved essential to allow trainees (most of whom are

[18] 'Missing foundations' refers to a situation within government characterised by insufficient institutional or procedural structures or resources to formalise policy or decision making, potentially leading to decisions that, among other things, are not sufficiently or systematically informed by evidence.

	contribution to this change (+++)	relatively junior) to apply their skills in the workplace. This has created some obstacles for trainees, and as a result VakaYiko has increasingly attempted to include more senior staff in its courses. There is some evidence that policy dialogues and knowledge cafés have encouraged knowledge sharing between stakeholders from different sectors, helping raise awareness of EIPM issues and specific policy areas.
UJ-BCURE (South Africa)	Reliable evidence of established change Evidence that BCURE made a crucial contribution to this change (+++)	The programme has, through its workshops, increased participants' understanding of EIPM terminology and skills and reinforced the importance of working in an evidence-informed manner. In almost all reviewed cases, both the individual and team mentorships deepen the mentees' ability and confidence to engage with evidence and there is evidence that some mentees are thinking and working differently and may emerge as future evidence champions. The AEN events and workshops provide platforms for sharing EIDM challenges and solutions.
ACD (Sierra Leone)	Strong evidence of moderate change Evidence that BCURE made a crucial contribution to this change (+++)	ACD has facilitated direct training for three Sierra Leone policy analysts and all CFPs as well as providing on-the-job support to the Cabinet Secretariat and CFPs in the implementation of new Cabinet procedures. Pre and post-test scoring for the international training for analysts shows overall positive change though some disaggregated data available for Sierra Leone participants shows some reverse in scores. Participants report and monitoring data shows high levels of satisfaction with both CFP and analyst training and workshops. The focus, especially for CFPs, has been on implementing the new procedures more than on use and quality of evidence. So far there has been only limited support to others beyond the CFPs and the three analysts in line ministries. Particularly important are the professional, technical staff in each ministry. Future training is planned to include them.

EQ 2: Interpersonal change **1.2 What were the observable changes in networks and relationships?**		
BCURE Programme	**Extent of observed change**	**Summary of evidence on outcome patterns**
Harvard (Pakistan)	Reliable evidence of moderate change Evidence that BCURE made an important contribution to this change (++)	There is reliable evidence to suggest policy dialogues are bringing people together, bridging the gap between members of the civil service, academia, think tanks and research institutions. In effect the policy dialogues are a platform from which interpersonal relationships between policy makers and researchers can be forged. Interviews with BCURE programme staff and one project participant suggested that the policy dialogues helped catalyse the Crime Mapping and Polio Vaccines project.
SECURE Health (Kenya)	Partial evidence of early change Evidence that BCURE made an important contribution to this change (++)	There is only partial evidence relating to examples of interpersonal change. It was not as strong a theme at Stage 2 as at Stage 1. This may be because the main interventions designed to promote networks and relationships – science policy cafés and research–policy events – were last held late in 2015. Nevertheless, SECURE has taken care to involve a range of participants from across MoH and parliaments in Kenya and Malawi in its interventions to build broad-based acceptance of its interventions, with emerging positive interpersonal change outcomes. These include collaborative linkages across teams in the Kenya MoH. However, the evaluation found very limited improvements in linkages among policy and research stakeholders arising from these interventions, which is a key focus of the programme.

ECORYS (Bangladesh)	No evidence of change	BCURE is not explicitly committed in its results statements to 'interpersonal level change'. There is no evidence of change at this stage although some of its early and planned activities are likely to yield benefits in this area over time, e.g. combining participants from different groups or sectors in workshops, training sessions and working groups, and sharing information on intervention benefits and lessons learnt with diverse stakeholders.
VakaYiko (Zimbabwe)	Strong evidence of moderate change Evidence that BCURE made an important contribution to this change (++)	Primary data and monitoring and implementation documents from Zimbabwe suggest the policy dialogues and knowledge cafés have encouraged interaction and, dialogue between stakeholders from different sectors, with some examples of new relationships directly attributable to the programme. ZeipNET has also contributed to ensuring the two policy dialogues held this year have concrete action points that can be tracked over time. Several trainees felt they had gained new networks or stronger relationships with other participants on the EIPM course, and two trainees gave examples of synergies between training and knowledge cafés or policy dialogues. However, there are few examples as yet of trainees using these new networks in their work.
UJ-BCURE (South Africa)	Reliable evidence of moderate change Evidence that BCURE made an important contribution to this change (++)	There is reliable evidence that the AEN events and strong evidence that workshops provide platforms for sharing EIDM challenges and solutions. Participation in these events validates the thinking of delegates, and reinforces their commitment to EIDM. Membership growth of the AEN is remarkable and largely driven by various BCURE funded events. There are early indications that the AEN may emerge as a hub for EIPM on the continent, but the value-added of the AEN is mostly still anticipated and not established. Long-term sustainability of the AEN is dependent on sustained membership growth, and high-quality material in the newsletters and on the website continues.
ACD (Sierra Leone)	Reliable evidence of moderate change Evidence that BCURE made an important contribution to this change (++)	There are two sets of relationship or interpersonal change to which the programme has contributed. The first is the relationship between CPRU and line ministries, particularly the CFP but in some instances also the minister. The participation of the ACD embedded national advisor in Cabinet meetings (as note-taker) and the efforts made by the CPRU team to create a sense of collaboration have built good relationships with CFPs. Second, within ministries, the procedure to create memos detailing proposals for Cabinet has potential to create a team approach to developing proposals. It is early days in the application of procedures and in places held back by CFPs' limited access to professional/technical personnel. In addition, the role of the embedded ACKN advisor in the CPRU has been crucial and there is uncertainty about the future of this role when funding ends.

EQ 3: Organisational change **1.2 What were the observable changes in individuals' knowledge and skills?**		
BCURE programme	**Extent of observed change**	**Summary of evidence on outcome patterns**
Harvard (Pakistan)	Reliable evidence of moderate change Evidence that BCURE made an important contribution to this change (++)	There is reliable evidence that the programme has made an important contribution to organisational change through pilot projects, alongside internal champions and national technical experts. The pilot projects are spurring short-term increases in evidence use, and providing the tools and systems on which long-term solutions for EIPM can be based. A number of the projects are providing a means by which diffuse data can be consolidated in a single location for easy interpretation and efficient decision making. There is also emerging evidence that the pilots are motivating the development of tools that could influence the systematic use of tools at organisational level. There is reliable evidence that training of trainers has made a crucial contribution to developing a pool of trainers who are currently providing BCURE training in Pakistan. This is an important step towards building a sustainable training of trainers model for fostering talent to create change at an interpersonal level. There is potential to build on the pool of

		trainers moving forwards but attention needs to be given to ensuring trainers are retained as the model seeks to respond to demand.
VakaYiko (Zimbabwe)	Reliable evidence of early change (in Zimbabwe – change is more pronounced in South Africa and Ghana) Evidence that BCURE made an important contribution to this change (++)	Primary data and implementation documents from Zimbabwe suggest ZeipNET has been successful in securing government buy-in and access – in part because of existing high-level interest in EIPM within government and in part because of ZeipNET's sensitive approach in a politically charged environment. Training has provided an 'entry point' for support to organisational reforms through Action Plans developed by trainees. Organisational reforms are in progress but it is too early to see evidence of concrete change, and the programme has not yet formally monitored progress on the mentorship. Trainees suggest they are starting to influence their managers in relation to EIPM, but as yet there are no clear examples of managers taking steps to promote EIPM as a result of this influence. Other than the allocation of staff to the Ministry of Youth, Indigenisation and Economic Empowerment Research and Policy Unit, there are also few signs that government departments are starting to invest their own resources in EIPM activities as a result of the programme, and it is unclear how feasible this is within the resource- constrained environment of Zimbabwe. Programme documentation suggests organisational change is more pronounced in Ghana (where training courses are being rolled out more widely) and in South Africa (where VakaYiko is facilitating a number of organisational reforms within DEA). ZeipNET has contributed to ensuring the two policy dialogues held this year have concrete action points that can be tracked over time.
UJ-BCURE (South Africa)	Reliable evidence of moderate change Evidence that BCURE made an important contribution to this change (++)	The application of learning results in specific products informed by evidence may change the way the organisation performs its work by developing new processes that promote more EIDM practice. The evidence suggests that evidence products may offer opportunities for more widespread change, if the products or processes are institutionalised throughout the organisation. At Stage 2, there is a suggestion that pressure for EIPM may create an opportunity for new evidence products when implemented within a conducive organisational environment. The data suggests that evidence products from senior champions for evidence may have further initial reach, but evidence products from champions lower in the organisation can affect change if the champion is sufficiently committed to EIPM and the environment encourages experimentation. These changes will be followed up at Stage 3.
SECURE Health (Kenya)	Strong evidence of established change Evidence that BCURE made a crucial contribution to this change (+++)	There is **strong evidence** to suggest the results of SECURE's interventions are coming together into consolidated change at the organisational level. Respondents see SECURE has provided MoH and parliamentary analysts with a 'roadmap' to implement EIPM and move from having a discourse about EIPM and *ad hoc* approaches to having a shared, comprehensive understanding about EIPM. SECURE has inspired some 'evidence champions', who are introducing new ways of working (albeit incrementally); developed policy guidelines with broad, senior-level ownership; and created the relationships and platform that the R&D team and the parliamentary team need to introduce reforms. The Malawi component on assessing the impact of the implementation of the country's national health research agenda of 2011 has been completed. However, work on the national health priorities framework in Kenya has been severely delayed. This could mean the MoH team could miss bidding for funds in 2017 to implement the framework, constraining the benefits of this important structural platform to support EIPM. The outcomes have a good chance of sustainability, given the new opportunities emerging with the WHO data initiative. It is reasonable to suggest that SECURE has equipped MoH with a strong set of baseline EIPM capacities to enable it to respond to World Health Organization and other opportunities (such as the ISO certification) for adopting EIPM into systematic procedures and systems. In Kenya, respondents confirmed that SECURE had been the only comprehensive support to EIPM in this period, so the results can be directly attributed to the programme as a **crucial contribution**, achieved within a relatively short timeframe.
ACD	Strong evidence of	The programme has introduced new Cabinet procedures that require the provision of evidence in proposals summarised in the memo that accompanies them to

(Sierra Leone)	established change Evidence that BCURE made a crucial contribution to this change (+++)	Cabinet. The procedures, notably the Cabinet memo template and pre-Cabinet briefing, are well established and being used, though there is significant room for progress in the extent to which they drive increased use of quality evidence in the policy making process. They have established a process that demands evidence and is led by the Cabinet secretary and reinforced by presidential support and the CPRU role. The significance of this established change rests on the extent to which that demand is met with high-quality and good use of evidence. This is the next necessary stage to achieve better use of evidence in policy. The role of the embedded ACKN advisor in the CPRU has been crucial and there is uncertainty about the future of this role when funding ends.
ECORYS (Bangladesh)	Reliable evidence of early change Too early to make contribution assessment	There are signs of early change in terms of BCURE's contribution to organisational-level change in Bangladesh. However, this is an area where the project is expected to make a difference in enhancing organisational capacity to support EIPM, especially in the CD, the national training institutions involved and the three pilot ministries. Overall baseline capacity for EIPM in the three pilot ministries was ranked at 60% (taking into account, e.g., skills and knowledge, use of analytical tools), although none of them presently has internal rules on policy formulation. The BCURE team intends to enhance organisational and other capacity shortcomings identified in its assessment processes through 'a comprehensive training programme that includes both on-the-job training in drafting of the policy proposals and classroom training'. While it is premature to measure substantive progress at the institutional or system level, BCURE has committed to an 'Improved institutional framework in support of EIPM' (Output 1) with various supporting activities: developing EIPM guidelines that support the use of evidence; provision of formal training to CD staff on the EIPM Guidelines; and provision of on-the-job training to CD staff on how to manage the introduction of EIPM across pilot ministries.

EQ 4: Institutional change **1.2 What were the observable changes at institutional level?**		
BCURE Programme	**Extent of observed change**	**Summary of evidence on outcome patterns**
Harvard (Pakistan)	N/A	This EQ is not reported as the programme is not directly attempting to influence change at this level. There is some suggestion that policy dialogues provide a platform for the BCURE programme to profile the work they are doing with a wide range of stakeholders.
SECURE Health (Kenya)	N/A	This EQ is not fully reported for SECURE because institutional-level changes are anticipated to emerge in Stage 3. SECURE's main intervention aimed at the institutional level is the national R4H Policy Framework in Kenya, and, to a lesser extent, the impact assessment of the national health research framework in Malawi. Related outcomes are currently reported under EQ 3 (organisational change), as the framework is still in development within the boundaries of the Kenyan MoH. If implemented, the R4H Policy Framework may result in outcomes in terms of changes in relationships and investments between governments at national and county levels, research agencies and development partners and donors. In turn, these changes may, over time, support institutional or system level changes. These changes will be followed up in Stage 3.
ECORYS (Bangladesh)	N/A	This EQ is not reported as the programme is not directly attempting to influence change at this level.
VakaYiko	Strong evidence of	Primary evidence from Zimbabwe and organisational capacity assessments of ZeipNET and GINKS suggest VakaYiko has increased the capacity of its national

(Zimbabwe)	established change Evidence that BCURE made a crucial contribution to this change (+++)	partners to act as EIPM players. This seems to owe to a combination of formal and informal support over time, within the context of international partnerships. ZeipNET is becoming a recognised actor in the Zimbabwean EIPM institutional space, and has received several requests for support and collaboration from new organisations. This success can be directly attributed to VakaYiko. However, more work is required to realise ZeipNET's goal of establishing a national evidence infrastructure, and there are still challenges relating to its size, the inactivity of the Board and sustainability in a context of severe public sector resource constraints. It is difficult to assess the contribution of knowledge cafés to their intended institutional level outcomes (increased interest in EIPM among the general public and the media, leading to growing pressure on government to make evidence-based decisions), given the more systemic and less instrumental nature of what they hope to achieve. VakaYiko has been active in global communications and lesson learning; publishing the EIPM Toolkit is a major highlight. Some lessons have also emerged from the small grants programmes, though it may be useful to consider how to more systematically share these.
UJ-BCURE (South Africa)	Reliable evidence of moderate change Evidence that BCURE made an important contribution to this change (++)	UJ BCURE has managed to build a network around its programme and has become well integrated with key role players in the evidence landscape. This may offer long-term opportunity to promote EIPM practice in the future. Emerging partnerships with established stakeholders involved in EIPM capacity building in South African may also contribute to long-term sustainability of the AEN by promoting the network at other EID/PM events.
ACD (Sierra Leone)	N/A	ACD did not focus on this level so there are no related outcomes. A planned launch of the Cabinet manual as part of the ACD programme has not taken place, which would provide one opportunity for wider engagement. Wider institutional change will be beneficial to the sustainability of ACD's outcomes, particularly if there is greater capacity in Sierra Leone's institutions to produce, interrogate and use evidence as part of policy processes. DFID is a key donor in Sierra Leone and is supporting a number of complementary initiatives, such as the GoSL Ebola recovery programme, which ASI is also involved with, and some programmes that promote the production and/or communication of research, for example those implemented by the International Network for the Availability of Scientific Publications (INASP). The evaluation did not find any links being made by DFID between these programmes at this stage; these may be beneficial for the future.

EQ 5: Policy process change

1.2 What were the observable changes in the nature and quality of policy processes in relation to evidence use?

BCURE Programme	Extent of observed change	Summary of evidence on outcome patterns
Harvard (Pakistan)	N/A	This EQ is not fully reported, as there is currently limited evidence of change at this level. The Pilot Projects are spurring short-term uses of evidence, but they are not contributing to a more systematic change towards quality policy processes. However, there is some evidence of government departments taking an interest, and considering integrating the systems and tools inspired by the Pilot Projects into institutional structures that could lead to more systematic changes in the quality of policy processes.

SECURE Health (Kenya)	N/A	This EQ is not fully reported here because changes in the quality of policy development processes are anticipated to emerge in Stage 3. Examples that will be followed up in Stage 3 include actions arising from the recommendations of the retrospective study, as this is considered by the R&D team to be a 'baseline' study to help it track evidence use and improvements in processes as part of performance management. The policies being followed by the prospective case studies, such as the health financing policies, are already considered to have been 'enriched' through SECURE's involvement (4-3), but substantive changes in process and content have yet to emerge. Another example will be anticipated improvements in quality assurance procedures arising from the roll-out of the EIPM guidelines and the ISO process in Kenya. These changes will be explored in depth in Stage 3.
ECORYS (Bangladesh)	N/A	It is too early to offer any meaningful commentary on this EQ in Bangladesh, especially since the policies to be addressed by the three pilot ministries have yet to be determined. Nevertheless, relevant foundations are being laid through the BCURE project, which should contribute to the quality of policy processes associated with the pilot initiatives to be carried out in the coming months.
VakaYiko (Zimbabwe)	Partial evidence of early change (in Zimbabwe; change is more pronounced in South Africa) Evidence that BCURE made an important contribution to this change (++)	Primary and secondary evidence suggests ZeipNET contributed to the decision to establish the Research and Policy Unit in the Ministry of Youth, Indigenisation and Economic Empowerment, and ZeipNET is providing valuable ongoing support to the functioning of this unit. There is strong evidence for this particular example of influence. Policy dialogues and (to a lesser extent) knowledge cafés have resulted in some concrete follow-up actions, and there are some examples of trainees influencing policy processes through applying the skills learned in training. Where policy change has happened, this seems to be because events were already in motion and the programme helped facilitate, catalyse or improve existing processes. However, the examples are generally drawn from single interviews, and so the evidence of VakaYiko's contribution is only partial at this stage. As yet, there is no evidence of changes to policy processes in Zimbabwe through the longer-term route envisaged in the programme Theory of Change. Secondary evidence suggests VakaYiko's work with DEA in South Africa is contributing to improvements in policy processes within the department.
UJ-BCURE (South Africa)	Reliable evidence of early change Evidence that BCURE made some contribution to this change (+)	There is early evidence that mentoring support focused on specific policy processes may lead to increased and better consideration of evidence in the initial design phase. Evidence suggests initial commitment and capacity for evidence use from the mentee and senior-level entry may be prerequisites for influencing high-level policy processes. There is some indication that external support to the policy process may become limited once the policy enters the decision making terrain.
ACD (Sierra Leone)	Strong evidence of early change Evidence that BCURE made a crucial contribution to this change (+++)	There are some positive changes taking place in the wider Sierra Leone policy processes as a result of the ACD programme. It is early days in this process, but new structures and processes that enable more proposal scrutiny, interaction between ministries, more Cabinet discussion and better-structured presentation of policy are enabling more debate about policy and the evidence to support it. However, the changes often emphasised the summarising and more streamlined presentation of evidence rather than a significant change in the quality or use of the evidence itself.

8. Evidence of outcomes for Stage 2 CIMOs

Annex 7 presents each Stage 2 programme evaluation assessment of change as a result of the BCURE programmes, at individual, interpersonal, organisational, institutional and policy level. This section provides a detailed map of the evidence underpinning the *specific outcomes* in Stage 2 CIMOs, at individual, interpersonal and institutional level. This aims to provide a fully transparent overview of the extent of primary and secondary evidence underpinning our Stage 2 findings.

The evaluation found evidence of a wide array of outcomes, especially at organisational level, which were not included in programmes' documented outcome objectives, for which primary data from the evaluation (mainly interviews) is the only source. The evidence for organisational level outcomes are not mapped below, because this evidence is derived solely from primary evaluation data, and the evidence base is therefore clear from the main report.

This mapping provides a starting point for the evaluation to define more specific outcome indicators for the summative evaluation at Stage 3.

CMIO 1: Evidence underpinning the outcome 'participants changing their behaviour and putting new knowledge and skills into practice in their work', as a result of information and opportunities to practise skills through training

VakaYiko	All 10 interviewed trainees cited concrete examples of how skills had translated into change in their day-to-day work. Trainees' managers and programme facilitators also gave examples of skills application.[19] This was triangulated with six-month follow-up surveys conducted by the programme: in Zimbabwe, 16 of 29 respondents provided qualitative examples of putting skills into practice in the workplace, and 18 of 29 trainees in Ghana.
Harvard	Data from pre- and post-training assessments conducted by the programme show statistically significant learning gains for trainees in India and Pakistan. Outcome diaries maintained by the programme provide some anecdotal examples of individual learning from training being used to inform policy decisions. However, these examples appear to be isolated, and the fieldwork found no supporting evidence to suggest widespread application of knowledge and skills. Logistical challenges meant it was possible to interview only three training participants, but all three stated that they had been unable to put their learning into practice.[20]
UJ-BCURE	Most interviewed workshop participants reported that they had been introduced to relevant terminology and methods, and increased or reinforced their understanding of the importance of evidence.[21] This is triangulated with programme monitoring data, which found improvements in self-reported knowledge and skills among workshop participants. Only two interview respondents said the workshops had led to behaviour change in the workplace.[22]
ECORYS	At the time of the evaluation, only initial awareness raising sessions and training needs assessments had been conducted. There is no evidence as yet of behaviour change.
SECURE	The programme's bi-monthly survey of participants in Kenya and Malawi reported that 35 trainees in Kenya and 18 in Malawi had applied EIPM skills acquired through training workshops. This was echoed by the interview data: 20 Kenyan respondents gave examples of improved skills and applications of these skills in the workplace.[23] There were triangulated reports that parliamentary researchers are digging deeper into issues to explore policy solutions and proactively researching trending topics.[24] However, monitoring data shows that 14 of 40 trainees had completed or were near completion of their policy brief by May 2016.

[19] 2-12, 2-5, 2-6, 2-7, 2-8, 2-10, 2-12, 2-13, 2-15, 2-17, 2-18, 2-24, 2-29, 2-30, 2-32
[20] 1-10, 1-15, 1-18
[21] 3-5, 3-6, 3-21, 3-14, 3-15
[22] 3-15, 3-16
[23] 4-1, 4-2, 4-3, 4-4, 4-9, 4-11, 4-12, 4-14, 4-17, 4-20, 4-21, 4-23, 4-25, 4-30, 4-36, 4-38, 4-39, 4-41, 4-42, 4-45
[24] 4-10, 4-21, 4-23, 4-26, 4-30

ACD	There was no independent assessment of CFPs' skills in use of evidence before or after the training. However, CFPs reported that training and meetings had helped them understand and apply the new procedures established in the Cabinet Manual,[25] and both ACD programme staff and CPRU staff report that there has been an improvement in the presentation of existing evidence (although there is a 'long way to go').[26]

CMIO 2: Evidence underpinning the outcome 'participants putting learning into practice and/or taking on new opportunities to use evidence or promote EIPM in the workplace', as a result of coaching through one-to-one mentoring and/or follow-up activities

VakaYiko	This resource was not provided through a stand-alone intervention and was therefore not monitored. Interviews found ZeipNET had provided *ad hoc* coaching support to trainees, and two respondents felt their work had benefited from this.[27]
Harvard	N/A – resource not provided in this programme.
UJ-BCURE	Six of the 13 mentees involved in the one-to-one mentorship programme were interviewed, along with their mentors. All of these stakeholders plus programme staff gave examples of the mentoring influencing mentees' day-to-day practices.[28] At the time of the Stage 2 data collection, the programme was compiling a mentoring report, but monitoring data on application of learning following mentoring relationships was not yet available. However, programme documents confirmed that most mentoring relationships were renewed after the six-week initial period, triangulating mentees' reports that the relationships were seen as useful.
ECORYS	N/A – resource not provided to date in this programme.
SECURE	All trainees received the opportunity to attend follow-up workshops on evidence access and use. However, programme reports show that these were poorly attended in both Malawi and Kenya. Following consultation with government stakeholders, it was agreed trainees would be organised into two groups to work on collective policy briefs, supported by a mentor. Interviews suggested this group mentoring approach resulted in an improvement in the engagement and completion rates of policy briefs.[29] However, a smaller number of trainees felt group training was less relevant to individuals.[30] In Malawi, programme reports suggest the group mentoring approach was not successful, but do not provide any reason for this. There is no monitoring data relating specifically to the group mentoring intervention.
ACD	N/A – resource not provided through a stand-alone intervention but as part of technical support; therefore evidence is discussed in relation to organisational change.

CIMO 3: Evidence underpinning the outcome 'new knowledge and awareness about the importance of EIPM, what EIPM is and how to conceptualise it within a particular national context, and how others have used evidence or dealt with challenges in different contexts', as a result of taking part in facilitated spaces for dialogue

VakaYiko	Seven interviewed stakeholders who had taken part in knowledge cafés and policy dialogues reported that these fora encouraged knowledge sharing and awareness raising.[31] Observational monitoring data collected by the programme did not directly monitor knowledge and awareness, but reported positive interactions between event participants, balanced discussion between different groups of participants and active engagement from the general public and female participants at both knowledge café events and one policy dialogue, although observations at the fourth policy dialogue suggested it was less successful.[32]
Harvard	There is no evidence from Harvard in relation to this outcome.

[25] 5-19, 5-28, 5-29
[26] 5-3, 5-12, 5-14
[27] 2-8, 2-15
[28] 3-2, 3-4, 3-5, 3-6, 3-7, 3-9, 3-16, 3-18, 3-20, 3-21, 3-22, 3-W, 3-W2
[29] 4-9, 4-11, 4-17, 4-20, 4-21, 4-38
[30] 4-12, 4-17
[31] 2-5, 2-18, 2-11, 2-20, 2-23, 2-25, 2-28
[32] The programme uses an observational rubric to measure 'interaction, facilitation, power dynamics and the use of research evidence' at its knowledge café and policy dialogue events.

UJ-BCURE	Four workshop participants interviewed for the evaluation suggested UJ-BCURE workshops provided opportunities for sharing experiences and challenges and learning from each other.[33] Programme monitoring did not measure this specific outcome.
ECORYS	N/A – resource not provided to date in this programme.
SECURE	Two interview respondents felt that science cafés raised awareness.[34] There is no monitoring data relating to this outcome.
ACD	Interviews with three participants in international annual meetings of Cabinet Secretaries across the region and international training for policy analysts suggested they had gained new knowledge or awareness of issues relating to EIPM as a result of participating in group conversations and dialogue.[35] However, there is no monitoring data relating to this outcome, and pre-and post-workshop skills tests for the Sierra Leone participants showed a reverse in their scores during the course of the workshop.

CIMO 4: Evidence underpinning the outcome: a) connections with stakeholders who can enable respondents to do more than they would be able to on their own and b) sharing knowledge that benefits individuals' work, or new organisational collaborations, as a result of facilitated spaces for dialogue

VakaYiko	Seven interview respondents said they had made new connections or gained better awareness of networks through knowledge cafés,[36] and two others felt policy dialogues had provided an opportunity for policy makers to hear from non-government stakeholders.[37] Five respondents gave examples of these new networks being used for either individual knowledge sharing or advice, or organisational collaboration over policy issues.[38] Observational monitoring data collected by the programme reported positive interactions between event participants, balanced discussion between groups of participants and active engagement from the general public and female participants at both knowledge café events and one policy dialogue in 2015/16, although observations at the fourth policy dialogue suggested it was less successful.[39] Participant lists from all networking events show a diversity of stakeholders attended the majority of events, including stakeholders from government ministries, civil society, industry and the media.
Harvard	Two interview respondents said policy dialogues helped bridge gaps between members of the civil service, academia, think tanks and research institutions, creating a platform from which relationships could be forged.[40] Both respondents also suggested policy dialogues helped catalyse specific pilot projects. Policy dialogue reports include programme observations that the dialogues were successful in bringing together key stakeholders, but do not include any reflections collected from participants.
UJ-BCURE	Several programme participants (including multiple participants in a programme workshop) reported that they had made new connections as a result of BCURE activities, including workshops and the AEN,[41] and the mentoring activities.[42] Two respondents gave examples of contacting/pursuing collaborations with people they had been introduced to through BCURE.[43] There is no monitoring data relating to this outcome.
ECORYS	N/A – resource not provided to date in this programme.
SECURE	Five interview respondents felt their personal networks within the ministry had improved as a result of their participation in the SECURE training (which brought together individuals from different units across the ministry).[44] Two respondents felt other SECURE activities (group

[33] 3-4, 3-20, 3-11, 3-15
[34] 4-25, 4-36
[35] 5-10, 5-13, 5-16
[36] 2-10, 2-11, 2-28, 2-32, 2-23, 2-25
[37] 2-5, 2-18
[38] 2-5, 2-11, 2-15, 2-28, 2-32
[39] The programme uses an observational rubric to measure 'interaction, facilitation, power dynamics and the use of research evidence' at its knowledge café and policy dialogue events.
[40] 1-6, 1-7
[41] 3-16, 3-4, 3-W
[42] 3-W2, 3-W, 3-7
[43] 3-7, 3-16
[44] 4-1, 4-9, 4-12, 4-14, 4-39

	mentoring and science policy cafes) had helped strengthen these links further.[45] However, the evaluation found little evidence that the science policy cafés and research–policy events had improved relationships between policy makers and researchers. There is no monitoring data relating to this outcome.
ACD	Three interview respondents reported that positive relationships were developed between the CPRU and line ministries, who were required to work together to complete the policy template.[46] There is no monitoring data relating to this outcome.
South Africa impact case	Three respondents felt DPME's support to the ECD Diagnostic Review had promoted increased inter-sectoral collaboration among ECD actors.[47]

CIMO 13: Evidence underpinning the outcome: establishment or strengthening of national institutional actors to promote EIPM, which act as a 'hub' for EIPM, are capable of running successful programmes to promote it and are potentially able to continue supporting EIPM once the programme has ended

VakaYiko	Programme documents show that a 'learning by doing' type capacity building approach was adopted. The ZeipNET and GINKS Year 3 Organisational Capacity Assessments report improvement in communication, M&E, project and financial management, pedagogical skills, networking and leadership. INASP has also supported ZeipNET to produce a sustainability plan. Interviews with programme staff provided examples of improvements in capacity of ZeipNET.[48]
Harvard	N/A – resource not provided to date in this programme.
UJ-BCURE	Three interview respondents reported that UJ BCURE had been successful in building a network around its programme, and had become well integrated with key role players in the evidence landscape.[49] One senior stakeholder felt UJ-BCURE presents a potential location for a dedicated centre for research synthesis (3-26).
ECORYS	N/A – resource not provided to date in this programme.
SECURE	Two stakeholders felt that AFIDEP had strengthened relationships and collaborations with sector stakeholders, as a result of the BCURE programme.[50]
ACD	N/A – resource not provided to date in this programme.

CIMO 14: Evidence underpinning the outcome: increased interest in and demand for partners to provide capacity building support for EIPM, including from actors not originally targeted by the programme, as a result of successful delivery of programme activities by local partners, and/or activities that explicitly aim to build relationships between partners and government departments

VakaYiko	Programme staff and civil society stakeholders provided several examples of ZeipNET receiving requests for support and collaboration from new organisations – e.g. invitations to attend and present at a number of national events, and requests from other ministries and non-governmental organisations.[51] Programme reporting supports this, documenting the requests received. However, interviews raised some concerns that resource constraints will pose a barrier to taking these opportunities forward, as ministries and external stakeholders lack the funds to pay for training.[52]
Harvard	N/A – resource not provided to date in this programme.
UJ-BCURE	Several stakeholders reported increased interest in and demand for BCURE activities.[53] Two stakeholders also emphasised the importance of BCURE's approach in working with the PSPPD and DPME, to strengthen relationships and find ways for the programmes to complement each other.[54] This interview testimony is supported by programme reporting, which documents the

[45] 4-1, 4-12
[46] 5-13, 5-32, 5-31
[47] 1-6, 1-7, 2-5, 2-11, 2-15, 2-28, 2-32, 3-7, 3-16, 3-W2, 4-W, 5-24, 7-1, 7-9, 7-12
[48] 2-1, 2-W, 2-W2
[49] 3-25, 3-18, 3-21
[50] 4-29, 4-37
[51] 2-W2, 2-21, 2-20
[52] 2-31, 2-W2
[53] 3-7, 3-20, 3-24, 3-W
[54] 3-1, 3-27

	additional requests for capacity support – for example the request from DPME to support the human settlements evidence map.
ECORYS	N/A – resource not provided to date in this programme.
SECURE	One programme staff member gave examples of growing interest in the programme from a new stakeholder not originally targeted by BCURE, manifested in elements of the programme being adopted into their work, and invitations to SECURE to present at a forum (4-37).
ACD	N/A – resource not provided to date in this programme.

9. Value for money analysis

VFM approach

A value for money (VFM) analysis of the BCURE programmes has been conducted at the level of individual programmes. This analysis is detailed in a commercially sensitive separate annex for DFID that is not publicly shared. Key overarching findings from the VFM annex are summarised here.

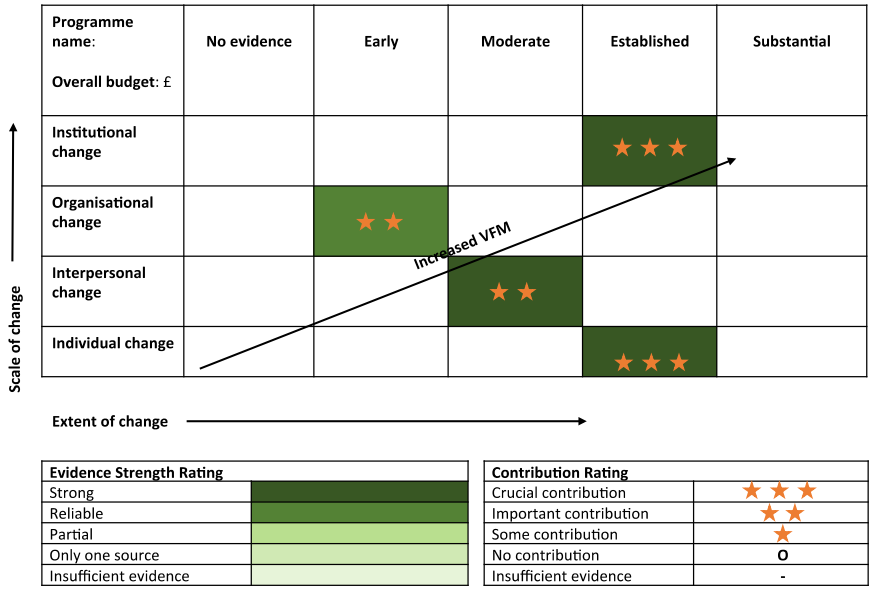
VFM within the BCURE programmes has been explored using the following approach:

- **VFM assumptions and framing:** Stage 2 synthesis evidence indicates that VFM in the BCURE programme arises from a combination of capacity development interventions implemented at different levels or scales of change (individual, interpersonal, organisational and institutional). These interventions need to work in parallel to influence changes in skills, behaviours and systems at all levels to enhance use of evidence and deliver VFM. For example, a project that has delivered on individual change but that has not influenced change at other levels will have lower VFM than a project that has delivered a degree of change at all levels.
- **The overall VFM of each project:** for each project, overall VFM was analysed using a matrix, as shown in Figure 6. As above, the VFM assumption underpinning the matrix is that the more change is stimulated at each of the levels (vertical axis), and the more established that change is (horizontal axis), the higher the VFM is of the BCURE project. The strength of the evidence is indicated by the shading intensity and the level of contribution of the BCURE project with the number of stars. A project that has delivered on individual change but that has not influenced change at other levels will have lower VFM than a project that has delivered a degree of change at all levels.
- **Effectiveness and efficiency** have been explored through evidence of specific examples of interventions, aiming to influence change at different levels, which could demonstrate the relationship between cost and value. Examples of effectiveness have been analysed using a combined framework that considers:
    - The outcomes at different levels identified by the evaluation team in the programme reports.
    - The judgements made on the extent of change influenced.
    - The strength of supporting evidence.
- **Economy** has been explored through looking at two measures: 1) the ratio between programme and administration costs; and 2) specific unit costs focusing on the main drivers of expenditure.

VFM judgements at Stage 2

In 2016, the BCURE evaluation has just completed Stage 2. This stage takes place when the programmes are moving beyond the mid-point phase of the programme. There is a mixed picture on VFM, with initial indications that some types of interventions may be providing good VFM, while other cases appear to be less clear-cut. However, outcomes are still emerging in many of the BCURE settings and so final judgements on VFM can be made only at Stage 3 of the evaluation, which will be summative. After three or more years of elapsed time since the projects started, the maturity of the outcomes observed, as long as they are supported by good evidence, should enable firmer conclusions regarding the VFM of the BCURE interventions.

Figure 6: Matrix of outcome level changes



Evidence Strength Rating	
Strong	
Reliable	
Partial	
Only one source	
Insufficient evidence	

Contribution Rating	
Crucial contribution	★ ★ ★
Important contribution	★ ★
Some contribution	★
No contribution	o
Insufficient evidence	-

Sources

For each project VFM analysis, the sources were the respective programme reports and financial data supplied by each team to the evaluation team on request. In some cases, the evaluation team has provided additional support to the teams to obtain financial data that fits the analytical framework.

Limitations to the VFM analysis

Limitations arise from two factors:

- The wide diversity of financial systems and contracting models in the portfolio means it has been challenging to disaggregate cost data against comparable indicators across the whole portfolio.
- Financial data has not been collected in ways that map to the scales of change in the evaluation framework (individual, interpersonal, organisational and institutional). Project financial reporting has been made against contract milestones, represented by 'bundles of activities' that cut across all the levels in the evaluation.
- Because of the challenges BCURE projects have had in extracting this data from their financial system, data in most cases has been collected only from the case study country. Although demonstrating valuable effects, the limitation of this approach is that it only accounts for only a proportion of the overall budget.

These limits have made it difficult to explore the relationship between cost, performance and value. In future, the unbundling of these activities according to the dimensions of change intended would enable the programme to demonstrate the shifts in investment over time as the programme evolved and to analyse overall cost-effectiveness at different levels.

Project summaries

Table 7 gives an overview of the programmes, overall budgets and timeframes, countries of operation and main interventions.

Table 7: BCURE programmes overview

Project name	Summary	Main interventions
INASP *VakaYiko Consortium*	The VakaYiko project is delivered by a consortium of organisations, led by INASP, and works in the following countries: Ghana, South Africa, Uganda and Zimbabwe. The budget value of the programme is £3,397,924 over three years (2013–16).	• Country level **capacity building** activities • Work to build and **document approaches** to supporting capacity to use research in developing countries • Formal and informal **organisational capacity strengthening** support to national organisations ZeipNET and GINKS

African Institute for Development Policy (AFIDEP) *Strengthening Capacity to Use Research Evidence in Health Policy (SECURE Health)*	The SECURE health programme is a three-year initiative (October 2013–October 2016) implemented by AFIDEP. It aims to increase the demand and use of research evidence in health sector decision making in Kenya and Malawi. It has a budget of £2,279,176.	• Strengthen institutional leadership and capacity for research evidence use • Enhance individual skills and capacity for policy makes in the health ministry and the legislature in accessing, apprising and using evidence
Harvard University *Towards a Culture of Evidence: Building Capacity for Evidence-Based Policy*	The Harvard programme aims to build capacity and culture of evidence informed decision making. It is a three-year programme that runs between 2013 and 2016. It spans four countries: Bangladesh, India, Nepal and Pakistan. It has a budget of £3,232,462.	• Assessment of needs and barriers to EIPM • Training on EIPM • Policy dialogues that foster collaboration and learning between policy makers and technical teams around specific policy issues • Pilot projects that demonstrate the practical systems and tools aimed at facilitating EIPM
University of Johannesburg *UJ-BCURE*	The UJ-BCURE programme strengthens capacity for evidence use in Malawi and South Africa through capacity building workshops, mentorships, and support of the AEN. It is a three-year project with a budget of £1,198,755.	• EIPM workshops • Mentoring and support to individuals and teams • Technical support within policy and decision making procedures to enhance the use of evidence • Convening and implementing the AEN
Adam Smith International (ASI) *African Cabinet Decision-making Programme (ACD)*	The ACD project aims to enable greater use of evidence in cabinet decision making in Africa, focusing on Sierra Leone, South Sudan and Liberia. It began in September 2013 and will run for 36 months plus a three-month no-cost extension. The total programme budget is £3,118,031. The project has four outputs.	• Cabinet processes to support cabinet secretariats' ability to oversee revised cabinet processes • Support to enhance the ability of ministers to interrogate the quality of proposals submitted to Cabinet • Support to enhance line ministries' development of evidence informed proposals • ACD programme guidance, advice and training materials are shared effectively with others, particularly African cabinet secretaries • Convening the Africa Cabinet Governance Network (ACGN)
ECORYS *BCURE Bangladesh*	The ECORYS programme focuses on Bangladesh and aims to increase the use of evidence in policy formulation, to contribute to better and more effective policies leading to greater poverty reduction and improved quality of life for citizens in Bangladesh. It runs from September 2015 to March 2017, with an overall budget of £1,628,000, approximately £1.4 million of which is for the implementation phase.	• Enhancing the *institutional framework* in support of EIPM • Strengthening of capacity to use EIPM through *formal training* • Strengthening of capacity to use EIPM through on-*the-job training* • Increased *awareness* of, demand for and understanding of the benefits of EIPM

Key VFM results

Overall, the evaluation has found evidence that BCURE activities are enhancing the use of evidence in the programme countries. However, there are some weaknesses, and also questions about how sustainable the positive results are, which may erode VFM in the future. For example, training and mentoring has been largely effective at building skills and influencing behaviour change across four programmes that have implemented training as a core activity. Organisational-level activities have also influenced positive change, for example official of adoption of EIPM procedures and guidelines. Other activities have had more limited results – for example the science policy dialogues three BCURE programmes have used as a core strategy.

At Stage 2, positive changes are still at an early stage and, as programmes are approaching completion of their three-year terms, this raises questions about sustainability. Three years is a relatively short time to have achieved sustainable reforms and behaviour changes in government settings. There is a risk that positive results will lose momentum after the BCURE projects have closed. A lack of sustainability would erode the VFM of projects, even if at the moment they are showing positive signs of delivering change. VFM could well be enhanced if BCURE interventions went on longer.

Synthesis findings at Stage 2 suggest that, to achieve VFM, capacity development for EIPM needs to go beyond building technical skills at an individual level, ideally encompassing a range of interventions at different levels to influence change. Capacity strengthening needs to be considered systemically, providing support at different levels. For example, investing only in training will yield limited value in terms of results, as individuals need a supportive organisational environment to put skills into practice, and individual-level changes are unlikely on their own to 'filter up' to create organisational change. Conversely, top-down reforms may create improved systems for planning and policy making, but may not necessarily change behaviours to use evidence unless individual change is also catalysed. Capacity support for EIPM should ideally encompass a range of interventions, from developing skills, values and norms to promote EIPM at an individual level, to supporting the adoption of organisational procedures, incentives and resources – financial and human – to create the hoped-for returns in terms of enhanced use of evidence.

VFM results: effectiveness and efficiency

We have provided summary overview of three types of interventions most BCURE programmes have implemented, in order to provide an illustrative comparison. It must be noted that these interventions are implemented in a wide diversity of countries and government contexts, which influences the costs. The three interventions are:

- EIPM training.
- Policy dialogues and science cafes.
- Development of EIPM guidelines and formal procedures.

Overview of EIPM training, coaching and mentoring

The training has been largely effective across four programmes that have implemented training as a core activity. Different contexts, target groups and delivery models mean there is some variation in costs. Delivery models include online and supported learning platforms reaching large cohorts across the civil service, modules released in blocks over time for a number of ministries and intensive residential courses lasting a few days for staff from within one ministry.

Training costs have included the development of curriculum and delivery platforms, as well as the training events. If the development costs are treated as sunk costs, then training costs in the programmes have ranged from £144 to £648 per person.

Value in training, coaching and mentoring interventions is generated not only through skills acquisition but also through people putting EIPM skills into practice, improving the quality of their work in policy and decision making processes and cascading EIPM to others. Positive results were observed in Kenya, South Africa and Zimbabwe on this, with more limited evidence suggesting application of skills has not been as good in Pakistan, possibly because of a lack of targeting to officials in policy development roles and the difficulties of introducing change in a large civil service. The evaluation also found examples in some countries where additional value has been generated through the creation of 'evidence champions', who then spontaneously go on to cascade EIPM training through onward training or the introduction of new ways of working that promote the use of evidence. Training and mentoring interventions can also provide an entry point for the partner to provide further support to the government organisation, which offers potential for more value to be generated (see the full Stage 2 Synthesis Report for more detail on outcomes).

Further value is likely to be gained from curricula and courses if these are incorporated into civil service training colleges, as is planned in a number of countries. In Bangladesh, Kenya and Pakistan, national colleges have agreed to incorporate the EIPM training into their programmes, such as in the Mid-Career Management Course in Pakistan.

Overview of policy dialogues and knowledge cafés

Science cafés and policy dialogues are a popular activity. However, evidence at Stage 2 suggests that, while these activities can influence positive changes, examples of this have been limited. Three programmes

have implemented policy dialogue and science cafés. Policy dialogues are formal meetings with the aim of exchanging and debating the evidence to foster dialogue and collaboration on a particular policy issue. Science cafés are similar but more informal, where participants from a range of organisations – government, non-government, research and practitioners – are brought together to review evidence on a policy issue, to facilitate collaboration and relationships. In one BCURE programme, knowledge cafés are open to the general public and focus on raising awareness about and demand for EIPM. There is some evidence to suggest that events that bring people together and create spaces for dialogue have multiple, interlinked outcomes – including raising individual awareness, catalysing policy processes that utilise evidence and strengthening the position and networks of BCURE partners to enable them to promote EIPM more widely. For example, in Pakistan, policy dialogue helped catalyse the Crime Mapping and Polio Vaccines projects. In Kenya, one event led to concrete policy actions being taken forward by the ministry. However, these positive results contrasted with examples where the events had not influenced tangible changes.

Costs for science café and policy dialogue events have varied across the BCURE programmes, again because of differences in contexts and the purpose. Most events have reached from 35 to 120 participants each, with costs ranging from £3,763 for a smaller event to £12,481 for a larger one, with only one programme showing significantly higher costs for a similar activity for which the reasons are not yet fully understood.

Based on the current evidence, it seems policy dialogues and knowledge cafés offer limited return on investment when conducted as stand-alone interventions: they should be treated as a *stepping stone* to EIPM, not as an end in themselves. The contrasting findings suggest that, to be considered a VFM investment, dialogues should be designed with a particular end in mind, or as part of a wider change strategy, with interventions before and follow-up after to catalyse tangible results. The dialogues should be used to generate debate and connections around an issue for which there is appetite or intent to take actions forward. However, when linked to other activities, they are a relatively low cost way of creating discussion and networks around EIPM that other, more focused and in-depth, activities such as training can build on.

Overview of development of EIPM guidelines and formal procedures

Three BCURE programmes have provided technical support to the development, adoption and implementation of EIPM guidelines and related formal government procedures, which the Stage 2 evidence suggests have generated good levels of return. In Sierra Leone, the BCURE programme has supported the development of new a Cabinet memo template and procedures for memo submission, which, together with technical support and monitoring of decisions, provides a new framework in which Cabinet operates and makes decisions. Stage 2 evidence suggests a positive trajectory for the roll-put of the system, with some improvement in the quality of memos being submitted to Cabinet, at least in the presentation and reference to evidence.

Costs of developing guidelines and procedures, including consulting and building ownership within government departments, have varied according to the scope of the intended roll-out. Costs range from £27,000 to £288,375. The wide range of costs aligns with the scale of the changes being introduced, for example introducing new procedures across the whole government system-wide (cabinet and all ministries) or within a single ministry.

The return on these investments is not yet clear, as it is too early to see results from implementation of guidelines and procedures. In Kenya, the EIPM policy guidelines have been signed off by the respective institutions and are considered by respondents to be a key tool to standardise EIPM at an organisational level. An additional result is that the guidelines have stimulated the production of formal standard procedures for policy development by the Kenyan MoH Policy and Planning Department and will be integrated into the formal quality procedures for ISO certification. In Bangladesh, EIPM guidelines have been developed but have not yet been signed off by Cabinet as official procedures.

VFM results: economy

Economy has been explored through looking at two measures: 1) the ratio between programme and administration costs; and 2) specific unit costs focusing on the main drivers of expenditure.

Across the portfolio, administrative costs are in a reasonable range of 17–25% of the overall budget. One programme has a significantly higher ratio, which may be because partner costs are included as administrative costs; this will be investigated further.

For all the programmes, the two main cost drivers are staff/consultancy costs and international travel. Costs in these categories reflect the nature of the support being given, which requires hands-on staff time and travel between programme

10. RAMESES standards for realist evaluation

Earlier this year, a set of reporting standards were developed for realist evaluations as part of the RAMESES II Project (Wong et al., 2016).[55] These standards aim to improve consistency, rigour and usability of realist evaluations. The table below sets out the standards, and indicates the relevant section of the BCURE evaluation report where each standard is addressed.

No.	Standard	Relevant section of report or annexes
1.	In the title, identify the document as a realist evaluation	See title
Summary / abstract		
2.	Journal articles will usually require an abstract, while reports and other forms of publication will usually benefit from a short summary. The abstract or summary should include brief details on: the policy, programme or initiative under evaluation; programme setting; purpose of the evaluation; evaluation question(s) and/or objective(s); evaluation strategy; data collection, documentation and analysis methods; key findings and conclusions. Where journals require it and the nature of the study is appropriate, brief details of respondents to the evaluation and recruitment and sampling processes may also be included. Sufficient detail should be provided to identify that a realist approach was used and that realist programme theory was developed and/or refined.	See Executive Summary
Introduction		
3. Rationale for evaluation	Explain the purpose of the evaluation and the implications for its focus and design	See Section 1.2. of the main report
4. Programme theory	Describe the initial programme theory (or theories) that underpin the programme, policy or initiative	Section 3.2 of the main report describes our approach to developing and refining theory. Annex 4 details the programme theory at the beginning of Stage 2.
5. Evaluation questions, objectives and focus	State the evaluation question(s) and specify the objectives for the evaluation. Describe whether and how the programme theory was used to define the scope and focus of the evaluation	See Section 3.3 of the main report. Further detail in Annex 3.3.
6. Ethical approval	State whether the realist evaluation required and has gained ethical approval from the relevant authorities, providing details as appropriate. If ethical approval was deemed unnecessary, explain why	See Annex 3.8.
Methods		
7. Rationale for using realist evaluation	Explain why a realist evaluation approach was chosen and (if relevant) adapted	See Section 3.1 of the main report. Further detail in Annex 3.1.
8. Environment surrounding the evaluation	Describe the environment in which the evaluation took place	See Section 2 of the main report.
9. Describe programme policy, initiative or product evaluated	Provide relevant details on the programme, policy or initiative evaluated	See Section 2 of the main report. Further detail in Annex 6.

[55] See http://www.ramesesproject.org/

No.	Standard	Relevant section of report or annexes
10. Describe and justify the evaluation design	A description and justification of the evaluation design (i.e. the account of what was planned, done and why) should be included, at least in summary form or as an appendix, in the document which presents the main findings. If this is not done, the omission should be justified and a reference or link to the evaluation design given. It may also be useful to publish or make freely available (e.g. online on a website) any original evaluation design document or protocol, where they exist	See Section 3 of the main report. Further detail in Annex 3.
11. Data collection methods	Describe and justify the data collection methods – which ones were used, why and how they fed into developing, supporting, refuting or refining programme theory. Provide details of the steps taken to enhance the trustworthiness of data collection and documentation	See Section 3.4 – 3.6 of the main report. Further detail in Annex 3.4 – 3.6.
12. Recruitment process and sampling strategy	Describe how respondents to the evaluation were recruited or engaged and how the sample contributed to the development, support, refutation or refinement of programme theory	See Section 3.4 of the main report. Full detail on the country case study sampling approach in Annex 3.4.3, and the impact case sampling approach in Annex 3.6.
13. Data analysis	Describe in detail how data were analysed. This section should include information on constructs that were identified, process of analysis, how the programme theory was further developed, supported, refuted and refined, and (where relevant) how analysis changed as the evaluation unfolded.	See Section 3.7 of the main report. Full detail on the data synthesis approach in Annex 3.7, and further detail on the country case study analysis approach in Section 3.4.5.
Results		
14. Details of participants	Report (if applicable) who took part in the evaluation, the details of the data they provided and how the data was used to develop, support, refute or refine programme theory	See Section 3.4 of the main report. Full detail on country case study participants in Annex 3.4.3, and impact case participants in Annex 3.6.
15. Main findings	Present the key findings, linking them to contexts, mechanisms and outcome configurations. Show how they were used to further develop, test or refine the programme theory.	See Section 4 of the main report. Annex 4 provides full detail on how the Stage 1 theories were refined at Stage 2 to lead to an updated programme theory.
Discussion		
16. Summary of findings	Summarise the main findings with attention to the evaluation questions, purpose of the evaluation, programme theory and intended audience	Summaries of the main findings are included throughout Section 4 in 'Overview' and 'Lessons' boxes.
17. Strengths, limitations and future directions	Discuss both the strengths of the evaluation and its limitations. These should include (but need not be limited to): (1) consideration of all the steps in the evaluation processes and (2) comment on the adequacy, trustworthiness and value of the explanatory insights which emerge. In many evaluations, there will be an expectation to provide guidance on future directions for the programme, policy or initiative, its implementation and/or design. The particular implications arising from the realist nature of the findings should be reflected in these discussions	See Section 3.8 of the main report.

No.	Standard	Relevant section of report or annexes
18. Comparison with existing literature	Where appropriate, compare and contrast the evaluation's findings with the existing literature on similar programmes policies or initiatives	This is done throughout Section 4 of the main report.
19. Conclusion and recommendations	List the main conclusions that are justified by the analyses of the data. If appropriate, offer recommendations consistent with a realist approach	See Section 5 of the main report
20. Funding and conflict of interest	State the funding source (if any) for the evaluation, the role played by the funder (if any) and any conflicts of interests of the evaluators.	See Section 1 of the main report. Further details on the evaluation team are contained in Annex 3.9.

11. References

Bandura, A., 1977. Self-Efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, *84*(2), 191–215.

Blamey, A., & Mackenzie, M., 2007. Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges? *Evaluation*, *13*(4), 439–455. http://doi.org/10.1177/1356389007082129

Broadbent, E., 2012. Politics of Research-Based Evidence in African Policy Debates: Synthesis of Case Study Findings. *EBPDN and MWANANCHI*, June.

Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M., 2015. What's in a Mechanism? Development of a Key Concept in Realist Evaluation. *Implementation Science: IS*, *10*(1), 49. http://doi.org/10.1186/s13012-015-0237-x

DFID, 2010. How To Note on Capacity Building in Research. Practice Paper. London: DFID.

du Toit, A., 2012. *Making Sense of 'Evidence': Notes on the Discursive Politics of Research and Pro-Poor Policy Making*. Working Paper 21. Cape Town: Institute for Poverty, Land and Agrarian Studies. Retrieved from http://www.plaas.org.za/plaas-publication/wp21dutoit

Faustino, J., & Booth, D., 2014. *Development Entrepreneurship: How Donors and Leaders can foster Institutional Change*. *Working Politically in Practice Series*. San Francisco, CA: The Asia Foundation. Retrieved from http://www.asiafoundation.org/resources/pdfs/DEODITAFWPIP2.pdf

IDRC, 2014. *Lasting Change – Strengthening Capacity to Improve Health Systems*. Ottawa: IDRC. Retrieved from https://www.idrc.ca/sites/default/files/sp/Documents EN/NEHSI-Capacity-Strengthening-INSIGHT.pdf

Jones, H., 2009. *Policy-Making as Discourse: A Review of Recent Knowledge-to-Policy Literature.* Joint IKM Emergent-ODI Working Paper.

Kirkpatrick Partners, n.d., The New World Kirkpatrick Model. Retrieved 8 August 2014 from http://www.kirkpatrickpartners.com/OurPhilosophy/TheNewWorldKirkpatrickModel/tabid/303/Default.aspx

Newman, K., Fisher, C., & Shaxson, L., 2012. Stimulating Demand for Research Evidence: What Role for Capacity-building? *IDS Bulletin*, *43*(5), 17–24. http://doi.org/10.1111/j.1759-5436.2012.00358.x

Noblit, G. & Hare, R., 1988. *Meta-Ethnography: Synthesizing Qualitative Studies*. London: Sage. Retrieved from http://www.uk.sagepub.com/booksProdDesc.nav?prodId=Book2416

Nutley, S. M., Davies, H., & Walter, I., 2002. *Evidence Based Policy and Practice: Cross Sector Lessons From the UK*. Working Paper 9. ESRC UK Centre for Evidence Based Policy and Practice.

Pawson, R., 2004. *Mentoring Relationships: An Explanatory Review*. ESRC UK Centre for Evidence Based Policy and Practice.

Pawson, R., & Tilley, N., 1997. *Realistic Evaluation*. London: Sage.

Pawson, R., & Tilley, N., 2004. *Realist Evaluation*. http://www.communitymatters.com.au/RE_chapter.pdf

Pope, C., Mays, N., & Popay, J. (2007). *Synthesising Qualitative And Quantitative Health Evidence: A Guide To Methods*. McGraw-Hill International. http://books.google.com/books?id=L3fbE6oio8kC&pgis=1

Punton, M., Brown, C., Echt, L., Fisher, C., Lloyd, R., Morton, S., & Vogel, I., 2016a. *How Can Capacity Development Promote Evidence-Informed Policy Making? Literature Review for the BCURE programme*. Brighton: Itad.

Punton, M., Vogel, I. & Lloyd, R., 2016b. Reflections from a Realist Evaluation in Progress: Scaling Ladders and Stitching Theory. *CDI Practice Paper*, 18, 1–11.

Sutcliffe, S., & Court, J., 2005. *Evidence-Based Policymaking: What Is It? How Does It Work? What Relevance for Developing Countries?* London: ODI.

Westhorp, G., 2014. *Realist Impact Evaluation: An Introduction.* Methods Lab. London: ODI. Retrieved from http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9138.pdf

Wong, G., Westhorp, G., Pawson, R. & Greenhalgh, T., 2013. *Realist Synthesis: Rameses Training Materials* (misc). London: Health Research Health Services and Delivery Research Program. Retrieved from http://www.ramesesproject.org/media/Realist_reviews_training_materials.pdf

Wong, G., Westhorp, G., Manzano, A., Greenhalgh, J., Jagosh, J. & Greenhalgh, T., 2016. RAMESES II Reporting Standards for Realist Evaluations. *BMC Medicine*, *14*(1), 96. http://doi.org/10.1186/s12916-016-0643-1

Yin, R. K. (2003). *Case Study Research: Design and Methods*. California: Sage.