Annexes for the Final Evaluation of the Building Capacity to Use Research Evidence (BCURE) programme

This document contains the annexes to the full report *"Final Evaluation of the Building Capacity to Use Research Evidence (BCURE) Programme,"* which is available here: http://www.itad.com/knowledge-and-resources/bcure/

Contents

1   Terms of reference and BCURE logframe

The original Terms of Reference and BCURE Logframe are presented below. Note that the project end-date moved to November 2017 as a result of various extensions, in order to allow sufficient time for projects to complete their activities.

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 policymakers 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 policymakers, 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 policymaking 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 12 were selected to develop into full proposals, including theories of change, work plans and logical frameworks. Of these 12 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 policymaking	Kenya and Malawi
C	Finalising contract	Develop online training on use of evidence aimed at policymakers	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 policymaking 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 policymaking organisations in developing countries use research evidence for decision making?

- What organisational structures, processes and systems help or inhibit the use of evidence by policymaking 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 policymaking 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 policymaking 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 policymaking 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 policymaking 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 policymaking 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 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.

---

[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.

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.

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 policymaking. 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 12 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 consortia 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)**	

			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)	Policymaking staff from selected countries have improved skills for and understanding of Evidence-Informed Policymaking (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 cafes 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 policymakers 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 policymaking 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 cafes 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 cafes, 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 policymaking 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	Five pilot mentoring relationships complete Two secondments complete Invitations to one key national-level meeting per quarter; membership of	To be agreed during inception phase	* 20 women and 20 men mentored * Other targets to be agreed during the inception phase

			meeting of National Evaluation Strategy (4/14); Invited to strategic review of PSPPD (5/14).	one strategic steering group			
		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, policymakers 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 cafes 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 cafes	* 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 cafes held (7 in Kenya and 5 in Malawi); at least 80% participants giving positive assessment of the policy science cafes	1) Enhanced evidence use in decision making will result in an increase in evidence-informed health policies 2) Increased capacity of mid-level policymakers 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 policymakers 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)	

		Baseline	Milestone 1	Milestone 2	Target (date)
Enhanced capacity of mid-level policymakers in MoH and Parliament in use of research evidence, as measured by: - Number of mid-level policymakers 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 programme	**Planned**	* 0 * 0 * 0	* 40 mid-level policymakers 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 policymakers 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 policymakers 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 policymaking 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 management and learning

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

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

Key audiences for 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:

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

An evaluation communications framework was developed to facilitate the contribution of the evaluation to the wider evidence base on EIPM, and a range of communication activities have been conducted and will continue following the publication of the final evaluation. Annex 10 provides further information.

---

[3] See https://bcureglobal.wordpress.com/

3   Evaluation design and methodology

This section presents the full evaluation design and methodology, expanding on the summarised version contained in the main report.

3.1   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 common theory of change (CToC). This proved to be unwieldy, and the framework was streamlined for Stage 2. It was decided to focus on five questions, built around four domains of capacity change (individual, interpersonal, organisational and institutional) within our programme theory, as follows:

Stage 2 Evaluation questions

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?
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?
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?
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?
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?

Several revisions to the evaluation design were undertaken at Stage 3 in response to comments from the evaluation Steering Committee and DFID independent quality review at Stage 2. One key piece of feedback was that framing the evaluation around the four levels of change potentially sacrificed clarity and the ability to understand and describe findings in a holistic, case-based way. There were also concerns that the scope of the evaluation was quite broad, involving exploration of a wide range of emerging outcomes in six BCURE case study countries. This produced broad evidence of outcomes but did not provide the depth of evidence to draw definitive conclusions. At Stage 3, it was therefore decided to focus on a smaller number of 'priority outcomes' rather than investigate all of the anticipated outcomes across the BCURE projects.

In line with this, the Steering Committee agreed to further revise the EQs. Rather than linking them to the different levels of change, they were linked to priority outcomes identified within the country programme theories.

Stage 3 Evaluation questions

EQ 1. To what extent have priority outcomes been realised and for whom, and how sustainable are they? Have the theorised changes happened? How far have these changes occurred across different sub-groups and organisations etc., reflecting on gender and equity issues? How sustainable are the changes?
EQ 2. How significant was BCURE's contribution to priority outcomes, alongside the contribution of non-BCURE factors? What is the evidence that BCURE contributed to causing the observed changes, and what is the evidence that non-BCURE factors contributed? What is the relative importance of BCURE and non-BCURE factors in explaining the observed changes?
EQ 3. How and why did BCURE contribute or fail to contribute to priority outcomes? Through which mechanisms, enabled by which features of the intervention and features of the (individual, interpersonal, organisational and institutional) context, did BCURE contribute to the observed changes?

To answer the three EQs, the Stage 3 evaluation gathered and analysed evidence from various sources against country-level theories of change, to first judge the extent to which an expected outcome had emerged (EQ 1), then establish the extent to which BCURE contributed to this outcome (EQ 2), and finally determine how, why, for whom, and in what circumstances the outcome had or had not happened (EQ 3). Figure 1 depicts the logical flow of the evaluation questions, which was used to structure the approach to data collection and analysis. As agreed with the evaluation Steering Committee, the evaluation questions were framed around case-specific priority outcomes and thus were answered at the level of the internal country case study reports. This overview report provides summary comparative judgements across the cases in relation to the EQs, but its purpose is not to answer the questions at a portfolio level.

Figure 1. Logical flow of the EQs

3.2   Approach to answering the evaluation questions

The BCURE interventions work in complex government contexts, with myriad contextual conditions influencing potential outcomes. These included diverse historical institutional trajectories; varied political and economic conditions, government systems and organisational cultures; and a wide range of participant characteristics (individuals' identities, gender and ethnicities). Quasi-experimental and counterfactual approaches are unsuited to evaluating this type of programme, as there is no possibility of establishing a control group or comparator (Stern et al., 2012). In addition, BCURE was likely to be just one of a number of factors influencing change in complex government systems, giving rise to the 'attribution problem' – the challenge of attributing a particular change to a particular programme when other factors are also contributing (Wimbush et al., 2012).

In order to address these challenges and answer the evaluation questions, the evaluation adopted a **realist evaluation** approach, drawing on elements of **contribution analysis** and taking a **political economy lens**.

Realist evaluation

A realist approach was selected because the primary aim of the evaluation was 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 future. DFID was interested in understanding not just *whether* BCURE worked but also *how and why* capacity building can contribute to increased use of evidence in policymaking in the very different contexts in which the programme is operating (EQ 3). 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 – see Box 1).

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' or CMO configurations (see Box 1) – the core analytical units of realist evaluation (Pawson and Tilley, 1997; Wong et al., 2013). The evaluation team decided to incorporate *features of the intervention* as an additional element to our CMO configurations, 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 training) when considering what might 'spark' a particular mechanism. This gives us the formulation C+I+M=O (CIMOs), used throughout this report.

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 2 provides an overview of the evaluation design.

Figure 2. Application of a realist approach in the BCURE evaluation

> **Box 1: Context, mechanism, outcome and programme theory**
>
> **Mechanisms** are the causal forces, powers, processes or interactions that generate change within an intervention – including the choices, reasoning and decisions people make as a result of the resources the programme provides. 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 and Tilley, 2004).
>
> **Outcomes** refer to intended and unintended short-, medium- and long-term changes resulting from an intervention.
>
> A **CMO configuration** is a theory or hypothesis about how a particular mechanism works in a specific context to lead to an outcome. They can usually be read as sentences – for example, 'Where training content is directly relevant to a person's day job (C), providing information about how evidence can improve policymaking can spark an "eye-opener" in which trainees recognise how evidence can add value (M), leading to increased use of evidence in their day-to-day work (O)'.
>
> 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). A realist programme theory is a variation on a ToC, which explicitly spells out the causal links between outcomes as CMO configurations. 'Assumptions' in a ToC are embedded as *theories to be tested* in the CMOs as contextual factors and/or conditions necessary for mechanisms to fire. Some ToC approaches also include 'risks to assumptions' – that is, factors that will prevent the assumptions from holding true. Again, realist programme theory integrates this into the CMO testing, by explaining the contextual or intervention factors that block mechanisms from operating.
>
> *Source: Pawson and Tilley, 1997; Westhorp, 2014; Punton et al., 2016b*

The first iteration of the BCURE theory drew on the evaluation team's existing knowledge and professional hunches about the nature of capacity building, and how capacity building can contribute to evidence use in policymaking. This was used to shape the research questions for the BCURE literature review, which identified additional theories in the wider literature about how capacity building can contribute to EIPM. These were used to develop our first iteration of CIMO configurations. Stages 1 and 2 of the evaluation then began to test and refine these CIMOs, contributing to a revised programme theory at each stage. At Stage 3, a prioritised set of theories have been tested and revised for a final time, and are presented in the report. Annex 4 contains a full explanation of how the BCURE theory has evolved over time, and lists the CIMOs tested at Stage 3.

Contribution analysis

In order to answer EQ 2, the Stage 3 evaluation drew on elements of contribution analysis. Contribution analysis is a theory-based evaluation approach that provides a systematic way to arrive at credible causal claims about a programme's contribution to change. It allows a robust assessment of cause and effect when it is not practical to design an experiment to measure the attribution of a particular change to a particular

programme (Mayne, 2012). The six steps of contribution analysis[4] provided a framework to help prioritise outcomes and causal links to investigate during Stage 3, and assess the contribution of the programme alongside the role of other factors, as follows:

▪ A country-level ToC was developed for each case study, allowing the underlying causal logic to be unpacked.

▪ Evidence from earlier stages of the evaluation was assembled, in order to assess the strength of the existing contribution story, and identify weaknesses and gaps.

▪ Priority outcomes and causal links to focus on at Stage 3 were then selected, based on a consideration of their importance to the overall contribution story, and utility and importance to stakeholders (Lemire et al 2012).

▪ Evidence about the extent of BCURE contribution was then collected through country case studies, including through incorporating questions about contribution in the interview topic guides, and examining other explanations for observed outcomes through the political economy lens.

▪ The country case study analysis then involved a systematic assessment on the extent of BCURE contribution against the country-level ToCs, described further below.

Political economy lens

The Stage 3 evaluation aimed to incorporate a stronger understanding of how political economy issues affect evidence use in policymaking, in order to unpack non-BCURE drivers of outcomes (EQ 2) and incorporate political economy dimensions into our explanations of why BCURE contributed or failed to contribute to outcomes – i.e. the 'C' in CIMOs (EQ 3). A light touch political economy analysis (PEA) exercise was conducted at both country level (to identify key overarching factors and trends that are shaping and influencing policymaking and evidence use) and sector level depending on the sectors targeted by the BCURE partner, as part of each country case study. This was guided by a framework incorporating a checklist of PEA questions, drawing from various pragmatic PEA tools (Hudson et al, 2016; Poole, 2011; Moncrieffe and Luttrell, (2005). The approach is described in more detail in Annex 3.4 below, and the framework is presented in Annex 5.

3.3   Evaluation components

The evaluation had four main components:

1. **Annual 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. At Stage 3, the evaluation refocused its resources to conduct four evaluations instead of six. This allowed the team to investigate a smaller number of priority outcomes in more depth.

2. **A realist literature review**, synthesising published papers and grey literature related to capacity building for EIPM.

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).

4. **A synthesis of findings**, drawing together insights on how and why capacity building for evidence use works or does not work in different contexts.

---

[4] These six steps are: setting out the cause-effect issue to be addressed; developing a theory of change; gathering existing evidence on the theory of change; assembling and assessing the contribution story and challenges to it; seeking out additional evidence; revising and strengthening the contribution story (Mayne, 2011).

Data collection and synthesis was 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 – see Figure 1 above. The four components are described in more detail below.

3.4   Component 1. Programme evaluations and country case studies

3.4.1   Approach

During Stage 1 and 2 of the evaluation, programme evaluation reports were produced for each of the six BCURE projects. At Stage 3 it was agreed with the Steering Committee that the evaluation would conduct four 'country case studies' instead of six programme evaluations, to enable a focus on 'depth' rather than 'breadth'. The reports performed two functions:

- **Providing internal management reports for each project**, which verified outcomes identified by the BCURE programme monitoring data (and identified additional outcomes), captured key lessons and recommendations and generated an assessment on programme effectiveness, value for money, sustainability and programme contribution to change, in order to inform decision making.

- **Collecting data on how and why BCURE projects contributed to different patterns of outcomes**. This data was then fed into the synthesis, in order to identify, test and refine theories about how and why BCURE interventions lead to, or do not lead to, change.

At Stage 3, the programme evaluations were reframed as 'country case studies', and focused primarily on the second function. Each programme evaluation / country case study consisted of an independent review of secondary monitoring data and implementation documents produced by the project team, and primary data collection by the evaluation team within one of the countries targeted by the project. Over the course of the evaluation, 15 programme reports have been produced (five programme evaluations at Stage 1, six at Stage 2, and four country case studies at Stage 3). These are all internal to DFID.

3.4.2   Selection of country case studies

BCURE worked across 12 countries. The evaluation was only able to cover six with available resources. The country case studies were selected during the inception phase 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*)
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*)

At Stage 3, it was decided to focus on four countries rather than six, in order to allow for a more in-depth investigation. Case studies were selected based on the following considerations:

▪ The feasibility of accessing data and stakeholders in the context, based on the potential receptiveness of partners given that most projects will have finished by the time the evaluation commences data collection, and also considering other issues that might affect feasibility such as elections.

▪ Aiming for a balance across African and Asian contexts.

▪ Aiming to provide insights from different delivery models, i.e. sectoral focus, entry point, number of ministries targeted etc.

▪ Focusing on DFID countries by spend.

Based on these factors, the following four countries were selected: Bangladesh, Pakistan, Kenya and Zimbabwe.

3.4.3 Methodology for Stage 3 country case studies[5]

The Stage 3 case studies were designed and conducted following six iterative steps:

Step 1. A country-level ToC was developed, drawing on the programme's own ToC, evaluation data from Stages 1 and 2 on outcomes and causal links (CIMOs), and insights from the wider literature.

Step 2. Existing evidence was assembled for each outcome and causal link (CIMO), and gaps and limitations were identified in the existing evidence base, including around political economy dimensions. Priority outcomes and causal theories were identified from this preparatory analysis. In some countries where interventions differed by sector, a sector-specific ToC was developed, which drew on the country-level ToC but reflected specific outcomes and causal pathways.

Step 3. A political economy analysis was conducted to contextualise the ToC within the risks and opportunities that the context posed for EIPM and the programme's desired outcomes.

---

[5] The methodology for the Stage 1 and 2 programme evaluations is detailed in the Stage 1 and 2 Synthesis Reports and annexes, available from http://www.itad.com/knowledge-and-resources/bcure/

Step 4. Based on the priority outcomes, a purposive sampling framework was then developed to gather secondary data and collect primary data. The sampling process was iterative, developed and revised throughout the data collection process. Data collection involved iteratively triangulating evidence of outcomes, as well as testing and modifying theories about BCURE's contribution to outcomes, the role of other factors, and how and why BCURE contributed or failed to contribute to priority outcomes.

Step 5. A small number of examples of potential policy-level influence were identified by interview respondents, and these were investigated in greater detail through one or two illustrative case studies per country case study.

Step 6. Primary and secondary data was then analysed against evaluation questions to establish the extent of: priority outcomes (EQ 1); BCURE's contribution (or non-contribution) relative to other factors (EQ 2); how and why BCURE contributed or failed to contribute (EQ 3); and examples of policy influence, in order to assemble a summative 'contribution story' for the country case study.

Step 1. Country-level theory of change

Based on the findings from Stage 2, a draft country ToC was developed in the design phase. This aimed to be as specific as possible about the outcomes anticipated by the programme, the critical political economy dynamics affecting the context, and the observed/theorised causal links, to provide a more concrete and contextualised framework for the country case study.

The ToC built on the BCURE programme's own ToC, evidence from Stages 1 and 2 of the evaluation and the broader literature, and consultation with programme staff. It was designed to enable prioritisation of outcomes and causal links for investigation at Stage 3, and systematic investigation of the evaluation questions. The ToC also incorporated our theories (CIMOs) from Stage 2, about how and why certain outcomes were expected to lead to other outcomes in particular contexts, based on evidence collected to date. The ToC was validated with programme teams prior to data collection.

Step 2. Assembling existing evidence for the country-level ToC, and identifying 'priority outcomes'

Drawing on advice about how to test programme theory and insights from contribution analysis, we assembled existing evidence for the country-level ToCs (Funnell and Rogers, 2011; Mayne, 2008; 2012). Data tables were designed to aid this process. Evidence from the Stage 1 and 2 evaluations was collated in relation to each outcome in the country-level ToC, against the three EQs. This allowed us to identify where there was already substantial evidence, and where there were weaknesses and gaps.

Contribution analysis provided a framework to help prioritise outcomes and causal links to investigate during Stage 3. Lemire et al (2012) suggest that prioritisation should be based on a consideration of:

- Fit with purpose of evaluation

- Importance to overall contribution story

- Utility and importance to stakeholders.

Based on these considerations and in consultation with DFID, we identified priority outcomes specific to each country's ToC. We prioritised longer-term outcomes crucial to the overall programme goal of improving the use of evidence in policymaking processes. We also collected evidence against shorter-term outcomes where there were significant evidence gaps that need to be addressed in order to strengthen the overall contribution story for the programme. Finally, we also collected evidence about the political economy dynamics that have shaped BCURE's contribution (or non-contribution).

Step 3. Conducting a political economy analysis

Political Economy Analysis (PEA) is concerned with the interests and incentives of different groups and how they generate policy outcomes; the role that formal institutions and informal norms play in shaping interaction; and the impact of values and ideas on political behaviour and public policy (DFID 2009). The

Stage 3 evaluation aimed to incorporate a stronger understanding of how political economy issues affect evidence use in policymaking. The political economy lens was linked to the revised EQs as follows:

- Helping to investigate *non-BCURE drivers of outcomes*, including the role of interests and incentives, formal institutions and informal norms, and values and ideas (EQ 2).

- Helping to incorporate political economy dimensions into our explanations of why BCURE contributed or failed to contribute to outcomes – i.e. the 'C' in CIMOs (EQ 3).

PEA was conducted at two levels:

1. Country-level, to identify key overarching factors and trends that are shaping and influencing policymaking and evidence use.

2. Sector level: guided by the country theory of change. We defined 1–2 sectors of interest within the country case study contexts, in collaboration with the programme teams and DFID.

A PEA framework was developed, incorporating a checklist of PEA questions drawing from various pragmatic PEA tools (Hudson et al, 2016; Poole, 2011; Moncrieffe and Luttrell, 2005). This was used to structure an initial review conducted by the national consultant prior to data collection, drawing on secondary data sources. Further information was collected through primary interviews with sectoral experts and government stakeholders during the main data collection stage.

Step 4. Developing a purposive sampling framework and collecting data

The priority outcomes guided our sampling and data collection for Stage 3. The aim was to achieve a sufficient degree of confidence about the extent to which priority outcomes had occurred (EQ 1), BCURE's contribution to the outcomes (EQ 2) and how and why BCURE contributed or failed to contribute (EQ 3).

Once priority outcomes were identified for each country case study, we began developing a purposive sampling framework. Our Stage 3 sampling followed four main principles:

1. **Sampling was driven by theory:** In line with our realist evaluation approach, sampling decisions were guided by our theory about the outcomes we expected to observe, and how and why these outcomes are expected to come about – in other words the country-level ToC and associated CIMOs.

2. **Sampling was iterative:** Following from this, sampling was iterative and sampling frameworks flexible, allowing for changes and additions during field work as theories developed and leads were followed. The sample therefore continually evolved throughout the data collection process.

3. **Sampling aided comparison between sub-groups:** A key element of our sampling strategy was comparison between different sub-groups of participants (e.g. more junior and more senior staff, and participants in different roles or ministries), in order to explain differential outcomes.

4. **Sampling sought to maximise triangulation of sources for each theory**: We aimed to triangulate evidence across a range of different stakeholders, through comparing insights from project participants with insights from knowledgeable 'outsiders' (informed by the PEA of who is influential in relation to the outcome), and through accessing secondary documentation where available. Our data sources are detailed below.

We identified stakeholders to interview in two ways:

- **Using previous samples, programme stakeholder lists, monitoring data, and staff recommendations.** During the case study design stage, an initial, incomplete list of interview respondents was identified through reviewing monitoring data and programme documentation (including complete participant lists where available), and conversations with programme staff. Interviews at Stage 3 built on the insights from Stage 1 and 2, and a substantial number of participants were consulted across two or more stages

to allow change to be tracked over time. Decisions about whether to interview the same participants again were based on the four main sampling principles above.

▪ **Using snowball sampling**. Once fieldwork was under way, interview respondents were asked to identify further individuals who could provide an insight into a particular outcome, or who were non-participants who could help to verify an outcome, for example, colleagues and unit managers. This strategy proved crucial in helping the team to identify knowledgeable non-participants in BCURE interventions, who were unknown to programme staff and therefore difficult to identify up-front.

Data sources

We drew on five types of data, with the aim of triangulating insights for each outcome from as many of the categories below as possible.

1. Monitoring data and other programme documentation, including pre- and post-training course test data, participant feedback on various programme activities, memoranda of understanding with government partners, activity reports, meeting minutes, and case studies written by BCURE partners. This was reviewed first to examine secondary evidence for theories, and to help identify relevant sub-groups of individuals to target for interviews.

2. Interviews and workshops with programme staff. This supplemented the monitoring data, helping understand gaps or areas where greater clarity was needed. They also aimed to explore the team's perceptions on observed changes with different groups, how and why the interventions contributed to change, and blockages to change. It also provided an insight into the areas project staff thought had been more or less successful and how and why, which helped to further develop our theories.

3. Interviews with project participants (individuals directly targeted by the activity / activities which aimed to contribute to the outcome). This generated self-reported insights about the extent to which outcomes had been achieved or not achieved, how and why, for different groups. We considered the possibility of collecting data from a larger number of project participants through a survey, but have rejected this as we felt it would be very difficult to get an acceptable response rate.[6]

4. Interviews with other knowledgeable stakeholders. These were stakeholders who did not participate in BCURE interventions, but who provided insights into (a) outcomes observed and the reasons for these outcomes; and (b) political economy issues that related to priority outcomes.

5. Secondary (non-project) documentation. This encompassed documentation not produced by the programme, which provided insights into outcomes and the reasons for outcomes: for example, government policy or guidance documents. We tried as much as possible to identify and secure potentially relevant documentation, (a) up-front when evidence was assembled; and (b) throughout the data collection phase, using interviews to attempt to secure documents that helped to triangulate insights from respondents.

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, identified through iterative snowball sampling).

▪ Between primary and secondary data sources.

---

[6] Several projects have already conducted surveys as part of their M&E, with medium-low response rates, and we felt we are unlikely to get anything better.

The total number of stakeholders consulted for the Stage 3 country case studies is summarised in Table 4. Full lists of respondents are included in the country case study reports. In total, 528 stakeholders were consulted across the six countries over three years.

Table 4: Total number of stakeholders interviewed at Stage 3

	Bangladesh	Kenya	Pakistan	Zimbabwe	Total
BCURE programme staff and direct implementing/consortium partners	6	5	8	6	25
BCURE programme participants	37	24	31	25	117
Non-participants, including government officials, civil society and other external actors	17	20	6	25	68
Total	**60**	**49**	**45**	**56**	**210**

Note: this table does not include interviews with BCURE programme managers

Interview tools

A set of semi-structured interview tools were developed (See Annex 7), designed to be customised to each individual stakeholder. Given the limited time available for interviews, it was necessary to prioritise which elements of the ToC and which CIMOs to test with different stakeholders, especially when respondents were involved in a range of different interventions, theorised to work in different ways.

In order to ensure that we explored outcomes, contribution and CIMOs systematically, we developed an 'Outcomes and CIMO question bank': a set of interview questions that covered the whole ToC. We decided, as part of the sample development and iteratively throughout the data collection, which outcomes and which CIMOs to discuss with which respondents based on relevance and consideration of data gaps. We designed unique interview guides for each respondent that aimed to test the most relevant theories, adding questions from the outcome and CIMO question bank into the generic interview templates.

The sampling spreadsheet was updated after each interview to keep track of which outcomes and which CIMOs had been discussed with which respondents, to ensure that we were testing the country ToC systematically. We used later interviews to corroborate and plug gaps in earlier ones. In addition to the CIMOs we were able to test explicitly, we were also able to infer information relating to our CIMOs from interviews during the analysis. The findings from both explicit and inferred testing were recorded as part of the analysis process.

The country case studies were scheduled to allow the interview guides to be tested by the team leader during the first case study fieldwork. The interview tools and question banks were iteratively revised over the course of the first case study, before subsequent case studies commenced.

Step 5. Embedded case studies on policy-level influence

The ultimate aim of BCURE was to influence the quality of policies in order to improve the lives of poor people. However, it was not feasible for the evaluation to systematically analyse evidence use in all decision-making processes in targeted BCURE organisations and ministries. First, BCURE programmes did not directly target specific policy processes but worked in a broader way to strengthen individual and organisational capacities and processes to enhance evidence use. Second, to focus on the decision-making processes taking place within a ministry would require considerable resources, disproportionate to the anticipated contribution of the programme.

Instead, we sought to (a) systematically catalogue examples of reported policy-level influence through the investigation of lower-level outcomes; and (b) conduct a small number of embedded case studies on a sub-
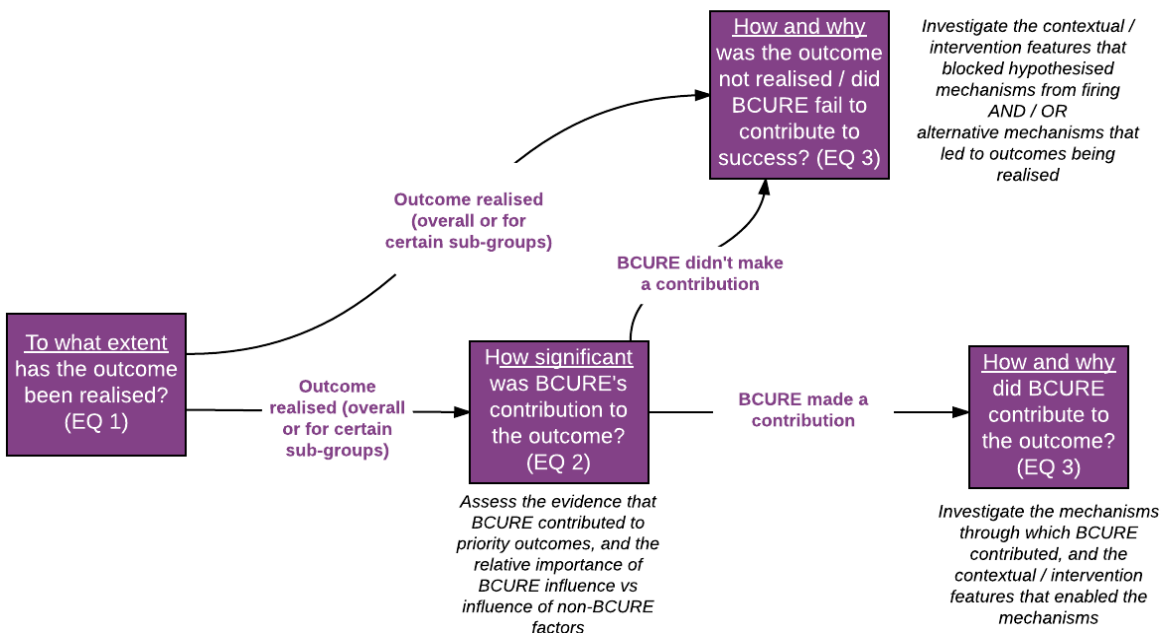
set of these examples. After an initial harvesting of examples of policy influence from respondents, we selected 1-2 examples per case study that were illustrative of an important pathway and appeared credible, triangulated them with supplementary interviews, and analysed them using the EQs.

Step 6. Analysis

Primary data from workshops and interviews was extracted evidence into a Microsoft Excel analysis spreadsheet, as follows:

1. Transcripts were reviewed for insights on the *outcomes* mentioned by respondents, in order to answer EQ 1. Each outcome was entered into a new row in the spreadsheet, in summary form, supported by a relevant quote from the transcript. Where a respondent had also been interviewed in Stage 2, programme leads reviewed the transcript from the previous year, to gain a sense of whether outcomes had been furthered or deepened.

2. Transcripts were the reviewed for insights on the *contribution* of the BCURE programme to the outcome, relative to other factors, including political economy issues, in order to answer EQ 2. This information was entered alongside the outcome in the same row, in summary form, with a supporting quote, as before.

3. For EQ 3, the transcript was then reviewed for the evidence arising from testing the CIMOs to explain 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 and contribution data, in the same row. 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.

The analysis followed the analytical logic laid out in the diagram below.

Secondary data: Documents were reviewed by the country case study 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:

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

Together, the primary and secondary data Excel sheets provided a catalogue of evidence enabling country case study leads to systematically and transparently assess the strength of evidence behind particular changes and identify how and why these changes were thought to have come about. Following discussions with the evaluation Steering Committee at Stage 2, the CIMO analysis has been embedded in the key findings sections in narrative form rather than presented as front-and-centre, in order to maximise the readability of the report.

3.4.4 Value for money analysis

A value for money (VfM) analysis was conducted as part of the Stage 3 country case studies, and integrated with the case study data collection and analysis. Given the summative stage of the evaluation, Stage 3 VfM analysis focused on cost-effectiveness – understanding the extent to which the investments made in the case study country had delivered value. At the country case study level, the focus was on assessing the cost-effectiveness of achieving priority outcomes. The cost-effectiveness of the overall investment made by BCURE partners in the case study countries will be made at a comparative level in the overview report.

Due to the nature of programme financial reporting it was not possible to identify the precise costs of programme activities. It was not a requirement for programmes to monitor the actual costs of activities and therefore the financial reports submitted to DFID did not provide an accurate picture. In many cases, the programmes had ended at the time the Stage 3 analysis was done, and so it was not possible for programme and financial staff to spend the time necessary to generate accurate data. As a result, the cost data was in most cases a rough estimation developed in consultation with BCURE programme staff.

Given the data limitations, it was not possible to conduct a robust quantitative VfM assessment, so qualitative judgements were made, were made through considering the following questions:

- Did the outcomes that were achieved justify the costs? Was the balance of investments across the priority outcomes appropriate?

- How institutionalised and/or sustainable were the reforms and outcomes observed?

3.5 Component 2. Literature review

A realist literature review (Punton et al., 2016a) was conducted during the early stages of the evaluation, in 2014–15.[7] 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 begin developing CIMO configurations. The findings informed the emerging theory and the development of the first iteration of CIMOs tested in Stage 1. A light touch literature review refresh was conducted in 2017 in order to generate additional insights on the Stage 2 programme theory, and the insights incorporated into the final comparative report.

3.6 Component 3. Impact case study

A non-BCURE impact case study was conducted, in order 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. This was designed to complement the BCURE programme evaluations through examining a non-BCURE capacity-building intervention that had been operating for a longer period of time,

---

[7] Punton et al., (2016a). Available from http://www.itad.com/knowledge-products/bcure-literature-review/

thus offering the potential to investigate how capacity building could contribute to changes in policy quality in the longer term.

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 policymaking 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 was: **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 looked specifically at two experiences with line ministries. The first is the updating of the government of South Africa's early childhood development policy following a DPME-facilitated diagnostic review in which the Department of Basic Education had a leading role. The second experience is the evaluation of the Department of Trade and Industry's Business Process Services programme and changes in the programme design arising from the evaluation.

There were 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 sought to explain how and why evidence produced through the evaluation/review of these policies/programmes was used in decision making. It also examined connections between evidence use and enhancement of policy processes in the two departments concerned.

The case study followed the same data collection and analysis methods as the Stage 2 programme evaluations, and took place during Stage 1 and 2 of the evaluation. It involved a 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. 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 the Stage 2 synthesis report and annexes. The final evaluation also drew on insights from an evaluation of the NES, due to be finalised in 2018. This was not yet published at the time of writing the final evaluation, but insights were drawn from presentations provided at the SAMEA conference in 2017.

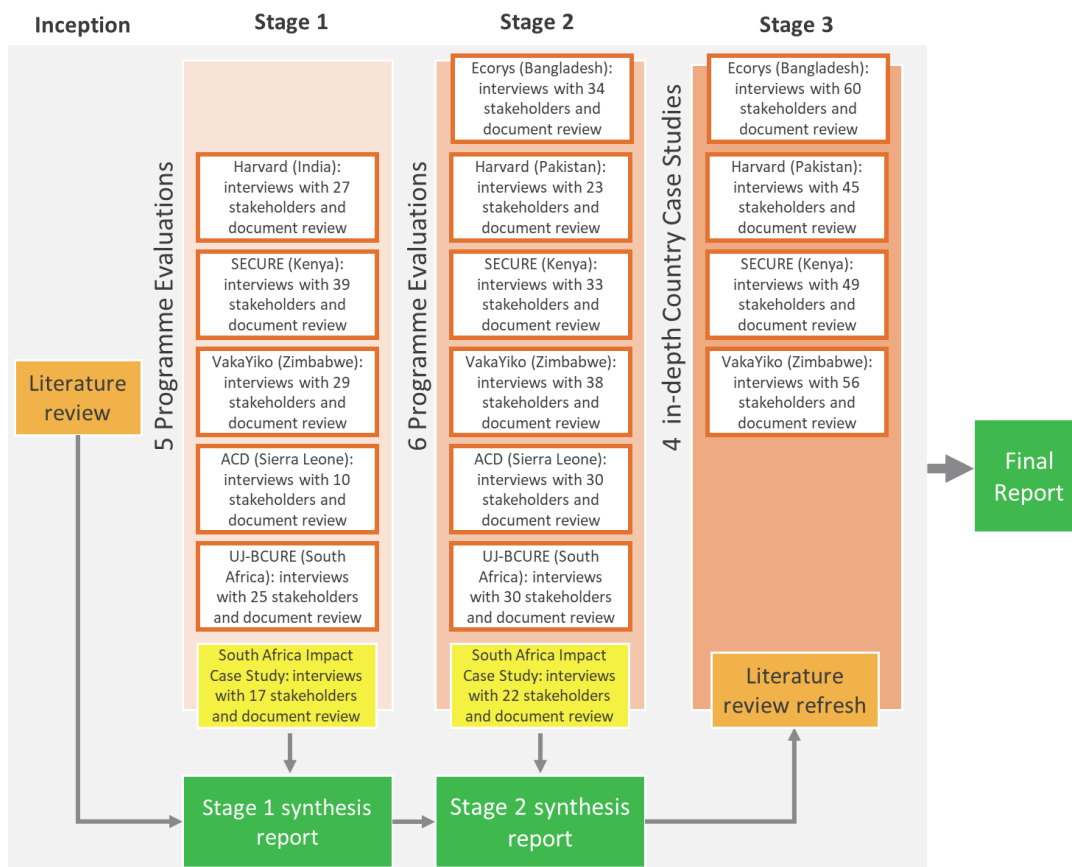**Table 5: Number of stakeholders consulted in impact case study**

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   Component 4. Overall synthesis

The overview report brings together the findings from the full three years of evaluation outputs: the Stage 1 and 2 programme evaluations and Stage 3 country case studies, the literature review and impact case study, and the Stage 1 and 2 synthesis reports. It aims to draw generalisable conclusions about how and why different BCURE interventions have contributed to different patterns of outcomes in different contexts, producing an evidence-based set of refined CIMOs and a refined programme theory.

Figure 3 presents a summary of the data from various evaluation components, illustrating how this fed into this final report. In total, 528 stakeholders were consulted across six countries and over three years.

Figure 3. Summary of data feeding into the final evaluation report

The synthesis process involved:

1. **Initial calls and workshops** with the country case study leads, to identify common concepts, themes or metaphors that applied across the cases, and interrogate differences. This enabled patterns to be identified and helped reveal nuances in the findings.

2. **Using a synthesis database** to combine relevant evidence from across the four Stage 3 case study reports about the outcomes achieved and not achieved (EQ 1), BCURE's contribution to these outcomes (EQ 2) and how and why particular outcomes were and were not achieved (EQ 3).

3. **Conducting a realist synthesis** across the cases, exploring how and why different BCURE interventions contributed to different patterns of outcomes in different contexts (EQ 3), in order to produce evidence-based set of refined CIMOs. As well as the Stage 3 case studies, this process also drew on the Stage 1 and 2 synthesis reports, the impact case study report and the literature review / literature refresh. This process applied realist synthesis techniques and additional insights from meta-ethnography in order to draw out meaning in a systematic way (see Box 2). 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.

4. **Checking and validating emerging conclusions**, through reviewing case study reports and where necessary interview data, to ensure that the evidence used to support, refute or refine the hypotheses

underlying the findings was relevant and sufficiently rigorous to support the inferences made (the 'translation' step in meta-ethnography, see Box 2). The two lead researchers cross-checked each other's analysis and conclusions, and shared drafts with other members of the core evaluation team to further validate and nuance findings.

During step 2, a few broad and cross-cutting patterns emerged that appeared to explain incidences of success across the BCURE portfolio. These patterns were discussed within the evaluation team, and then systematically analysed by developing matrices that drew together relevant insights from across the sources, and applying the synthesis techniques described in Box 2. This analysis suggested the importance of three broad 'ways of working' when seeking to build capacity for EIPM, described in Section 5.

Box 2: Qualitative synthesis techniques

Realist synthesis is an iterative process of theory building. It aims to generate the best possible explanation of evidence through retroductive analysis: moving between inductively building theories, and deductively testing them, while (in line with the realist philosophy) acknowledging that evidence and the resulting theories will always be partial and incomplete. Retroductive analysis applies a range of techniques to draw out insights from data, including: (Michaelis and Westhorp, 2016; Pawson, 2006)

- **Juxtaposing** insights from one case study to make sense of an outcome pattern noted in another.
- **Reconciling** contradictory insights through unearthing contextual or implementation differences and showing how these lead to opposing outcomes.
- **Adjudicating** between contradictory findings from different cases, to unearth strengths and weaknesses in the original conclusions that may explain these contradictions.
- **Consolidating** different results into multi-faceted explanations of success.
- **Situating** different results in their contexts – e.g. by exploring how one mechanism might operate in context A while another may operate in context B.

Meta-ethnography has much in common with realist synthesis. It is also an interpretive synthesis method, involving the transfer and translation of ideas, concepts and meanings across different sources (Noblit & Hare, 1988). Two of its steps were helpful as additional techniques for the synthesis:

- **Determining how evidence was related**: identifying points of comparison or opposition within the case studies, and identifying 'lines of argument' – inferences that cut across cases – through "*comparing and sorting interpretations, examining similarities and differences, and then integrating or framing these within a new interpretation*" that applied across cases (Pope et al., 2007).
- **Translation**: periodically revisiting case study reports and interview data to attempt to 'translate' evolving concepts or theories back into the source data, checking to see how far they accurately reflected case study findings, and scrutinising conceptual differences.

Throughout the overview report, insights on 'what worked, for whom, and why' have been drawn out. These represent 'empirical' CIMOs, which explain specific **outcomes** (O) from across the BCURE projects in terms of the **mechanisms** (M) that were (or were not) sparked by resources provided by BCURE, and the **context** (C) and **intervention** (I) factors that enabled or constrained the mechanisms. In the conclusions, these empirical formulations are brought up to the level of middle-range theory,[8] representing our final tested theory about what works to build capacity for EIPM, for whom, and in what circumstances.

3.8 Judging strength of evidence and extent of contribution

'Strength of evidence' relates to the internal validity of the evaluation findings. Our aim through the Stage 3 evaluation was to achieve a *sufficient degree of confidence* about the extent to which priority outcomes have occurred (EQ 1), BCURE's level of contribution to the outcomes (EQ 2) and our theories (CIMOs) about how

---

[8] This is theory that is *"detailed enough and 'close enough to the data' that testable hypotheses can be derived from it, but abstracted enough to apply to other situations as well"* (Wong et al., 2013).

and why BCURE contributed or failed to contribute (EQ 3).[9] Confidence in our conclusions about outcomes, contribution and CIMOs is underpinned by three broad considerations:

1. The extent of triangulation across stakeholders, participants/non-participants, and/or data sources. Within BCURE, triangulation has been pursued on several levels:

   o **Within interviews, by asking for examples.** If a stakeholder claims to have observed an outcome, confidence that this is true is increased if they are able to give specific examples.
   o **Across stakeholders and types of stakeholders**. Confidence that an outcome has occurred is stronger if more people, across different groups, claim to have observed it. Where possible, this has included seeking out and comparing insights from programme participants with non-participants, who have less of a stake in the programme being perceived as successful, and who, due to their position, have independent insights that provide corroboration and contextual information.
   o **Across data sources**: We have sought to triangulate insights from primary data collected through interviews with M&E data collected by the programme, and where possible with documents (e.g. policy documentation) produced by BCURE participants.

2. **A consideration of the position, knowledge, analytical capacity, reflexivity, and potential biases of primary informants.** In line with our realist approach, sampling decisions were purposively and iteratively guided by our (existing and emerging) *theory* about the outcomes we expected to observe, and how and why these outcomes were expected to come about. Stakeholders are therefore not considered in terms of homogenous categories (participants / non-participants), but as individuals positioned in unique ways in relation to the programme, with different levels of knowledge, capacity and reflexivity, and different incentives that may lead to bias. Weighing the strength of evidence requires a consideration of these issues, rather than simply considering the number of respondents who confirmed a particular outcome or CIMO. For example:

   o Different people can be expected to know different things about an expected outcome or change process. In some cases, only a small number of people are likely to know about an outcome, BCURE's contribution, and how / why it happened. Weighing the strength of evidence requires the evaluators to judge whether those who can be expected to know about the issue have confirmed that things happened in a certain way.
   o Different respondents have different levels of capacity (and interest) in scrutinising how and why something happened – particularly when this requires them to consider why they themselves have (or have not) changed their attitudes or behaviours – and this affects the weight that should be given to their responses.
   o Different stakeholders will have different incentives which may lead to biased responses; most obviously an incentive to 'tell the evaluator what they want to hear' in order to paint the project in a positive light and potentially secure future funding, leading to confirmation bias.
   o The position of a respondent in relation to the programme gives them a particular perspective which needs to be considered, overlapping with all of the above considerations. An external sectoral stakeholder may be able to provide important independent insights about broader political economy issues, but may not know much about the specific individuals or teams who took part in the programme (and therefore their opinions should be weighed accordingly). A senior civil servant may have good insights into outcomes but may be unwilling to speak openly about the realities of incentives and power structures within their ministry, and although they may not have participated directly in the programme they still stake in its success which implies the need to mitigate possible bias.

---

[9] This draws on thinking from process tracing and contribution analysis. Process tracing in particular offers useful insights into how to qualitatively weigh evidence in order to 'increase our confidence' that an intervention had an impact in a particular way. See: Befani, B. and Mayne, J. (2014). Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation. *IDS Bulletin*, *45*(6), 17–36. http://doi.org/10.1111/1759-5436.12110

The evaluators considered these issues both during the sampling process (when making decisions about who to interview), and during the interview write up and analysis (taking note of issues on the analysis spreadsheets in order to feed these considerations into the write up).

3. **A consideration of the broader context.** At Stage 3, the evaluation took a more explicit look at the broader political economy factors that enable and constrain EIPM in the countries and sectors under examination, and which provide opportunities and risks to the programme. This has provided more detailed insights into the contextual dynamics of BCURE country programmes, helping ensure that explanations of change are grounded in an understanding of the political context and are not over-reliant on the explanations of programme participants. This also helped identify other (non-BCURE) explanations of change, in order to help guard against over-attributing change to BCURE.

These three considerations were used to develop a qualitative approach to assessing the strength of evidence, described in the table below. This is not a rigid framework, but a way to ensure the evaluative judgements were made systematically and are comparable across the four case study reports.

Strength of evidence	EQ 1	EQ 2	EQ 3
Strong evidence	High level of confidence that the outcome occurred…	High level of confidence that BCURE contributed to the outcome…	High level of confidence that the outcome occurred / did not occur as a result of *x* mechanism, operating in *y* context and as a result of *z* features of the intervention…
	• …Based on a good degree of triangulation a) within interviews, b) across stakeholders and types of stakeholders, and/or c) across data sources • …Taking into account the position, knowledge, analytical capacity, reflexivity, and potential biases of primary informants • …and also taking into account what we know about the broader context through the PEA insights		
Some evidence	More confident than not that the outcome occurred…	More confident than not that BCURE contributed to the outcome…	More confident than not that the outcome occurred / did not occur as a result of *x* mechanism, operating in *y* context and as a result of *z* features of the intervention…
	…But confidence is reduced by: • Shortcomings with regards to triangulation, and/or • Concerns that the position, knowledge, analytical capacity, reflexivity, and potential biases of primary informants lowers the reliability of evidence, and/or • What we know about what is happening within the broader context		
Limited evidence	Low level of confidence that the outcome occurred, given that…	Low level of confidence that BCURE contributed to the outcome, given that…	Low level of confidence that the outcome occurred / did not occur as a result of *x* mechanism, operating in *y* context and as a result of *z* features of the intervention, given that…
	• …Evidence comes from a small number of sources with limited triangulation, and/or • …there are major concerns that the position, knowledge, analytical capacity, reflexivity, and potential biases of primary informants lowers the reliability of evidence, and / or • … there are contradictory insights into what is happening within the broader context		

Judging extent of contribution

In relation to EQ 2, 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	Details
Crucial contribution	Evidence that programme made a crucial contribution to observed change; i.e. change would not have happened without the programme. OR observed change is directly attributable to the programme
Important contribution	Evidence that programme made an important contribution alongside other factors
Some contribution	Evidence that programme made some contribution alongside other factors, but was not the most important cause
No contribution	Evidence of no contribution, or no improvement evident
Insufficient evidence	Insufficient evidence to make an assessment

3.9   Stakeholder engagement throughout the evaluation

The BCURE evaluation has been designed and implemented in close collaboration with the DFID evaluation Steering Committee, through regular meetings and calls, as well as numerous internal approach papers which offered an opportunity for DFID to review and comment on emerging design choices and suggested report structures. This regular engagement has facilitated annual revisions to the design in order to ensure the evaluation is meeting DFID's needs, particularly at Stage 3 where a substantial redesign was conducted (described above). The Steering Committee was also consulted on the selection of priority outcomes and CIMOs to test at Stage 3, based on the issues and questions most relevant to the design of future programmes.

BCURE partners have also been engaged at various points throughout the evaluation. Annual BCURE learning events offered an opportunity for the evaluation team to share emerging findings and interim analysis, with comments from partners fed into synthesis reports. In-country workshops with project partners provided an opportunity to hear the views of implementation teams and test CIMOs against their understanding of how and why change was (or was not) happening. At Stage 3, draft country ToCs were also shared and discussed with BCURE partners, and revised accordingly. Where possible during country visits, the evaluation leads also conducted debrief interviews or workshops with project staff, to share emerging findings at the end of the fieldwork, answer partner questions, and sense-check interpretations. Finally, draft programme evaluation and country case study reports were shared with partners to provide an opportunity for comments before the reports were finalised. These reports are internal, in order to protect the confidentiality of respondents and the relationships of BCURE partners with government stakeholders. However, synthesis reports and other publicly available evaluation products have been shared with interviewed stakeholders.

3.10   Ethics

The key ethical issue faced in the evaluation was protecting and managing the confidentiality of government documentation and stakeholder views at the local level. A number of the BCURE partners were operating at a high level within government and as such had access to government policy processes as they unfold. Access to these processes and the actors involved was 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.

The evaluation team sought to collect data in an appropriate and respectful manner, taking into account cultural and ethical concerns. Where possible, potential interview respondents were contacted several weeks in advance to enable the evaluation to fit into busy government schedules. Access to government institutions was facilitated by BCURE partners and national consultants, who had local knowledge about the protocols and etiquette involved in accessing stakeholders at varying levels of seniority, and who briefed international consultants on this in advance. Field trips were scheduled to allow sufficient time for researchers to be able to change their plans in order to fit in with the fast-changing schedules and commitments of government stakeholders. Researchers were also respectful of participants' time, and frequently cut interviews short or changed venues to enable stakeholders to participate. Researchers were also mindful of questions that might be inappropriate in particular contexts.

We ensured informed consent was 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 Annex 7). Unique anonymous interview codes were attached to each transcript and referenced in the country case studies. Where the content of quotes had the potential to identify an individual, this information was removed.

The BCURE country case studies are confidential reports viewed only by DFID and by the programme teams. The overview 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.

3.11 Evaluation team

The evaluation was undertaken by a core team from Itad, in partnership with Stellenbosch University.

Core team members

Isabel Vogel (Itad Associate)	Team Leader and lead evaluator for Kenya case study
Rob Lloyd (Itad Associate Director)	Project Director
Melanie Punton (Itad Senior Consultant)	Lead evaluator for Zimbabwe case study, support to Bangladesh case study, methods advisor
Gregory Gleed (Itad Consultant)	Lead evaluator for Pakistan case study
Joe Bolger (independent consultant)	Lead evaluator for Bangladesh case study, support to Pakistan case study, lead researcher on impact case study
Teresa Hanley (independent consultant)	Lead evaluator for Sierra Leone case study (Stage 1 and 2), support to Kenya case study
Babette Rabie (Stellenbosch University)	Lead evaluator for South Africa case study (Stage 1 and 2), support to Zimbabwe and impact case studies
Fanie Cloete (Stellenbosch University)	Lead evaluator for India and Zimbabwe case studies (Stage 1), advisory support

Country case studies were supported by national consultants Maheen Sultan (Bangladesh), Alfred Odour and Susan Mathai (Kenya), Rafiq Jaffer (Pakistan), Munhamo Chisvo (Zimbabwe), Andrew Lavalei (Sierra Leone), Benita Williams (South Africa) and Milindo Chakrabarti (India). Research assistance was provided by Alexina Jackson, Greg Smith, Verdiana Biagioni, Louise Horton and Katharine Hagerman. Communications support was provided by Clare Gorman and Emmeline

3.12   Limitations to the synthesis

The evaluation team was able to work freely and without interference, and there are no conflicts of interest to report.

The Stage 3 evaluation attempted to address a number of limitations identified at Stage 2. While this was successful to a large extent, certain issues proved difficult to address, giving rise to important lessons for future realist evaluations and evaluations of EIPM capacity development programmes.

The realist evaluation approach has been challenging to implement across the complex BCURE programme. In particular, we have faced challenges with 'breadth vs depth', including how to systematically prioritise outcomes and theories to assess within the limited time available for interviews. The Stage 1 and 2 evaluations generated a large number of theories (CIMOs) about how and why BCURE might be contributing to change at individual, interpersonal, organisational, institutional and policy levels. At Stage 2, it became clear that it was not possible to systematically test theories across the whole BCURE theory of change with the resources available for the evaluation. This was mitigated at Stage 3 through conducting a smaller number of more in-depth case studies, and prioritising a smaller number of outcomes and CIMOs for investigation. Developing country-level ToCs rather than relying on an overarching ToC helped identify case-specific outcomes and CIMOs that were less well-evidenced through earlier stages of the evaluation, and which were most important for achieving longer-term outcomes. This approach proved largely successful, and highlights the importance of realist evaluations prioritising the most interesting and important causal links in enough depth to draw useful insights, rather than trying to investigate everything. Case-specific theories, rather than (or as well as) a single overarching theory, can help facilitate this, through building an in-depth understanding of how and why a programme is expected to unfold in a specific case.

It has also been challenging to encompass complexity within the CIMO framework, including features of the macro political context and how they give rise to or inhibit mechanisms of change. There is a risk that CIMOs become overly linear explanations of how and why change happens ('this intervention feature, in this context, sparks that mechanism to lead to this outcome'). This was mitigated by presenting the final CIMOs in a more narrative way, which allowed the nuances and interconnections to be unpacked. The Stage 3 CIMOs also contain multiple features of context and multiple mechanisms, illustrating how these work together to lead to outcomes.[10]

In Stages 1 and 2 of the evaluation, it proved much easier to identify 'micro' features of context (e.g. around the characteristics of trainees) than 'macro' features (e.g. around the nature of government systems, the influence of power, politics and high-level incentives). This was mitigated through incorporating a specific PEA step in the Stage 3 methodology, which significantly enhanced the richness of the analysis. However, while the interviews provided a wealth of insights into the risks and opportunities that the context posed for EIPM and the programme's desired outcomes, it was not possible to gain insights into certain important issues likely to affect evidence use, including actual distribution of power and decision making, and some of the individual and organisational incentives that affect evidence use by senior government stakeholders. This suggests the importance of building in an explicit PEA lens from the outset in future realist evaluations working in government contexts, and considering how the evaluation design and team can be structured to best gain access to information on power distribution and incentives.

The evaluation has by necessity relied on interview data for evidence of outcomes, and there is a real risk of positive (confirmation) bias of respondents. With some exceptions, BCURE project monitoring systems were not set up to capture evidence of outcome-level change (including behaviour change and changes in policy processes or content). This has proved a major challenge for the evaluation, suggesting the importance of ensuring future programmes build monitoring systems that go beyond measuring outputs such as self-reported increases in knowledge and skills. The evaluation explored the possibility of conducting

---

[10] This follows the example of Leavy, Boydell and Mcdowell (2017).

large-scale surveys to capture insights from a broad cohort of participants, but given the high risk of low response rates this was not pursued. As a result, the evaluation has relied primarily on qualitative interviews with a select number of participants in order to provide evidence of longer-term outcomes. This carries a risk of confirmation bias, given the power dynamics of interviewing government stakeholders in low and middle-income countries. There is also the risk that participants may genuinely feel that the programme contributed to a positive outcome, when in fact other factors were more important – and this risk is heightened due to the complexity of the interventions, which makes it challenging to unpack contribution. We attempted to mitigate this at Stage 3 as follows:

1. **Triangulation:** We aimed for no more than 60% of the sample to consist of project participants and programme staff, with the remainder consisting of knowledgeable non-BCURE participants. Increasing the number of non-participant interviews helped to triangulate insights from project participants with the perspectives of individuals with less stake in the programme and potentially less incentive to tell the evaluators what they felt we wanted to hear.

2. **Conducting a more in-depth investigation into priority outcomes, and identifying and testing non-BCURE influences of change.** Focusing on a small number of priority outcomes enabled us to interrogate stakeholder testimony in more depth, helping us gain more detailed insights into what had happened and what had enabled or inhibited change. Our PEA exercise provided insights into country and sectoral contextual dynamics, helping ensure that explanations of change were grounded in an understanding of the political context, were not over-reliant on the explanations of programme participants, and were fair to programmes working in challenging settings. This helped to guard against over-attributing change to BCURE, as well as contextualising shortfalls in programme achievements.

3. **Dedicating more resources to finding monitoring and other documentary sources in order to triangulate interview data.** This included policy documents or government guidance that would help us validate stakeholder testimony about improvements in evidence access, appraisal and use. While this had some success and most case studies were able to view at least some documentation, there were ongoing challenges in accessing this data as the majority of stakeholders were unable to share internal government documents. This challenge was somewhat mitigated through interviewing a wide range of participants, and where possible their colleagues and managers, to triangulate insights.

It proved very challenging to secure interviews with government officials across all four settings – both participants and non-participants – particularly in Bangladesh and Pakistan. Challenges in securing and conducting interviews was a result of high workloads, adverse weather, security concerns (in Pakistan), and also the fact that most BCURE projects had largely finished activities in both settings, providing limited incentive for participants to volunteer their time. These challenges were mitigated through dedicating substantial efforts to contacting and following up with respondents, and through extending the length of country visits; however, in a number of cases the interviews were very short and it was only possible to explore a small number of outcomes and theories. This is reflected in the depth of analysis and strength of conclusions drawn in the country case study reports, and subsequently this overview report.

It has been challenging to ensure consistency of data collection and analysis across a diverse programme team. Time and budget constraints meant it was challenging to train the team comprehensively on the evolving programme theory, the principles of conducting realist interviews, and the approach to analysing data in a realist way. This created issues with ensuring a consistent approach to testing CIMOs and analysing interview data across the cases. At all three evaluation stages, we have attempted to mitigate this through a team workshop prior to data collection, involving a full introduction to the programme theory and basic training on realist interviewing and analysis. Programme leads provided training in-country to national evaluators prior to data collection, and additional analysis was conducted at synthesis stage by the team leader and methodological lead to capture insights that may have been missed during the case study analysis. At Stage 3, we also revised the team structure so that country visits were conducted by two core team members rather than one, which helped improve consistency across the cases. However, our major reflection is that realist evaluations require a different approach to team structuring and capacity building. Realist interviews and analysis require team members to have an in-depth understanding of not only the methodological approach, but the theory that the evaluation is trying to test. In order to ensure consistency

and understanding, a realist evaluation requires a more participatory model, which involves in-depth and ongoing engagement and capacity building.

Finally, the evaluation draws on evidence from only six of the 12 BCURE countries, and the short time frame of the programme limits the potential to record longer-term results. The evaluation is limited in what it can say about how BCURE worked across all of its settings, because its focus on six country case studies means it has not captured the full range of outcomes across the whole portfolio. Given the country-level focus, it also does not capture outcomes from the international and regional networking components that were part of various projects.

Finally, the programme was relatively short given its aim to generate systemic change in government settings – the shortest programme, in Bangladesh, had only two years of implementation time. This has limited the ability of the evaluation to identify longer-term results.

4   Programme theory and CIMO refinement

This section details the evolution of our programme theory and CIMOs from Stage 1 to Stage 3, documents the changes made and the rationale behind these changes, and presents our refined programme theory at the end of the evaluation. The Stage 2 Synthesis Annex contains further information about the evolution of theory from Stage 1 to Stage 2.[11]

Stage 1 BCURE ToC: Unpacking 'capacity development' to create a unifying framework

The six BCURE projects were highly diverse, taking different approaches to enhancing skills and systems for evidence use, in complex government contexts. As BCURE did not have a portfolio-level ToC, the evaluation developed an initial common theory of change (CToC) through synthesising the original project ToCs with key insights from the literature. The Stage 1 CToC unpacked capacity development into four levels of capacity change (See below),[12] which helped to bring the diverse BCURE approaches into a unifying framework for the evaluation. The four levels conveyed the concept of capacity development as multidimensional, and capacity as a function of different factors and processes working together and reinforcing each other at different levels. The BCURE ToC at Stage 1 stated that multidimensional change across these four domains would contribute to routine change in the use of evidence by government, in turn contributing to improved *quality of policy development processes,* as the overall impact.

Four levels of change

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: the systems, policies and procedures, practices, culture or norms within a governmental organisation that exist above the level of individual actors, and which incentivise, support (or inhibit) evidence access, appraisal and application in decision making. This includes 'system-level' factors within government that affect EIPM, such as national or sub-national laws, policies, regulations,

---

[11]   Available   here: http://www.itad.com/reports/building-capacity-use-research-evaluation-bcure-realist-evaluation-stage-2-synthesis-report/

[12] There are many definitions used in the literature to describe levels of capacity change. We 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 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 that '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.
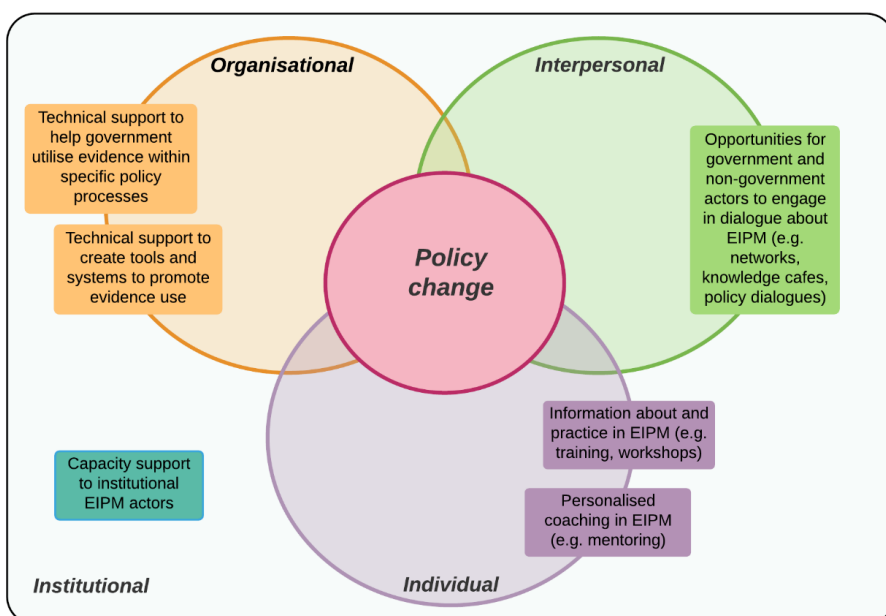
governance systems and 'institutional rules of the game'. Our definition of 'government' includes government administration and parliamentary scrutiny functions (including elected opposition politicians).

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 and 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.

Stage 2 ToC: Unpacking 'evidence use' and EIPM as a system working on multiple levels

The Stage 2 evaluation confirmed our theory that changes emerging at different levels (e.g. individual skills and organisational systems) seemed to reinforce each other, and that changes at different levels were required to make progress towards improvements in the quality of policy products and processes. The BCURE interventions worked through different 'entry points' at different levels. Some projects initially targeted *individuals* with information and opportunities to practise skills, others provided spaces for dialogue between different *groups* of stakeholders, others delivered technical support to *organisational* systems and processes and/or develop the capacity of *institutional* actors to promote EIPM. This led us to formulate the programme theory emerging from Stage 2 findings as a set of interlocking domains, with entry points at individual, interpersonal, organisational and institutional levels.

Figure 3: Different entry points of the BCURE interventions

The Stage 2 programme theory is described below, and represented in diagrammatic form in Figure 6.

> **Stage 2 programme theory and CIMOs**
>
> When the programme 'entry point' is through interventions targeting *individuals*...
>
> - 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 attempting to improve *interpersonal* links or relationships…

- Facilitated spaces for dialogue (e.g. between policymakers, 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).
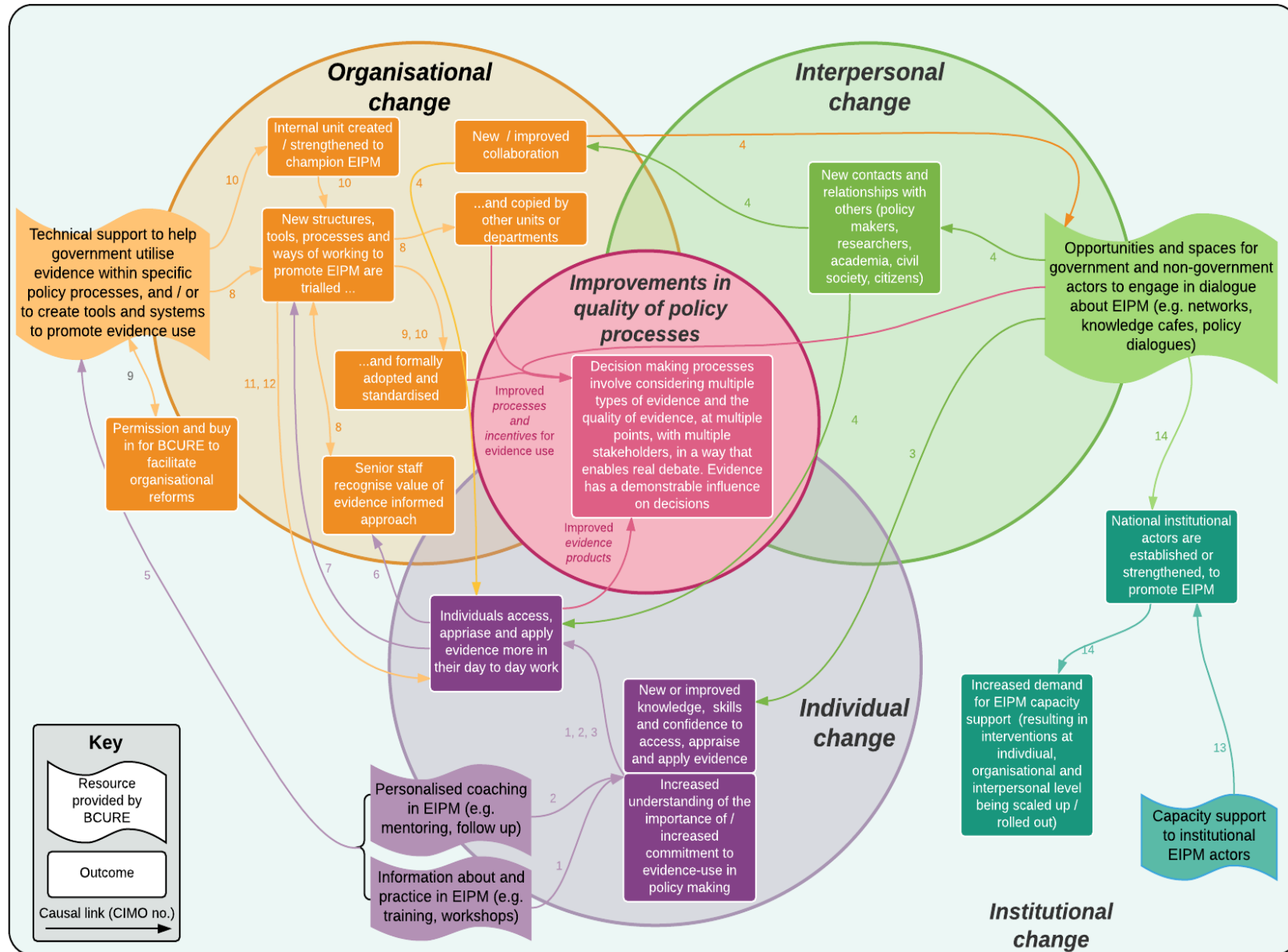
When the 'entry point' is through interventions at *institutional* level…

- 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.

Figure 4.   Diagram representing BCURE programme theory at the end of Stage 2

Stage 3 design phase: Unpacking 'evidence use' and EIPM as a system, and defining 'policy quality'

At the end of Stage 2, it became clear that the evaluation needed to gain a deeper understanding of key political economy dynamics in order to explain BCURE's emerging outcomes. In order to engage with this, at Stage 3 we developed country-specific ToCs, drawing on the Stage 2 CToC and insights from the first two years of the evaluation, to provide a more concrete and contextualised framework for the country case studies. This allowed us to explore the critical political economy dynamics affecting the observed and hypothesised causal links. The synthesis level CIMOs were mapped on to these country-specific ToCs, providing a common set of theories to explore in different contexts.

The definitions of 'policy quality' and overall impact were also revised, to enable an assessment of progress towards impact in the final evaluation. We unpacked our working definition of 'policy quality' from Stage 2 into four priority outcomes and an impact statement, to reflect *embedded, transparent, conceptual and instrumental* uses of evidence. Our definition proposed that, for evidence use to promote critical thinking, a decision-making process needs to be *transparent* about the limitations of evidence by engaging explicitly with diverse perspectives and values and multiple types of evidence, and it should be transparent about the extent of evidence and its quality. In this way, productive debate and discussion on the issues raised by evidence can be encouraged and evidence is likely to have a demonstrable influence on the decisions made, whether conceptual or instrumental. However, occasional uses of evidence are not enough to achieve the impact. A key part of BCURE's intended impact was for evidence use to become *embedded* in decision-making routines, supported by organisational systems and incentives to use evidence. These concepts linked to emerging DFID thinking on measuring the VfM of evidence-into-use interventions, and also linked to key insights in the EIPM literature (see Box 3). In this framing, the evaluation recognised that dimensions a-d may also contribute to *strategic, tactical* and *political* uses of evidence, as well as potentially evidence *misuse*, which would fall short of BCURE's anticipated impact.

Box 3: Insights from the literature: Understanding 'evidence use' in policy processes

The BCURE literature review highlighted the different ways that evidence is used in policy design, decision making and implementation. Weiss (1972, 1980, 1982) emphasised that **instrumental** use of evidence, where specific evidence directly shapes policy choices, is only one way – and is often quite rare. More common is where evidence generates a slow 'enlightenment' as concepts and theories from research gradually percolate through society, *'coming to shape the way in which people think about social issues'*. This was labelled **conceptual** use of evidence by Nutley et al. (2007). However, evidence may just as frequently be used to justify or refine a position that has already been reached, which can be thought of as **strategic**, **tactical** or **political** use. There are also examples of unambiguous **misuse**, when poor quality findings are used, or tactical use of evidence intentionally justifies a bad practice (Nutley et al., 2007). Finally, there are examples of **over-use**, where a set of findings may become a new fad and be applied uncritically and wholesale.

Several sources in the literature review emphasise that evidence itself is not a neutral product – first because it reflects pre-existing views, values and beliefs of researchers and commissioners involved in producing it; and second because it rarely points to an obviously optimal solution, so that contestation over its meaning is inevitable (see for example, du Toit, 2012; Waldman, 2014). This suggests the importance of considering the nature of the decision-making process, and how different evidence sources and stakeholder perspectives are consulted and interpreted.

Table 1. Dimensions of policy quality

Dimension a	Government officials routinely consider a range of evidence and the quality of evidence when developing policy products (embedded use)
Dimension b	Appropriate policy development processes engage with evidence from diverse stakeholders and multiple perspectives (transparent use)

Dimension c	Routine evidence use is *facilitated by evidence tools*, which allow officials to access, identify and critically appraise the evidence base and apply it to decisions, being transparent about the evidence base behind decisions (transparent use)
Dimension d	Routine evidence use is reinforced, incentivised and monitored through processes and standards, supported by senior managers (embedded use)
Impact	Together, a–d are expected to contribute to conceptual and in some cases instrumental use of evidence: evidence indirectly shapes the way in which people think about social issues and in some cases has a demonstrable influence on the decisions made … and ultimately evidence-informed decisions are implemented

During the Stage 3 design phase, the decision was made in consultation with the Steering Committee to prioritise a focus on the 'longer-term' theories crucial to explaining how and why the projects did and did not contribute towards the impact. This also reflected the need to limit the investigation to a smaller number of CIMOs in response to the challenges of breadth vs depth noted in Annex 3.12. Individual-level change was only investigated insofar as it helped to contribute to longer-term, more routine shifts in evidence, or contributed to changes at the organisational level. Theories relating to interpersonal change were not prioritised in part because the Steering Committee was less interested in understanding change at this level, and in part because Stage 2 suggested change at this level was important as part of the context that enabled individual and organisational change, rather than a standalone outcome. Finally, most changes in the institutional domain were beyond the scope of the evaluation to investigate in depth, and also relatively minor part of most project activities, so theories relating to this were only investigated in some contexts.

The table below details which CIMOs were prioritised for testing at Stage 3.

Stage 2 CIMOs	Status at Stage 3
CIMO 1. Self-efficacy 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	Not prioritised for investigation
CIMO 2. Coaching 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	Not prioritised for investigation
CIMO 3. Learning from similar challenges 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	Not prioritised for investigation
CIMO 4. Facilitated spaces for dialogue Facilitated spaces for dialogue (e.g. between policymakers, 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	Not prioritised for investigation

CIMO 5. Foot in the door 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	Tested
CIMO 6. Filtering up 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	Tested
CIMO 7. Cascading 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	Tested
CIMO 8. Showcasing 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	Tested
CIMO 9. Accompaniment 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	Tested
CIMO 10. Adoption 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)	Tested
CIMO 11. Facilitation 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	Tested
CIMO 12. Reinforcement 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	Tested
CIMO 13. Sustainable national hub 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	Tested only in Zimbabwe, which was the only context that aimed to build a sustainable national partner
CIMO 14. Relating and attracting 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	

Stage 3 final report: identifying impact pathways towards improved use of evidence

The Stage 3 synthesis process compared the contextualised country case study ToCs to identify how the tested CIMOs had played out in the different countries. This highlighted three main 'routes' towards EIPM taken by BCURE partners, at different levels of the government. They build on the idea of 'entry points at different levels' articulated earlier in the evaluation, but structure the findings in a more holistic, case-based way rather than disaggregating our findings according to levels of change. We have termed these three routes towards EIPM 'impact pathways':
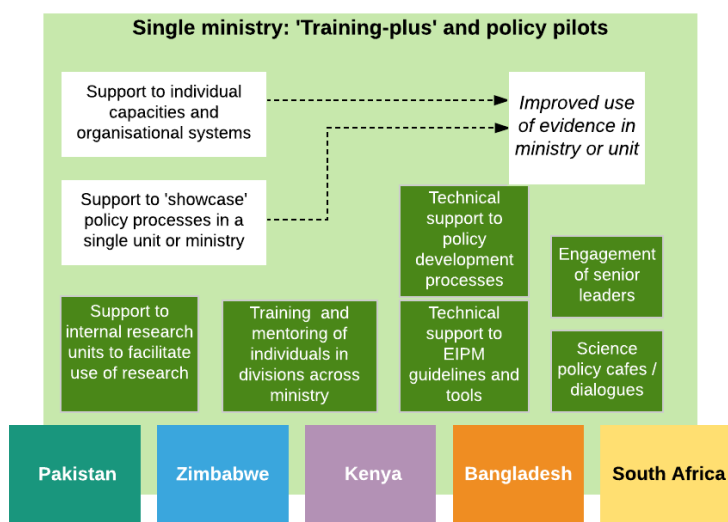
- **Impact pathway 1:** Support to a single ministry or unit
- **Impact pathway 2:** Working at a government-wide scale
- **Impact pathway 3:** Support to parliament

The impact pathways are archetypal programme theories (Funnell and Rogers, 2011), presenting a sequence of activities and outcomes from short term to long term, with evidenced causal explanations of how and why change has come about through BCURE. They are not mutually exclusive – most projects worked across two or more. The three impact pathways take the place of an overarching, portfolio-level ToC or programme theory, providing a rich and context-specific explanation of how and why capacity support can promote EIPM through entry points at different levels.

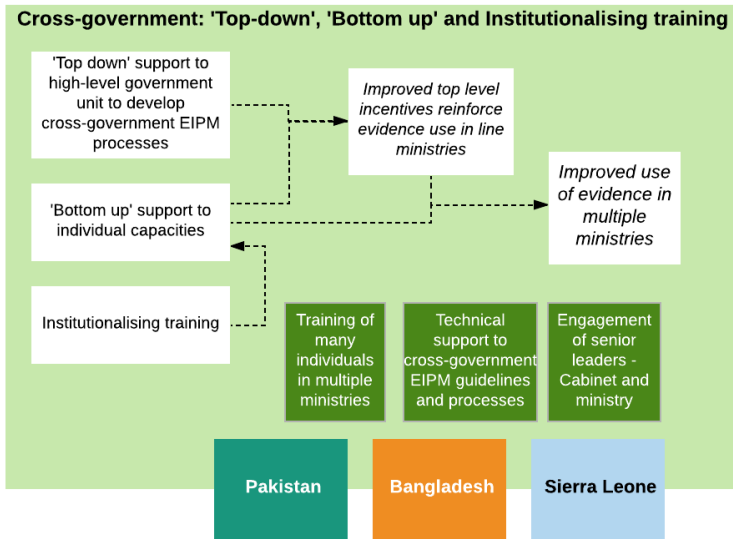### Impact Pathway 1: Support to a single ministry or unit

The single ministry pathway incorporates two approaches: 'training-plus', and technical support to pilot policy processes or EIPM tools. In the 'training-plus' approach, training on EIPM was given to technical officers responsible for policy formulation, who were then provided with follow-up support, to help sustain the application of new EIPM skills. Organisational support was also given to tools and guidelines that were intended to be adopted by ministries in order to facilitate and in some cases, incentivise and reinforce individuals to use evidence more routinely and more skilfully. In the second approach, some

BCURE projects provided technical support at an organisational level to accompany pilot policy processes, in order to 'model' systematic, evidence-informed approaches to policy development within the ministry, provide EIPM trainees with opportunities to apply their skills, and produce evidence-informed policy proposals. Other projects provided technical support to the development of data and evidence tools that aimed to showcase the value of evidence for decision making, intending for them to be adopted or replicated by government partners to help facilitate and embed evidence use in the ministry or unit.
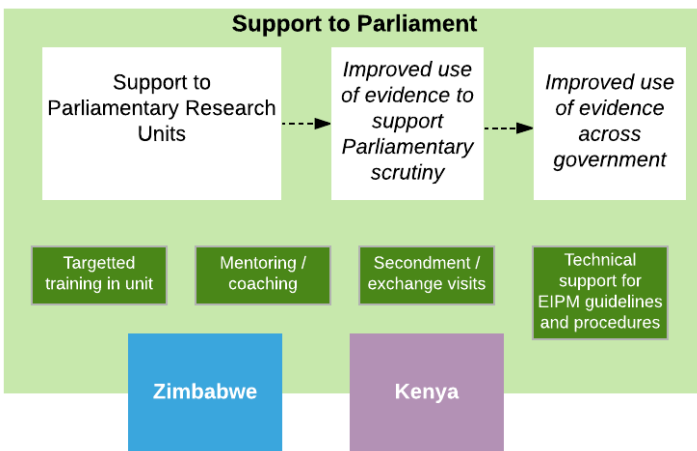
Impact pathway 2: Cross-government approach

The cross-government pathway incorporates three approaches: 'top down', 'bottom up' and 'institutionalising training' to promote EIPM. In the 'top down' approaches, BCURE supported activities such as working with cabinet to develop and roll out EIPM guidelines and procedures, often aiming to establish common cross-government standards for EIPM, facilitating various learning events, exchanges and training with senior government stakeholders to build high-level buy-in for EIPM, and supporting (mainly senior) officials in line ministries to develop evidence-informed policy processes, systems and structures in their ministries. In the 'bottom up' approach, BCURE programmes developed EIPM training courses and delivered it to large numbers of civil servants. The third approach involved BCURE also working to institutionalise EIPM training embedding EIPM curricula within national public sector training institutes. In all three projects working across government, at least two of the three approaches were adopted simultaneously.

Impact pathway 3: Support to parliament

This setting poses a different set of issues and challenges – as parliaments do not make policy, but can play an important role in interrogating it and holding line ministries to account. This pathway incorporates training research staff within a parliamentary research unit (as an entry point to parliaments), combined with follow-up individual and organisational support to strengthen parliamentary use of evidence in oversight and scrutiny functions. The two BCURE projects taking this approach promoted EIPM by focusing on individual and organisational change in parallel, as in the single ministry pathway. Training on EIPM was provided to a cohort, or all, research officers responsible for supporting MPs and committees with impartial and accurate analysis of policies and bills, and with evidence for committee enquiries. Trainees were then provided with follow-up support to help sustain the application of new EIPM skills, as well as offered learning exchange opportunities with other parliaments to further enhance learning about EIPM. BCURE also provided flexible, ongoing technical support, in collaboration with senior managers, at the level of the research unit to support trained researchers to cascade skills to non-trainees and develop procedures, ways of working, tools and guidelines that were intended to be adopted by parliament in order to facilitate researchers, and sometimes MPs themselves, to use evidence more routinely and more skilfully.

Testing and revising CIMOs

Within the overarching impact pathway frame, evidence on how and why changes had and had not happened were synthesised across the cases (method described in Section 3.4). This allowed us to formulate 'empirical CIMOs,' which unpacked 'what worked, for whom, and why' within each impact pathway in terms of the **outcomes** (O) that had come about in different case study contexts through various **mechanisms** (M), and the **context** (C) and **intervention** (I) factors that enabled or constrained the mechanisms. See below for an example.

Establishing 'top down' cross-government tools and structures: What worked, for whom, and why?

In **Bangladesh** and **Sierra Leone**, Cabinet Division and the Cabinet Secretariat decided to adopt and endorse new EIPM tools and systems (M, O) because these institutions had clear ownership over and buy-in to the process (C), in part a consequence of the support of high-level champions (C), and in part because of they had a mandate for reform aligned with BCURE's objectives (C). In **Bangladesh**, the framing of EIPM as a technical approach to improve policy formulation was a key selling point (I). Ownership was also promoted through BCURE's implementation approach, which was sensitive, flexible, and tailored to the local context (I): an approach that can be characterised as 'accompaniment' (M).

Following a high-level directive from Cabinet Division / Secretariat (C), EIPM tools and guidelines have been adopted by line ministries in **Sierra Leone**, and seem likely to be adopted in **Bangladesh** (M), where the policy pilots have to some extent successfully showcased (M) their value. However, insights from **Sierra Leone** and **South Africa** suggest that a one-off directive is not enough: ongoing engagement through 'carrots and sticks' (C) is necessary to ensure new tools are actually used (O).

These empirical formulations across the three impact pathways were then brought up to the level of middle-range theory, again following the synthesis approach described in Annex 3.4. The synthesis identified six key mechanisms that, when catalysed, led to positive changes around the use of evidence, although not all of them were always present in any one project. The key mechanisms do not operate in insolation, but instead work together to catalyse change, and build on each other so that where one mechanism operates it often creates a conductive context for another mechanism to 'fire'. These mechanisms are derived from well-established theories from psychology, sociology, development studies and governance – referenced throughout the report.

- **Accompaniment:** where an external partner provides tailored, flexible and responsive support to a government institution through a process of reform, characterised by a high level of trust, as opposed to a more traditional supplier / consumer model where ad hoc support is provided through one-off interventions. This often involves co-producing tools, systems or policy products.

- **Self-efficacy:** where providing information, opportunities to practise skills, coaching or technical support builds individuals' confidence in their ability to do their jobs or achieve a particular goal. This is akin to feeling of 'now I know how... (to find the evidence I need, to weigh up sources, to communicate evidence effectively).'

- **Facilitation:** where a tool, system or process for EIPM facilitates government officials to do their jobs or undertake a task more easily or efficiently.

- **Reinforcement**: where rewards or other forms of control create incentives that motivate officials to work in a particular way. Positive reinforcement includes rewards and encouragement, while negative reinforcement includes reminders, audits and mandatory requirements.

- **Showcasing**: where providing good examples of evidence tools or processes demonstrates the value of an evidence-informed approach, which leads to them being adopted elsewhere.

- **Adoption**: where senior government stakeholders decide to adopt a new tool, system or process for EIPM to help standardise EIPM within a government institution. This can be on a small scale (a unit

deciding to adopt a new template to standardise policy briefs) or a large scale (a government deciding to adopt a revised procedure for policymaking across all its line ministries that requires engagement with evidence). Adoption can happen for many reasons, and there is a risk that it might be a case of 'isomorphic mimicry' – where a new tool or system is adopted on the surface in order to access donor resources, without actually changing day-to-day practice.

The evaluation also identified a further mechanism that implicitly underpinned several BCURE project theories, but which has not (yet) catalysed in practice:

▪ **Critical mass**: where changes in practice among a sufficient number of government officials diffuse out to influence colleagues' behaviour, and the rate of adoption of new behaviours becomes self-sustaining. This diffusion may happen through cascading, where government officials formally cascade their new knowledge on EIPM through introducing new ways of working or new structures and processes. Or it may be through filtering out or filtering up: where improvements in evidence use by government officials leads to recognition of the value of an evidence-informed approach among colleagues (filtering out) or senior management (filtering up) which in turn influences' colleagues behaviour, or increases senior-level support for evidence-informed ways of working and/or organisational reforms to promote EIPM.

The final evaluation found that the key mechanisms do not operate in insolation, but instead work together to catalyse change, and build on each other so that where one mechanism operates it often creates a conductive context for another mechanism to 'fire'. Our revised CIMOs reflect this, representing our final tested theory about what works to build capacity for EIPM, for whom, and in what circumstances.

CIMO 1. Where there is genuine interest in partnership from high-level government stakeholders, existing incentives for evidence use in policymaking, and a window of opportunity to catalyse reform (C), an external partner can accompany EIPM reforms (M) in a participatory and collaborative way, providing tailored, flexible and responsive ongoing support that evolves over time (I) in response to emerging challenges and opportunities (C). This mode of working is greatly helped by the presence of high-level, enthusiastic and committed champions (C), and can create a conducive context for the other EIPM mechanisms to operate through encouraging government ownership (O) and building trust in the partner to work alongside government (O).

CIMO 2. Where information is provided about the importance of EIPM and how to access, appraise and apply evidence, alongside opportunities to practise skills, this can generate self-efficacy (M) and lead to individual behaviour change (O). Behaviour change is more likely to be sustained where there are clear incentives that motivate participants to apply their learning and reinforce changes in practice (M) – this includes management support to encourage and provide space for participants to access, appraise and apply evidence (C). Behaviour change is also more likely where activities are closely targeted to individuals who can apply their learning because it is directly relevant to their day-to-day work (I), and where activities are practical and participatory (I), provide practical tools (I) that facilitate trainees to do their jobs more easily (M), incorporate a focus on soft skills as well as technical skills (I), use knowledgeable, patient and confident facilitators (I), and tap into incentives to encourage participation (I).

CIMO 3. Where a cohort of officials start accessing, appraising and applying evidence more effectively, this can diffuse out to influence colleagues' behaviour (O) through a 'critical mass' effect (M). This is more likely when the cohort consists of a good number (I) of well-connected and clustered officials (C) in a unit with some reach and influence within the broader organisation (C), and where there are clear organisational incentives to use evidence (C) and senior management support to cascade learning (C) – potentially supported by a 'training of trainers' strategy (I).
Tentative theory, based largely on insights about blocking factors from Kenya and Zimbabwe, and insights from the wider literature.

CIMO 4. Where technical input is provided to support a policy process in an evidence-informed way, or develop a tool to improve evidence access, appraisal or use, this can generate high quality policies or

products (O) that showcase the value of evidence for quality, performance and delivery (M) and lead to adoption (O) and diffusion (O) of the procedure or tool. This is more likely where external actors 'accompany' government partners to co-produce policies or tools in a flexible, responsive and collaborative way (I), where policies are high priority or tools address a recognised problem (C), and where tools are intuitive and interactive (I) and genuinely facilitate officials to make decisions and do their jobs better and more easily (M). However, adoption can be stymied by many factors including shifting political priorities or resource constraints (see CIMO 5).
Strong evidence in support of theory, from Kenya, Pakistan, South Africa and Bangladesh (see single ministry pathway).

CIMO 5a. Where capacity support succeeds in showcasing the value of an evidence-informed approach, training course, tool or product (M) and /or generating tools that facilitate staff to do their jobs more easily (M), this can lead to a high-level decision to formally adopt the initiative to help standardise EIPM (O). Meaningful adoption is more likely where reforms have been co-produced by government and external partners through a flexible and collaborative process of accompaniment (C), and where there are high-level institutional and individual champions with a clear mandate for and interest in reform (C) who have access to resources to scale up or roll out the initiative (C). Adopted tools and processes, when attached to high-level incentives and encouraged through ongoing support rather than just a one-off directive (C), can then help reinforce (M) changes in practice at an individual and organisational level (O).
Strong evidence in support of theory from Bangladesh, Sierra Leone, Kenya, South Africa and Pakistan (see single ministry and cross-government impact pathways). Insights on factors that blocked adoption in Zimbabwe also support this theory.

CIMO 5b. Where capacity support succeeds in catalysing high-level ownership and buy-in to EIPM (C), it can position a national governmental unit to carry on promoting EIPM into the future (O), provided it is able to access resources (C) and buy-in is not eroded by high-level changes in priorities or staffing (C).
Tentative theory based largely on insights from the South Africa impact case study, and early observations in Bangladesh

The table below describes how the Stage 2 CIMOs were accepted or rejected, and how they translate into the final CIMOs at Stage 3.

Stage 2 CIMOs	Status at Stage 3	Corresponding Stage 3 CIMOs and supporting evidence
CIMO 1. Self-efficacy	This theory was not directly investigated at Stage 3, as there was strong evidence to support it at Stage 2, and the final evaluation prioritised theories relating to organisational change. However, several additional insights were gained through Stage 3 relating to the contextual and implementation factors that help this operate, including about how this mechanism interrelates with 'facilitation' and 'reinforcement' to help lead to routine behaviour change	CIMO 2
CIMO 2. Coaching	Not prioritised for investigation	N/A
CIMO 3. Learning from similar challenges	Not prioritised for investigation	N/A
CIMO 4. Facilitated	Not prioritised for investigation	N/A

Stage 2 CIMOs	Status at Stage 3	Corresponding Stage 3 CIMOs and supporting evidence
spaces for dialogue		
CIMO 5. Foot in the door	Tested: rejected. The Stage 3 evaluation unpacked a much more nuanced explanation of why BCURE partners gained a 'foot in the door', linking to 'thinking and acting politically' and accompaniment. These factors proved significantly more important in providing a foot in the door than starting the programme through training, across all four Stage 3 cases.	N/A
CIMO 6. Filtering up	Tested: Very limited evidence that this theory held; however insights on blocking factors from Kenya and Zimbabwe	CIMO 3. These theories are both facets of 'critical mass,' a theory implicitly underpinning several BCURE projects. Although there is only tentative evidence for them, there are insights from some contexts and the wider literature on blocking factors that inhibit the theory from holding.
CIMO 7. Cascading	Tested: Some evidence, at a relatively small scale, in the Kenyan MoH and parliament, and insights on blocking factors from Kenya and Zimbabwe	
CIMO 8. Showcasing	Tested: confirmed and nuanced with additional insights from Stage 3.	CIMO 4. Strong evidence in support of theory from Kenya, Pakistan, South Africa and Bangladesh
CIMO 9. Accompaniment	Tested: confirmed and nuanced with additional insights from Stage 3.	CIMO 1. Strong evidence in support of theory from across all six settings
CIMO 10. Adoption	Tested: confirmed and nuanced with additional insights from Stage 3.	CIMO 5. Supported by strong evidence in support of theory from Bangladesh, Sierra Leone, Kenya and Pakistan. Insights on factors that blocked adoption in Zimbabwe also support this theory
CIMO 11. Facilitation	Tested: confirmed and nuanced with additional insights from Stage 3.	Rather than standing alone, this mechanism was an important part of the context that helps spark routine behaviour change (CIMO 2), showcasing (CIMO 4), and adoption (CIMO 5). Supported by insights from Pakistan, South Africa, Kenya, Sierra Leone and Bangladesh
CIMO 12. Reinforcement	Tested: confirmed and nuanced with additional insights from Stage 3.	Rather than standing alone, this mechanism was an important part of the context that helped spark routine behaviour change (CIMO 2) in Kenya, Zimbabwe and Sierra Leone, was deemed important to incentivise behaviour change in Bangladesh, and was one of the main reasons for many trainees *not* changing their behaviour in Pakistan. This mechanism can also spark as an outcome of adoption (CIMO 5)
CIMO 13. Sustainable national hub	Tested in Zimbabwe, which was the only context that aimed to build a sustainable national partner. Theory was broadly confirmed, with some additional insights captured, explored in the VakaYiko case study report. However, this is not captured in the Stage 3 CIMOs given the limited evidence base or applicability to the other projects	N/A
CIMO 14. Relating and attracting		

5   Political economy analysis template for Stage 3

This template was provided to national consultants to complete prior to primary data collection

PEA purpose: Mapping of country and sector-level institutional arrangements

PEA focus: [Country] [Ministry] [Cabinet] [Parliament]

Purpose: To produce a descriptive mapping of national and sector-level political dynamics that affect policymaking in the focus sector.

PEA focus: Country – Specific Sectors- e.g. Ministry of Health and Parliament

1. **Objective:** To produce a **descriptive map** of the institutional and political dynamics around policymaking and parliamentary scrutiny of policy decisions. This should use as a starting point the contextual summary provided in the Stage 2 programme evaluation report.

2. **Process:** Please respond to the questions on Part 1 and 2 of the question framework below. Use the headings in bold to structure the document. This is an internal document, so please only provide informative bullet points or short sentences to provide information for each heading and sub-bullet. No need for a polished narrative.

    Also, please add any additional information that you feel is important but doesn't fit neatly under any headings.

3. **Input and output**: The document should be between 3-6 pages long. You should spend no more than 3-4 days on it.

4. **Data sources:** The consultant will need to use secondary sources, and possibly a key informant interview with the BCURE team.

Secondary sources need to be selected with consideration as to their relevance, reliability, accuracy, independence, timeframe and sourcing of the information. Wherever possible attempts should be made to corroborate the information used across independent sources, to ensure accuracy.

All sources cited need to be referenced in footnotes, with weblinks, following the Harvard style (see https://www.imperial.ac.uk/media/imperial-college/administration-and-support-services/library/public/harvard.pdf), and listed at the end of the document.

Secondary sources could include the following:

- Information produced by international sources, e.g. international non-governmental organisations country reports on health sector
- Information produced by bilateral and multilateral organisations, e.g. WHO,; World Bank; and/or United Nations Organisations' country reports; USAID, other donors, e.g. https://www.healthpolicyproject.com/index.cfm?id=country-kenya
- Information produced by other governments, for example country analyses produced by the UK government
- Independent media reports, e.g. *Financial Times*
- Information produced by the SECURE programme
- Information produced by the Kenyan government on the functioning of the Ministry of Health, and the various acts it is implementing

A key informant interview with the BCURE team should be conducted to gather sector-specific information that is not in key documents, and to enquire about additional sources and possible interviewees.

Identifying key informants

Some of the below questions will need to be explored through interviews, as the information will not be available through secondary sources. Through speaking to the BCURE partner, please identify 6–7 individuals who might be able to give insights into the questions below. (Some may be informants who have knowledge across the government system, while others have sector-specific knowledge relating to the Ministry of Youth / parliament. These may include civil society stakeholders, key political analysts, academics and government political reporters.)

PEA topic areas and questions

Please provide informative bullet points, no need for a polished narrative.

PART 1: General update on national-level issues affecting the [sector] since June 2016
To be completed before the country visit starts.

Overall events and trends in the [country] context since June 2016:
- What regional, national or international events / issues are having a major influence on life in [country]?
- What major issues and events have affected government and policymaking? E.g. economic, political, social, socioeconomic, environmental, health etc.?
- Any new government-wide initiatives introduced, e.g. anti-corruption measures; regulations; transparency?
- Have political incentives, ideologies or values changed and how is this affecting behaviours of politicians and citizens?
- What is the influence on today of historical legacies and change processes?

PART 2: Specific situation at a sectoral level (focusing on 2013–16). To be completed as far as possible before the country visit starts.

[Ministry]	[Parliament]
Actors and key players: *The actors (individuals or collectives) involved in making, influencing and delivering policy, including actors at different levels of government including those at sub-national and regional level*	
- What is the official status, role and the mandate of the Ministry? - How is the ministry structured and what are the key departments? - Who have been the dominant individuals within the Ministry, what is their role? - Have there been any reorganisations within the Ministry of since 2014? - Where does the **[xxx unit]** fit in? What services does it provide? - Who are the other key players that have an influence on youth policymaking, implementation and priority-setting, beyond the ministry? *Please include both national, county and international* - Where does the Ministry of Youth sit in relation to national and county-level policymaking and priority sectors? - How important and/or influential is youth as a sector in national politics?	- What is the official status, the role and the mandate of the parliament? - How is parliament structured and what have been the key committees (select committees; standing committees etc.) from 2013-16? - Who have been the dominant individuals within parliament, what is their role? - Where does the **[xxx unit]** fit in? What services does it provide? - How much influence does parliament have in government and public life, and what is the basis for this? - What power do other actors have to scrutinise parliament?

[Ministry]	[Parliament]
• How is the Ministry portfolio viewed by political appointments, and why? E.g. desirable, problematic, lucrative?	
Institutions: *The rules, norms, practices, relationships which influence individual and collective behaviour. These may be formal or informal, and exist at different levels within governments and departments*	
• Are the rules, roles and responsibilities for youth policymaking formally set out in the constitution? • Which entity at which level (national, sub-national etc.) is responsible for leading the policymaking process? • What changes have there been in the balance of power between national, county-level and international youth actors between 2013 and 2017? What are the implications of these for youth policymaking? • How are finances and resources allocated in Min of Youth, e.g. which actor approves budgets, e.g. cabinet and/or parliament? • Which key players provide the financing, e.g. international donors?	• What is the role played by parliament in national policymaking and priority sectors? (Its position from 2013-16, e.g. minor sector or major player in cabinet) • What is parliament's role in budgeting process and procedures, who approves budgets? • What are the formal and informal rules for scrutiny of policymaking and implementation between parliament and the executive? • What are the main Parliamentary bodies that provide scrutiny over the Ministry of Youth?
Policy networks: *the relationships between actors responsible for policy decisions and those who seek to influence it; the level of influence these groups have and the level of government consultation with them*	
• Have there been changes in the players/actors and/or interest groups that have influence on youth policymaking now? • What are the interest groups that the [Ministry] responds to? E.g. private sector health providers; international donors etc.? • What is the role of external actors on government policymaking (e.g. international donors; lobby groups; civil society groups)? • Which groups does the government consult most, and which coalitions seem to have the most influence (both in the youth sector and more broadly?) • What are the mechanisms for consultation with citizens? How are citizens are involved in policy development and monitoring (e.g. referendum, opinion surveys)?	• What inter-linkages are there between parliament and other organisations inside and outside government, e.g. policy analysts; universities; parliamentary committees, others? • What is the role of external actors on policy analysis and scrutiny (e.g. international donors; lobby groups; civil society groups)? • What are the mechanisms for consultation, participation and inclusion in policy processes and the way in which citizens are involved in policy development and monitoring (e.g. referendum, opinion surveys)?
Context: *The socioeconomic, demographic, and geographic factors that policymakers take into account when making decisions, and the routine (e.g. elections) and non-routine events with the potential to shift attention or provide an impetus to policy change*	
• Elections 2018: How is youth being presented and talked about in the media and social media as an issue in Zimbabwe in the run up to the elections? • What have been the events /issues that have affected specifically **[sector] policymaking** and **service delivery** in [country], e.g. scandals, strikes, protests etc. How has the government responded to events, negative and positive?	• Elections 2018: How is the run up to elections influencing parliament? • What have been the events /issues that have affected parliament specifically in [country], e.g. scandals, strikes, protests etc.?
Ideas: *ways of thinking and the extent to which they are shared within groups, organisations, networks and political systems. This includes the interplay between different forms of knowledge underpinning*	

[Ministry]	[Parliament]
action, the often deeply held beliefs of actors, and the ability of actors to persuade others to act in a particular way (e.g. through framing issues in specific ways)	
• What have been the main **[sector] policy priorities** between 2013 and 2016? • What are the main citizens' concerns about youth? • What are the narratives being presented in the media and social media about youth policies and service delivery? • Have there been changes in the beliefs, ideologies and values which shape the youth sector?	• What are the predominant values, narratives and perceptions that influence policy discussions in parliament, and what is the source of these narratives?

PART 3: Questions to be integrated into interview topic guides

Leadership, management and organisational culture:
• What are the leadership and management structures in the wider [Ministry], and what is their basis?
• What are the incentives and motivations that influence staff? Formal and informal, positive and negative?
• How do resource, capacity and skill levels vary across the organisation, including among managers and leaders, and with what consequences?
• How hierarchical is the organisational culture? To what extent can technical staff and political appointments challenge peers and seniors and/or express alternative views on policy issues?

Institutions and rules around policymaking:
• What is the influence of leadership, management on the content and direction of policy?
• In what specific ways does the **[xxx unit]** contribute to policy processes?
• What is the influence of dominant or prominent personalities on policymaking processes?
• What are the predominant values, narratives and perceptions that influence policy formulation, and what is the source of these narratives?
• How do power relations influence policy negotiation processes?
• What role does evidence, data and evaluation play in policy and decision making?

6 Sampling guidance

The below guidance was provided to BCURE lead evaluators prior to the Stage 3 country case studies

BCURE Stage 3 sampling guidance

Core principles of sampling

The sampling for Stage 3 is purposive, guided by the priority outcomes selected within the country-level programme theory. The aim is to achieve a *sufficient degree of confidence* in our hypotheses about the extent to which priority outcomes have occurred (EQ 1), BCURE's contribution to the outcomes (EQ 2) and how and why BCURE contributed or failed to contribute (EQ 3).

Our Stage 3 sampling approach is guided by four main principles. Bear these in mind and try to keep them front-and-centre when developing and iteratively revising the sample!

1. **Sampling will be driven by theory.** The starting point is priority outcomes and evidence tables. Who do we need to speak to in order to generate evidence for hypotheses?

2. **Sampling will be iterative,** allowing for changes and additions during field work as theories develop and leads are followed. The sample will therefore continually evolve throughout the data collection process.

3. **Sampling will aid comparison between sub-groups:** A key element of our sampling strategy will be *comparison* between different sub-groups of participants, in order to investigate how change has or has not occurred for different people / units / ministries etc., and to help explain how and why these differences exist. Sub-groups will evolve over time as our understanding develops. Although we will be limited by resources as to the number of sub-groups it will be possible to explore, we will be guided by the emerging evidence on what seems to be most important in explaining the outcomes.

4. **Sampling will seek to maximise triangulation of sources for each hypothesis**: We will aim to triangulate evidence against our hypotheses across a range of different stakeholders, through comparing insights from project participants with insights from knowledgeable 'outsiders' (informed by the PEA of who is influential in relation to the outcome), and through accessing secondary documentation where available. Our data sources are detailed below.

Five categories of data sources:

The aim is to triangulate insights for each priority outcome from as many of the below categories as possible.

1. **Monitoring data and other programme documentation**. This will be reviewed first to examine secondary evidence for hypotheses. It will also help identify relevant sub-groups of individuals to target for interviews.

2. **Interviews and workshops with programme staff**. This will supplement the monitoring data, helping understand gaps or areas where greater clarity is needed. It will also provide an insight into the areas project staff think have been more or less successful and how and why, which will help further develop our theories.

3. **Interviews with project participants (individuals directly targeted by the activity / activities which aimed to contribute to the outcome)**. This will generate self-reported insights about the extent to which outcomes have been achieved or not achieved, how and why, for different groups.

4. **Interviews with other knowledgeable stakeholders.** These are stakeholders who did not participate in BCURE interventions, but who can provide insights into (a) outcomes observed and the reasons for these outcomes; and (b) political economy issues that relate to priority outcomes. This group will be considerably enlarged this year, in order to address concerns of the Steering Committee and EQUALS review that the Stage 2 report did not sufficiently deal with potential confirmation bias from project participants.

5. **Secondary (non-project) documentation.** This is documentation *not* produced by the programme, which provides insights into outcomes and the reasons for outcomes. In previous years, it has not proved possible to access this documentation, and there remains a major risk that important documents will remain impossible to access this year. However, in previous years we have been unable to dedicate much time to identifying and attempting to secure relevant documents, rather relying on BCURE programmes to share documents they had access to. This year, time will be dedicated to identifying and securing potentially relevant documentation (a) up-front when evidence is assembled; and (b) throughout the data collection phase, using interviews to attempt to secure documents that can triangulate insights from respondents.

Sampling rules of thumb

- Aim for roughly **60** interviews in total. However, don't feel the need to interview people for the sake of it! If it is not possible to reach this many people with knowledge of the programme, you might want to do additional interviews with PEA informants, or focus more resources on trying to get hold of (and then reviewing) secondary documentation.

- Around **30–35** interviews with *programme participants* (predominantly government, but also civil society and non-government stakeholders where they have been direct targets of the programme).

  o Consider the 'rule of three': where possible, aim to speak to 3 people from each relevant 'for whom' 'sub-group' – see below

- Around **20** interviews with <u>people who were not involved in the programme </u>but who can give insights into whether outcomes were achieved / how and why, and PEA factors.

  o This should include around **5–7** interviews with people who may have no knowledge of the programme but who can give insights into political economy factors relating to the relevant sectors (e.g. political analysts, academics, think tank or civil society stakeholders, DFID staff)

- You should conduct an in-country workshop (and if you think relevant, individual interviews) with BCURE staff.

- If relevant to understanding activities conducted since last year it may be helpful to conduct a small number of interviews with trainers, mentors, and other facilitators of activities.

- Interviews don't all need to be lined up in advance – there should be flexibility to add new stakeholders once in-country.

- Sampling decisions should be transparent: documented clearly using the sampling spreadsheet. The sampling spreadsheet should also be used to identify in advance which outcomes and CIMOs will be tested with which people, to ensure we are being systematic.

Sub-groups

There are two types of 'sub-group' we are interested in:

1. Different sub-groups **targeted by the programme**. These may be individuals from different target ministries, units or departments, male and female participants, more junior and more senior participants.

2. Sub-groups **associated with differential outcomes**. These may or may not be the same as the sub-groups targeted by the programme. This is essential for testing our theories (CIMOs) about how and why BCURE works, and understanding *for whom* BCURE works. For example, did some senior staff demonstrate more buy-in to an EIPM agenda following an intervention than others? If so, it will be important to try to speak to individuals (and if possible their colleagues / peers) from the 'more buy-in' and 'less buy-in' sub-groups in order to understand what it was about these individuals or the wider context that enabled or constrained buy-in. Iteration is vital – as our theories and our

understanding of differential outcomes evolve over the course of data collection, new sub-groups will emerge and others disappear or subsume into broader groups.

Steps in the sampling approach

1. Start with last year's sample. Either save a new version of last year's spreadsheet and amend it to reflect the Stage 3 sample template (saved in Dropbox), or copy relevant stakeholders into the Stage 3 template. Please do make sure you're using the Stage 3 rather than Stage 2 template, as changes have been made to the stakeholder categories etc.

2. Work through each priority outcome in turn and consider who will be able to give insights into the EQ 1, 2 and 3 hypotheses, and then add them to the sample. You might know these people by name from last year's stakeholder lists, or you might just indicate their role at this stage [e.g. 'someone high level in the MoY']. Note that you don't need to do this separately for outcomes 17a-d, and impact level change – this should be covered through considering all of the previous outcomes.

    i)   Make sure you've included both <u>participants</u> in BCURE activities and <u>non-participants</u>, for each priority outcome. Non-participants might include:

        ▪   Managers or colleagues of participants

        ▪   Senior stakeholders from the department

        ▪   Members of other units or teams who work with participants' teams

    ii) Make sure you've thought about <u>potentially relevant sub-groups of participants</u> in relation to that outcome, and where feasible tried to include 3 members of each.

        ▪   E.g. three junior and three senior trainees from the Ministry of Youth

    iii) Start compiling a list of potentially useful secondary sources of evidence relating to that outcome, to try and track down.

3. Draw on the following sources in order to continue building the sample in line with the theory

    i)   The evidence we already have for each of the priority outcome hypotheses (in the evidence tables). This should give insights into the additional data needed (on top of what we already have) to collect to test each part of the theory.

    ii) The document review, which should give more insights into who might be important to test particular outcomes, and which may also include participant lists to draw on when developing the sample.

    iii) The PEA review from the national consultant, which may give insights into people who can provide insights into various outcomes or into PEA issues.

4. Share draft sample with partner for comments, additional suggestions and contact details, and ideas about who should be a priority to talk to.

5. Once fieldwork is under way, ask interview respondents to identify further individuals who can provide an insight into a particular outcome, or who are members of a particular sub-group that is emerging as important. This strategy will be crucial to identify knowledgeable non-participants in BCURE interventions, who may be unknown to programme staff and therefore difficult to identify up-front. Also make sure you ask respondents about any potentially relevant documentary evidence.

Sampling spreadsheet

This spreadsheet was used by country case study leads, to record details of potentially relevant stakeholders and aid with the iterative and purposive sampling process.

	Stakeholder type	If programme participant - which interventions?	Priority outcomes to test / PEA stakeholder	CIMOs to test	Include/ exclude decision	Rationale for inclusion / exclusion	First name	Family name	Sex	Organisation	Job title	Government stakeholder?	Location	Interviewed at Stage 1?	Interviewed at Stage 2?	Email	Phone	Comments
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		

7   Topic guides

7.1   Instructions for customising topic guides, for case study leads

Topic guides will need to be contextualised for individual stakeholders, around the outcomes and CIMOs prioritised for investigation. At Stage 3, we are speaking to three broad categories of stakeholders:

Programme participants	Individuals (government or non-government) directly participating in BCURE interventions (training, mentoring, workshops, knowledge cafes, policy dialogues, discussions around organisational systems development etc.)
Non-programme participants	Other knowledgeable stakeholders, who did not participate in BCURE interventions, but who can provide insights into a) outcomes observed and the reasons for these outcomes, and / or b) political economy issues that relate to priority outcomes
Programme staff	Individuals managing the programme, in-country and in the UK, including consortium partners. This also includes individuals hired by the BCURE partner to deliver training and mentoring, facilitate sessions etc.

- Prior to data collection, you will need to develop specific outcome probes for each of the priority outcomes in the country-level ToC. The aim is to find out what happened, and test the extent to which the outcome in the ToC actually emerged. The document review should help guide the formation of the questions (e.g. mentioning specific outcomes that you want to validate).
- As your understanding of the context and emerging outcomes develops, the outcome tables are likely to need refinement and new questions added to test emerging and more specific outcomes.
- We have developed CIMO tables to help probe specific priority CIMOs across the cases. These can largely be copied and pasted as-is into the topic guides, although the introductory questions and prompts may need some contextualisation. You'll need to decide which CIMOs are most relevant to which interviews in advance.
- You should decide in advance, as part of the sample development, which outcomes and which CIMOs to discuss with respondents.
- We have found that it is possible to probe a minimum of 2 outcomes and 2 CIMOs in depth within one interview (not including the longer-term outcomes 17a-d which have a separate set of questions at the end of the guide). Sometimes it is possible to test a much larger number of outcomes and CIMOs – but this is dependent on the flow of the interview and the extent to which change has or has not been observed (where change is minimal, it is easy to run through a larger number of outcomes).
- Use the sampling spreadsheet to keep track of which outcomes and which CIMOs you have discussed with which respondents and make adjustments if necessary, to ensure that you are testing the theory of change systematically.

Guide to developing probes for outcomes

EQ 1 probes

- Insert questions to examine the extent to which priority outcomes have come about.
- You should aim to ask a question for every link in the ToC that you want to test (e.g. if two arrows point to an outcome representing two specific causal pathways, and you want to test both links, you should aim to ask about them both).
- You will need to adjust and add new questions as the data collection progresses, to test emerging and more specific outcomes. For example if someone mentions a very specific outcome, it will be important to test this with others in order to triangulate insights.

EQ 2 probes

- Insert questions to examine the contribution of BCURE to priority outcomes.
- You should aim to ask 'why' or 'why not' for every outcome you are testing in the interview.
- Sometimes generic questions will be sufficient:
    - *What were the drivers and influences that led to this / prevented this from happening?*
    - *What do you think caused these changes / what is inhibiting change?*
    - *What was BCURE's contribution to these changes / initiatives?*
    - *Apart from the BCURE programme, has anything else contributed?*
- But sometimes it will be helpful to ask specific contribution questions relating to the outcome, especially for outcomes expected to emerge directly from BCURE activities. Other contributory factors may emerge as important during the data collection, which you may want to probe:
    - *Apart from BCURE, have you attended any other training courses / learning exchanges relating to evidence use?*
    - *Have you heard of xxx initiative? Do you think this contributed?*

Example outcome probes

Outcome 3. Government stakeholders apply, promote and communicate evidence routinely in their day-to-day work due to training…leading to…   *Outcome 8. Wider cohort of officials (beyond initial trainees) accessing, appraising and applying evidence more*
EQ 1 questions   **[For programme participants]**   • *Individual.* Has anything happened about how <u>you</u> work with evidence in your day-to-day work since the programme started? Can you give me some examples? Are there any written examples of work you can share with me?   • What have you noticed about how <u>your colleagues</u> are working with evidence on a day-to-day basis – have there been any changes? Can you give me some examples?   • *Non-trainees in unit / dept*: Have you seen any signs that the training has influenced people who weren't actually trained? Can you give me some examples?   • *Senior staff:* Have you seen any signs of the training influencing senior staff? Can you give me some examples?   **[For non-participants]**   • *Non-trainees in unit / dept.* Has anything happened about how <u>you</u> work with evidence in your day-to-day work since the programme started? Can you give me some examples? Are there any written examples of work you can share with me?   • *Trainees*: What have you noticed about how <u>your manager or your colleagues who went on the training</u> are working with evidence on a day-to-day basis – have there been any changes? Can you give me some examples?   • *Senior staff:* Have you seen any signs of the training influencing senior staff? Can you give me some examples?
EQ 2 questions   • What do you think caused these changes?   • Apart from the BCURE programme, has anything else contributed?   • What other training programmes or capacity-building opportunities are available within your unit / Ministry? Have you taken part? What was the content / what did you learn?   • Are you involved in any donor programmes at the moment? Do you think this has contributed?

CIMO question tables

CIMO 5: Foot in the door

One idea we have is that starting with a relatively neutral intervention like training might have provided a 'foot in the door' for BCURE. In other words starting with training generated permission and buy-in for them to begin implementing organisational reforms.

IF YES	IF NO
• Why do you think this was? • What was it about the way BCURE engaged or provided the training that was important? Is it important to be collaborative? Is it important to be flexible? • Would BCURE have been able to come straight in and work at that level? What is it about the context that makes this type of approach important? (Probe from PEA framework)	• Why do you think that didn't happen? *(Probe for C and I factors that might have blocked the mechanism)*

CIMO 6: 'filtering up'

One idea we have is that when enough people begin using evidence in a department, this can 'filter up' and make senior staff and peers recognise the value of an evidence-informed approach. Has it worked at all like that here?

IF YES	IF NO
• Can you give examples? • What helped this 'filtering up' to happen here? • What other incentives were there, from the organisation or management? • How important was it that a group of people were trained at the same time? • How did peers/junior colleagues [who were trained] inspire / help you to work with evidence? • What other factors were at play? ○ *Need to probe for PE factors from the framework*	• Why do you think that didn't happen? *(Probe for C and I factors that might have blocked the mechanism)*

CIMO 7: 'Cascading'	
One idea we have is that when enough individual people, sometimes in mid-level roles, have been trained in using evidence, they can cascade new skills or introduce new ways of working with evidence to their teams. Has it worked at all like that here?	

IF YES	**IF NO**
- Can you give examples? - What helped this 'cascading' to happen here? - What other incentives were there, from the organisation or management? - How important was it that mid-level managers [who were trained] were committed and passionate about EIPM? - How important was it that they had senior management support? - What was it about how junior colleagues [who were trained] that helped them to inspire you to work with evidence? - What other factors were at play? *Need to probe for PE factors from the framework*	Why do you think that didn't happen?

CIMO 8: 'Showcasing'	
One idea we have is that tools or systems that help staff to use evidence more effectively (for example xxx) can act as what we call 'showcases' - good examples that promote and highlight the value of evidence. Has it worked at all like that here?	

IF YES	**IF NO**
- Can you give examples? - What helped this 'showcase' to happen here? - How important was it that the showcase used innovative approaches? - How important was it that the tool provided opportunities to 'learn-by-doing' for staff? - How important was the tool development was undertaken in a collaborative way by the partner? - What other incentives were there, from the organisation or management? - What factors helped the showcases inspire other reforms or new approaches? - What other factors were at play? *Need to probe for PE factors from the framework* - Did it matter which policy was chosen for support?	Why do you think that didn't happen?

o Who decided?	

CIMO 9: 'Accompaniment'

One idea we have is that when a government unit has developed trust in an external partner through a few collaborative activities, they allow the partner to 'accompany' policy processes and help embed evidence use. We call this 'accompaniment' - basically providing close-up, tailored and flexible technical support. Has it worked at all like that here?

IF YES	IF NO
- Can you give examples? - What helped this 'accompaniment' to happen here? - How important was it that there was already a pressure to improve policy development processes from senior levels? - What other incentives were there, from the organisation or management? - What factors promoted trust in the partner? - Does this 'accompaniment' also help to optimise the government unit and strengthen their abilities to champion EIPM internally? In what ways? - What factors helped the internal unit to optimise its own work to promote EIPM internally? - What other factors were at play? *Need to probe for PE factors from the framework* - Did it matter which policy was chosen for support? - Who decided?	Why do you think that didn't happen?

CIMO 10: 'Adoption'

One idea we have is that if an external partner provides technical support to co-produce tools and processes for using evidence, this can spark a high-level decision to formally adopt them as official procedures to help standardise and embed evidence use within the organisation. Has it worked at all like that here?

IF YES	IF NO
- Can you give examples? - What helped this 'adoption' to happen here? - What other incentives were there, from the organisation or management? - How important was it that the tools/procedures were collaborative? - How important was it that they were backed by senior stakeholders/managers? How did that support come about? - How important was it that the EIPM procedures link to other formal processes? - What other factors were at play? *Need to probe for PE factors from the framework*	Why do you think that didn't happen?

CIMO 11: 'Facilitation'	
One idea we have is that evidence tools can provide practical assistance to people, essentially helping them do their jobs better or more easily, which means that tools actually get used. Has it worked at all like that here?	
IF YES • Can you give examples? • What helped this 'facilitation' to happen here? • How important was it that the tool helped you to do your job better? [benefit] • How did it do that, can you give me an example? • What other incentives were there, from the organisation or management to use the tool? • What other factors were at play? o *Need to probe for PE factors from the framework*	**IF NO** • Why do you think that didn't happen?

CIMO 12: 'Reinforcement'	
One idea we have is that when there are incentives to use evidence tools or procedures [e.g. xxx], this 'reinforces behaviour' – basically providing positive or negative incentives that lead individuals to change how they work with evidence. Has it worked at all like that here?	
IF YES • Can you give examples? • What helped this 'reinforcement' to happen here? • What other incentives were there, from the organisation or management? • How important was it that the procedures had senior management backing/authority? • How important was it that the procedures include monitoring use of evidence? • What other factors were at play? o *Need to probe for PE factors from the framework*	**IF NO** • Why do you think that didn't happen?

CIMO 13: 'Institutional local actor catalysed'
One idea we have is that local organisation, delivering EIPM technical support as part of an international consortium, can lead to that local actor becoming optimised as a 'hub' for EIPM support nationally, beyond the end of the programme. Has it worked at all like that here?

IF YES	IF NO
- Can you give examples? - What helped this to happen here? - Is it important that the actor is part of an international consortium? Does that help to build credibility / provide access? - How important was the 'learning by doing' aspect (through being part of an international consortium?) - What other incentives might have stimulated this, from the national sector or internationally? - What other factors were at play? *Need to probe for PE factors from the framework*	Why do you think that didn't happen?

7.2   Guide for workshop with programme implementing teams

1. Aims
    - For the evaluator to understand fully what interventions have been implemented
    - To check and validate the country ToC produced by the evaluator
    - To explore EQs 1, 2 and 3 in relation to the 'priority outcomes' in the country ToC

The workshop will focus solely on the evaluation case study country, and will concentrate on the sectors / ministries / activities relating to the priority outcomes identified through the evidence mapping.

2. Set up and materials
    - The session will be informal and participatory.
    - With participants' consent, we would like to record the session to ensure we have an accurate record of the team's insights.
    - It would be helpful to have access to a whiteboard, or a screen / wall where it is possible to put up post-its and flip chart paper.
    - We will bring post-its, but if flipchart paper and marker pens are available this would be helpful. If not we can bring these with us.
3. Practical tips for facilitators
    - Recording the conversation is a good back-up for detailed notes taken in the session.
    - The workshop can be treated as a structured group discussion, but post-its and CIMOs on flipcharts are useful visual prompts.
    - It may be worth preparing post-its or flip chart in advance, with lists of activities and outcomes drawn from the Stage 2 reports and the document review.

4.   **Agenda and process**

Session	Details	Instructions for facilitator
Session 1: **Introduction** **10 mins**	Introduction Discuss the aim of the workshop and its role in the data collection Update on revised evaluation approach for Stage 3	**During introduction session:** - Explain purpose of the session – describe aims - Get consent to record - Introductions around the room - If you think relevant, potentially repeat the Stage 2 Icebreaker – card sort – how do you feel about the evidence-informed decision-making landscape in this country now that the programme has ended / is shortly ending?
Session 2: discussion of programme activities **30 minutes**	Discuss key programme activities in the case study country, additional activities since the Stage 2 evaluation, and rationale for any changes	Map out programme activities using post-its / flip chart. Begin with a list of activities from Stage 2 / drawn out from the document review, and verify these – using the opportunity to clarify any questions that may have arisen from the document review about what the programme has done and why *Check that there are no gaps in our understanding of what has been done.* *Clarify understanding of stakeholder groups – both department and role, and more conceptual (higher vs mid-level gov; colleagues from same dept vs people from diverse depts). Also probe for the programme team's rationale for targeting particular groups, through these*

Session	Details	Instructions for facilitator
		interventions. Why were the interventions combined and sequenced in this way? *How does gender come into your understanding of stakeholder groups?* *Looking back to the project plans this time last year – has there been any evolution in terms of your focus or activities?*
Session 3: Discussion of EQ 1 **45 minutes**	Discussion of EQ 1: **To what extent have priority outcomes been realised and for whom?** *a) Have the changes hypothesised in our country ToC happened?* *b) How does change differ for different sub-groups, organisations etc., reflecting on gender and equity issues?* *c) How sustainable is the change?* Also highlighting any examples of policy processes that the programme may have influenced, which may be the focus of the embedded policy case studies Throughout the discussion, highlighting relevant data sources (individuals to speak to and documents) that will help the evaluation to evidence outcomes and programme contribution.	Map out outcomes observed at Stage 2. Suggest preparing this in advance using post-its / flip chart. Put outcomes up alongside activities (starting to build the ToC) Ask: what else has happened since last year? Have any of these outcomes deepened, or led to further outcomes? It may be helpful to have pre-prepared post-its representing higher-level outcomes in our country ToC – to bring into this conversation in order to verify the ToC (is this how you saw things happening? Or was it something else?) Has anything *not happened* that you hoped would happen, or has anything taken longer or been more difficult than expected? (Again, post-its representing outcomes in the country ToC might help facilitate this discussion) Put up our priority outcomes 17a-d. Discuss whether the programme sees progress towards those outcomes, and where the links are from earlier outcomes. **Discuss sustainability:** • Looking at the changes observed so far, what do you hope / expect will happen now that / once the programme has ended? • What is it about your programme that you hope will make the changes sustainable?
Session 4: Discussion of EQ 2 **30 minutes**	**How significant was BCURE's contribution to priority outcomes, alongside the contribution of non-** **BCURE factors?** *a) What is the evidence that BCURE contributed to causing the observed change, and what is the evidence that non-BCURE factors contributed?* *b) What is the relative importance of BCURE and non-BCURE factors in explaining the observed change?*	Focus on the longer-term, priority outcomes. Ask: • How significant do you think BCURE's contribution was to this? • What else was going on, that might have contributed to this outcome? Use insights from the document review and PEA analysis as prompts (e.g. other programmes, supportive elements within the context, etc.)
Session 5: Discussion of EQ 3	**How and why did BCURE contribute or fail to contribute to priority outcomes?**	In advance of the workshop, write up prompts for relevant ICMOs on flipchart / print them out on A3 [A3 prompts saved in Dropbox].

Session	Details	Instructions for facilitator
1 hour	*a) Through which mechanisms, enabled by which features of the intervention and features of the (individual, interpersonal, organisational and institutional) context, did BCURE contribute to the observed change? (Testing and confirming/rejecting CIMO configurations)* *b) Where hypothesised change did not happen, or where BCURE did not contribute to observed change, how and why was this the case (through which mechanisms, features of the intervention and features) of the context)? (Testing and confirming/rejecting CIMO configurations)*	Starting with the priority outcomes, ask: How and why do you think the programme contributed / struggled to contribute to this change? Write up on post-its / flipchart, or annotate our CMO prompts to show how intervention leads to change • *What is it about the programme that led to the change?* • *What did the programme provide that was new? (Information, skills practice, opportunities for collaboration, technical support, access to evidence sources, etc.?)* • *What is it about the way the programme is implemented that made a difference, or failed to?* • *What is it about this place / context that makes the intervention work or made it not work?* • *Has change happened in the same way for all participants? In what ways it differed, and for whom? What is driving these differences?* *Bring CIMO prompts into the conversation when relevant, and ask the team to reflect on them. Aim to test all of the CIMOs identified as a priority in the evidence mapping process.*
Close **10 mins**	**Wrap up and final reflections** What else do you think we need to know, to really understand how this program has worked here?	Thanks for participation Repeat what we'll do with the info

7.3   Programme participant topic guide

Case study and country	
Interviewee name	
Position and organisation	
Interviewer name	
Date of interview	

Introduction

- We are independent researchers investigating the [xxx] project, which is funded by the UK Department for International Development. We want to hear your thoughts on this project.
- The interview will last about 1 hour.

Consent

- Everything you tell us will be confidential, and your name will not be used in any of our reports. However, we would like to use your thoughts and some anonymised quotes from the interview in our findings, if you are happy with this?
- Do you mind if we audio record the interview? This is for the researchers' reference and will allow us to check that we have we recorded your views correctly.
- Do you have any questions about the research, or concerns you would like to raise before we start?

Aim of the interview

- We'd like to talk to you about what has changed, if anything, since the start of the programme. However, this interview might be slightly different to others you may have had in the past.
- We're not just interested in whether the programme has been successful – we want to know how and why. So I'm very interested in your ideas about how and why things have changed, or not.
- We have some initial ideas but we're not sure if they are correct or not, so we will share these with you during the interview and get your thoughts.

Role and involvement in policy / decision making

- Could you please introduce yourself and your role within the organisation?
- Can I briefly check – how would you describe your role in relation to [or how are you involved in] policy and decision making? Can you summarise that for me please?
- Can you tell me what your involvement in (or contact with) this programme has been?
  - *What were the specific activities and when did you participate?*
  - *When was your first contact, and when was your last contact with the programme?*

Description of project (for stakeholders without much knowledge of it)

- The [xxx] programme aims to [encourage the use of evidence in policy and decision making / insert local description here].
- In [xxx country] [insert local description here: e.g. AFIDEP has been leading the SECURE Health programme, providing training and coaching to staff in parliament and MoH, providing technical support to health policy, convening science policy cafes and strengthening policy and research networks.]
- The programme as a whole is funded by DFID and worked across 11 countries in Africa and Asia.
- We are evaluating the programme in order to investigate how effective the programme has been, but also to understand more about how and why different types of approaches can help support evidence-informed policymaking in different contexts.

EQ 1 and 2: Outcomes and contribution

- What do you consider the outcomes of the programme to have been for you personally?
- What do you consider the outcomes of the programme to have been for [the organisation / Ministry]?

- o *Probes: last year you said that xxx had happened. Is this still the case?*
- What do you think caused these changes?
- Apart from the BCURE programme, has anything else contributed?
  - o *Probe for other factors using other PE prompts, e.g. other initiatives; we have heard that the President is keen on EIPM, to what extent so you think this has been a factor*

<mark>Insert outcome probes here. Use them to make sure all of the priority outcomes for this respondent have been explored.</mark>

- Are there things that **did not** happen as a result of the programme [that you hoped would], or results that were **more limited** than you hoped?
- Do you think that the outcomes have been the same for all [people within the specific stakeholder group – e.g. trainees, mentees]? In what ways have they been different?
  - o *Probe for examples of people who have been less engaged with the programme / haven't felt the benefits of the programme. Attempt to get names.*
- Have you noticed any differences in outcomes that relate to gender?

Policies being directly influenced by evidence

- [Where respondents have given examples of changes to practice] Do you know of any examples where a policy or bill (that you've worked on since you've taken part in the training?) has been directly or indirectly influenced by evidence?
- Can you tell me a little about it and who was involved please? *[Ask some of the following questions if time, prioritising questions that can only be answered by this respondent.]*

  - o What was the purpose / goal of the policy / bill?
  - o What was the outcome?
  - o Did the training / support from ZeipNET feed in? How?
  - o What were the other drivers of success?
  - o Did it face any obstacles or blockages?
  - o Who else was involved? (Government stakeholders, civil society?) – can you give us names / contact details?
  - o Are there any documents we could look at?

EQ 3: How and why did BCURE contribute or not contribute?

For each of the changes and non-changes mentioned throughout the interview:

- You said that xxx has happened and that the programme contributed to that. Why do you think the programme made a difference to xxx?
  OR
- You said that xxx hasn't changed / the programme didn't contribute. Why do you think the programme **didn't** influence this?

<mark>Insert CIMO tables here. Use them to make sure all the CIMOs prioritised for this respondent have been covered</mark>

Longer-term outcomes/sustainability

You may not need to ask these questions separately. It may be possible to ask about 17 a, c and d in one question, as they are interlinked. Always ask about outcome 17a as this applies in all BCURE contexts. Ask about the other outcomes where relevant activities have been undertaken / results observed within the programme.

Outcome 17a	You've mentioned x, y, z [in relation to individual behaviour change]. How <u>consistently</u> do you think officials [in xxx department] use evidence [in their day-to-day work / when developing [policy briefs, concept notes etc.]? How far would you

Always ask this question	say the [department] is along this journey from evidence use being ad hoc vs being more consistent and routine? *Underline:* What else needs to happen for this to become ***sustained*** as a routine? What other factors might influence this? E.g. role of external actors on policy analysis and scrutiny (e.g. international donors; lobby groups; civil society groups)?
Outcome 17b *Ask if there have been activities and results relating to consultation process etc.*	You've mentioned x, y, z [in relation to policy development processes.] How far do you think policy processes in [the parliament] in general engage with evidence from different stakeholders and perspectives? What else needs to happen for this to become ***sustained*** as a routine? E.g. What are the mechanisms for consultation, participation and inclusion in policy processes and the way in which citizens are involved in policy development and monitoring (e.g. referendum, opinion surveys)?
Outcome 17c *Ask if there have been activities and results relating to evidence tools*	You've mentioned x, y, z [in relation to tools]. Are these tools being consistently used by people? Or is the use of the tools more ad hoc? What else needs to happen for this to become ***sustained*** as a routine? What other factors might influence this?
Outcome 17d *Ask if there have been activities and results relating to evidence processes / standards*	You've mentioned x, y, z [in relation to processes / standards]. Do you think they are helping to promote consistent evidence use by people in [the Ministry / x department]? Are they supported by senior managers? What else needs to happen for this to become ***sustained*** as a routine? What other factors might influence this?

Wrap up

- Finally, if you could change something about this [intervention] to make it work more effectively here, what would you change and why?
- Is there anything else you think we should know about the programme that we haven't already covered?
- [If you need to identify additional stakeholders for interviews] We are interested in speaking to a number of people from [xxx department], but also some people from outside the department, to give us a 360 degree picture of how the programme has interacted with it.
    - Can I please check with you which of your colleagues also participated in the programme?
    - Who is your line manager / who manages the [xxx unit]
    - Which other units / departments do you work with regularly? Can you suggest anyone in these units / departments I could speak to?

Thank respondent for their time, remind them about any documents they said they could share with you, and ask them if they would mind you getting in touch again if you have any follow-up questions.

Interviewers' reflections on interview *(consider respondent's attitude towards interview / programme; potential issues that may affect how much weight to give claims made by respondent such as motivations, plausibility of claims, inconsistencies in respondent's account):*

7.4   Programme non-participant topic guide

Case study and country	
Interviewee name	
Position and organisation	
Interviewer name	
Date of interview	

Introduction

- We are independent researchers investigating the [xxx] project, which is funded by the UK Department for International Development. We want to hear your thoughts on this project.
- The interview will last about 1 hour.

Consent

- Everything you tell us will be confidential, and your name will not be used in any of our reports. However, we would like to use your thoughts and some anonymised quotes from the interview in our findings, if you are happy with this?
- Do you mind if we audio record the interview? This is for the researchers' reference and will allow us to check that we have we recorded your views correctly.
- Do you have any questions about the research, or concerns you would like to raise before we start?

Aim of the interview

- We'd like to talk to you about the role of evidence in policymaking in [xxx sector] AND / OR [specific outcomes the respondent should have an insight into]. However, this interview might be slightly different to others you may have had in the past. I'm not just interested in what is happening, but also in your ideas about how and why things have changed, or not, over the past few years
- We have some initial ideas but we're not sure if they are correct or not, so we will share these with you during the interview and get your thoughts.

Role and involvement in policy / decision making

- Could you please introduce yourself and your role within the organisation?
- Can I briefly check – how would you describe your role in relation to [or how are you involved in] policy and decision making? Can you summarise that for me please?
- Have you heard about the [xxx] programme?

Description of project (for stakeholders without much knowledge of it)

- The [xxx] programme aims to [encourage the use of evidence in policy and decision making / insert local description here].
- In [xxx country] [insert local description here: e.g. AFIDEP has been leading the SECURE Health programme, providing training and coaching to staff in parliament and MoH, providing technical support to health policy, convening science policy cafes and strengthening policy and research networks.]
- The programme as a whole is funded by DFID and worked across 11 countries in Africa and Asia.
- We are evaluating the programme in order to investigate how effective the programme has been, but also to understand more about how and why different types of approaches can help support evidence-informed policymaking in different contexts.

Political economy analysis discussion [could have this conversation up-front, towards the end, or interspersed throughout the interview]

- I'd like to get your thoughts on the goal of [programme's] work. xxx is trying to promote better use of evidence in policymaking, through building the capacity of civil servants. What are your thoughts on this goal in the [country context]?
- Have you worked or come into contact with the [specific sectors of interest in the case study]? Other than capacity, do you have any insights into the main dynamics that affect evidence use in the [xxx sectors]

Insert questions from the PEA template here. Make sure all of the key questions have been covered, prioritising questions that are difficult to answer through secondary document review.

- Who else do you think we could speak to, to get an insight into the political dynamics in xxx sector? Do you have any documents or sources that we could draw on?

EQ 1: Outcomes and sustainability

- [If respondent has heard about the programme] What is your general impression of the xxx programme?
- Over the past three years, have you noticed any changes in the way [xxx sector / department] thinks about or uses evidence in decision making? What kinds of changes? Can you give me an example?

Insert outcome probes here. Use them to make sure all of the priority outcomes / ToC links prioritised for this respondent have been explored.

- Do you think that the changes have been the same for all [people within the specific stakeholder group – e.g. senior stakeholders]? In what ways have they been different?
  - Probe for examples of people who have been less engaged / have resisted change. Attempt to get names.
- Are there things that **are not changing** in relation to how [xxx sector / department] uses evidence? Or changes that are happening more slowly?

EQ 2: What was BCURE's contribution to observed changes?

For each of the changes mentioned under EQ 1:

- You said that xxx has happened / changed. What do you think caused that change?
- Apart from the BCURE programme, has anything else fed into this?
  - Link back to the initial PEA discussion

EQ 3: How and why did BCURE contribute or not contribute?

For each of the changes and non-changes mentioned under EQ 1:

- You said that xxx has happened and that the programme contributed to that. Why do you think the programme made a difference to xxx? **OR** you said that xxx hasn't changed / the programme didn't contribute. Why do you think the programme didn't influence this?

Insert CIMO tables here. Use them to make sure all the CIMOs prioritised for this respondent have been covered

Policies being directly influenced by evidence

- Have any of these changes you've mentioned led to a specific policy or bill being influenced by evidence?
- Can you tell me a little about this and who was involved please?

Longer-term outcomes/sustainability

You may not need to ask these questions separately or at the end of the discussion. It may be possible to ask about 17 a, c and d in one question, as they are interlinked, or to ask about them under EQ 1 if the opportunity arises. Always ask about Outcome 17a as this applies in all BCURE contexts. Ask about the other outcomes where relevant activities have been undertaken / results observed within the programme

Outcome 17a	You've mentioned x, y, z [in relation to individual behaviour change]. How consistently do you think officials [in xxx department] use evidence when developing

	[policy briefs, concept notes etc.]? How far would you say the [department] is along this journey from evidence use being ad hoc vs being more consistent and routine? __What else needs to happen__ for this to become *sustained* as a routine? What other factors might influence this? E.g. role of external actors on policy analysis and scrutiny (e.g. international donors; lobby groups; civil society groups)?
Outcome 17b *Ask if there have been activities and results relating to consultation process etc.*	You've mentioned x, y, z [in relation to policy development processes.] How far do you think policy processes in [the parliament] in general engage with evidence from different stakeholders and perspectives? __What else needs to happen__ for this to become *sustained* as a routine? E.g. What are the mechanisms for consultation, participation and inclusion in policy processes and the way in which citizens are involved in policy development and monitoring (e.g. referendum, opinion surveys)?
Outcome 17c *Ask if there have been activities and results relating to evidence <u>tools</u>*	You've mentioned x, y, z [in relation to tools]. Are these tools being <u>consistently</u> used by people? Or is the use of the tools more ad hoc? What else needs to happen for this to become *sustained* as a routine? What other factors might influence this?
Outcome 17d *Ask if there have been activities and results relating to evidence <u>processes / standards</u>*	You've mentioned x, y, z [in relation to processes / standards]. Do you think they are helping to promote <u>consistent </u>evidence use by people in [the Ministry / x department]? Are they supported by senior managers? __What else needs to happen__ for this to become *sustained* as a routine? What other factors might influence this?

Wrap up

- Is anything else you think we should know about the programme that we haven't already covered?
- If there anything else you think we need to know, to really understand the role that evidence plays in policymaking in this context?

Thank respondent for their time, remind them about any documents they said they could share with you, and ask them if they would mind you getting in touch again if you have any follow-up questions.

Interviewers' reflections on interview *(consider respondent's attitude towards interview / programme; potential issues that may affect how much weight to give claims made by respondent such as motivations, plausibility of claims, inconsistencies in respondent's account):*

7.5 Programme staff topic guide

Case study and country	
Interviewee name	
Position and organisation	
Interviewer name	
Date of interview	

Introduction

- We'd like to talk to you about your perceptions of the BCURE programme. As you know, we're not just interested in what is happening, but also in your ideas about how and why things have changed, or not, over the past few years. We'd like to share our initial ideas with you during the interview and get your thoughts.
- The interview will last about 1 hour.

Consent

- Everything you tell us will be confidential, and your name will not be used in any of our reports. However, we would like to use your thoughts and some anonymised quotes from the interview in our findings, if you are happy with this?
- Do you mind if we audio record the interview? This is for the researchers' reference and will allow us to check that we have we recorded your views correctly.
- Do you have any questions about the research, or concerns you would like to raise before we start?

Role and involvement in BCURE

- Could you please introduce yourself and your role within the organisation?

Political economy analysis discussion

We are interested in understanding the national and sector-level political dynamics that affect policymaking in [xxx] sector.

- In your opinion, what are the main issues and dynamics that affect evidence use in [xxx sector / Ministry]?

Insert questions from the PEA template here. Make sure all of the key questions have been covered, prioritising questions that are difficult to answer through secondary document review.

- Who else do you think we could speak to, to get an insight into the political dynamics in xxx sector? Do you have any documents or sources that we could draw on?

EQ 1: Outcomes and sustainability

- What do you consider the outcomes of the programme to have been for [each of the main stakeholder groups the programme is working with]? Can you give examples?

Insert relevant outcome probes here, attempting to build on the insights from the programme team workshop / plug gaps.

- Do you think that the changes have been the same for all [people within the specific stakeholder group – e.g. senior stakeholders]? In what ways have they been different?
  - *Probe for examples of people who have been less engaged / have resisted change. Attempt to get names.*
- Are there things that **did not** happen as a result of the programme [that you hoped would], or results that were **more limited** than you hoped?

EQ 2: What was BCURE's contribution to observed changes?

For each of the changes mentioned under EQ 1:

- What role do you think BCURE played in promoting [xxx change]?
- Apart from the BCURE programme, has anything else fed into this?
    - *Link back to the initial PEA discussion*

EQ 3: How and why did BCURE contribute or not contribute?

For each of the changes and non-changes mentioned under EQ 1:

- You said that xxx has happened and that the programme contributed to that. Why do you think the programme made a difference to xxx? **OR** you said that xxx hasn't changed / the programme didn't contribute. Why do you think the programme didn't influence this?

==**Insert CIMO tables here, attempting to build on the insights from the programme team workshop / plug gaps**==

Policies being directly influenced by evidence

- Have any of these changes you've mentioned led to a specific policy or bill being influenced by evidence?
- Can you tell me a little about this and who was involved please?

Longer-term outcomes/sustainability

You may not need to ask these questions separately or at the end of the discussion. It may be possible to ask about 17 a, c and d in one question, as they are interlinked, or to ask about them under EQ 1 if the opportunity arises. Always ask about Outcome 17a as this applies in all BCURE contexts. Ask about the other outcomes where relevant activities have been undertaken / results observed within the programme

Outcome 17a	You've mentioned x, y, z [in relation to individual behaviour change]. How <u>consistently</u> do you think officials [in xxx department] use evidence when developing [policy briefs, concept notes etc.]? How far would you say the [department] are along this journey from evidence use being ad hoc vs being more consistent and routine? <u>What else needs to happen</u> for this to become *sustained* as a routine? What other factors might influence this? E.g. role of external actors on policy analysis and scrutiny (e.g. international donors; lobby groups; civil society groups)?
Outcome 17b *Ask if there have been activities and results relating to consultation process etc.*	You've mentioned x, y, z [in relation to policy development processes.] How far do you think policy processes in [the parliament] in general engage with evidence from different stakeholders and perspectives? <u>What else needs to happen</u> for this to become *sustained* as a routine? E.g. What are the mechanisms for consultation, participation and inclusion in policy processes and the way in which citizens are involved in policy development and monitoring (e.g. referendum, opinion surveys)?
Outcome 17c *Ask if there have been activities and results relating to evidence <u>tools</u>*	You've mentioned x, y, z [in relation to tools]. Are these tools being <u>consistently</u> used by people? Or is the use of the tools more ad hoc? What else needs to happen for this to become *sustained* as a routine? What other factors might influence this?
Outcome 17d *Ask if there have been activities and results*	You've mentioned x, y, z [in relation to processes / standards]. Do you think they are helping to promote <u>consistent </u>evidence use by people in [the Ministry / x department]? Are they supported by senior managers?

relating to evidence processes / standards	<u>What else needs to happen</u> for this to become ***sustained*** as a routine? What other factors might influence this?

Wrap up

- Is anything else you think we should know about the programme that we haven't already covered?
- If there anything else you think we need to know, to really understand the role that evidence plays in policymaking in this context?

Thank respondent for their time, remind them about any documents they said they could share with you, and ask them if they would mind you getting in touch again if you have any follow-up questions.

Interviewers' reflections on interview *(consider respondent's attitude towards interview / programme; potential issues that may affect how much weight to give claims made by respondent such as motivations, plausibility of claims, inconsistencies in respondent's account):*

8   Analysis frameworks

As described in Section 3, analysis frameworks were developed in Excel to guide a systematic and transparent analysis of the evidence against the EQs, at country case study and synthesis level.

Country case study analysis database template

#	Background details					EQ 1: what happened, for whom?				EQ 2: BCURE contribution			EQ 3: How and why (CIMOs)						Other
	Interview respondent/ document name	Reflections on potential bias / position in relation to programme	Stakeholder type	Relevant intervention	Organisation	Outcome no.	Outcome	Evidence	Reflection on strength of evidence	Detail of BCURE contribution	Details of other factors	Reflection on strength of evidence / bias	EQ 3 hypothesis (CIMO)	Explicitly tested or inferred?	CIMO confirmed, rejected, revised, or new CIMO suggested?	Evidence for intervention factors	Evidence for contextual factors	Mechanism	Any other thoughts or comments

Overview report synthesis database template

Background details			EQ 1: what happened, for whom?				EQ 2: BCURE contribution			EQ 3 - How and why did BCURE contribute? (CIMOs)							Other	
Case	Setting / impact pathway	Level *individual, organisational, institutional?*	Outcome	Extent outcome achieved	Evidence for outcome (narrative and quotes)	Reflection on strength of evidence	EQ 2 - BCURE contribution - details (narrative and quotes)	Details of other factors, factors)	Reflection on strength of evidence / bias	EQ 3 hypothesis (CIMO)	Evidence for intervention factors	Evidence for contextual factors	Mechanism	CIMO confirmed, rejected, revised, or new CIMO suggested?	Revised CIMO (reflecting what actually happened at Stage 3)	Reflection on strength of evidence	Comments	Lessons

9 Additional analysis tables

Detailed summary of results from BCURE training approaches

	Training approach and scale	**Improvements in individuals' capacity** *Kirkpatrick Level 1*	**Changes in individuals' behaviour** *Kirkpatrick Level 2*
Pakistan *Evidence from Stage 3*	Large scale EIPM training course (1780 participants), limited targeting, no follow-up	Strong evidence through pre- and post-training tests of significant gains in technical knowledge. Participants and trainers felt that training quality was enhanced by the use of contextually relevant case studies, and felt some of the practical tools like cost-benefit analysis were particularly useful, although some participants at Stage 3 complained that the sessions were quite academic and not particularly interactive	Limited evidence of widespread behaviour change, although the evaluation was only able to interview a small proportion of the whole cohort. Where trainees had applied their learning, the training had generally helped trainees frame their thinking on how to use evidence or data to address a specific task in the workplace (M), and gave them tools (e.g. cost-benefit analysis) to facilitate specific types of analysis (M) in a context where they were tasked with working on an issue that required data (C), where senior managements were supportive and actively encouraged them (C), and where they were able to draw on resources to access and analyse evidence in response to a specific problem (C). This was enabled by participants' existing skills, motivation and experience, including soft skills that enabled them to present analysis to senior members of the government to influence reform (C). Trainees who hadn't applied their learning often did not perceive the training as relevant to their current professional role, often because they were not involved in policy formulation (C). This was a consequence of the training being delivered to whole cadres of staff as part of mandatory requirements for promotion rather than based on a decision about relevance and need (I). Interviews also suggested missing incentives in the workplace to change practices towards more evidence-informed policymaking, including uninterested senior managers themselves lacking an incentive to consider evidence (C), against a backdrop of endemic corruption providing a motive to ignore or suppress evidence (C)
Bangladesh *Evidence from Stage 3*	Large scale EIPM training course (400 participants), for civil servants involved in policy formulation in three pilot ministries, with content aligned to EIPM guidelines developed at Cabinet Division Level. On-the-job follow-up support to a small number of trainees, to apply guidelines in policy pilots	Strong evidence of increased knowledge and understanding of EIPM, through pre- and post-training tests corroborated by interviews Participants felt the training was useful and high quality because it provided a systematic procedure for policymaking, it was practical and incorporated local case studies, and the trainers were local experts with relevant experience	Strong evidence that training led to new or improved knowledge, skills and confidence and improved trainees' self-efficacy (M), leading to changes in the way evidence was considered in policy formulation (O), where trainees were supported to apply skills through policy pilots after the training (I). A key factor was the fact that the EIPM guidelines were seen as providing a helpful structure to facilitate policy formulation (M) in a context where such guidance was lacking (C). Limited evidence (as yet) that trainees who did not receive follow-up support had had an opportunity to apply skills (C) – although the top down EIPM guidelines are likely to facilitate skills application (M) if they are adopted by line ministries and trainees are requested to use them (M). However, some trainees were not involved in policy formulation roles, suggesting there may be limited opportunities to apply learning in future (C). Several stakeholders suggested that without follow-up (for example refresher training, or permanent EIPM focal points within ministries) (I) there is a risk that trainees will forget what they learned or will lack the confidence to apply their learning (O)
Zimbabwe	Small scale EIPM training course (49 participants) delivered to	Strong evidence (though training follow-up survey and interviews with	Strong evidence for sustained change in the work of trainees in the Ministry of Youth as a result of BCURE. Training generated self-efficacy (M) and contributed to sustained behaviour change (O) because it helped trainees perform in their new

	Training approach and scale	Improvements in individuals' capacity *Kirkpatrick Level 1*	Changes in individuals' behaviour *Kirkpatrick Level 2*
Evidence from Stage 3	technical staff in targeted institutions, followed up by technical support to implement organisational reforms and a Parliamentary exchange programme	significant number of participants) that trainees had gained new knowledge and skills (although less so for participants outside the research department in parliament). Participants in the Ministry of Youth felt the training was high quality because it was practical, hands on and participatory, and because it imparted soft skills as well as technical skills	roles as research officers in a newly-established research unit (C), as most did not have background in research (C), and given the resource-constrained context where other training was unavailable (C). Skills application was supported by the two Directors in the unit, who were interested in evidence and supportive of trainees applying their skills (C). However, opportunities to apply skills in research work were limited by the small scale of the research unit – which is shrinking due to staff rationalisation (C), and the fact that it interacts with only part of the Ministry (C), and that officers are often engaged in administrative rather than research work (C). Limited evidence that the training or exchange programmes made a significant contribution to changes in practice in parliament (O). The training content was insufficiently tailored to the parliament-specific needs of staff (I) (particularly those outside the research department), and therefore was not relevant to trainees' needs (C). The plethora of other training and exchange opportunities available in parliament may also explain why some participants felt the training or exchange visits did not offer much that was new (C)
Kenya *Evidence from Stage 3*	Small scale EIPM training course (45 participants), delivered to technical staff in targeted institutions, followed up by mentoring support and an overseas secondment	Strong evidence (through training follow-up survey and interviews with significant number of participants) of increases in knowledge and skills. Participants felt the training was useful and high quality because it combined theory with practical application and provided the space to work on a live policy topic, and the facilitators were high calibre, knowledgeable, patient, skilled and committed; although some felt that course had been insufficiently tailored for Parliamentary staff	Strong evidence (from triangulated interviews with participants and managers) that substantial numbers of trainees in both parliament and the MoH had been able to use learning in their work, and this behaviour change had been sustained up until the final evaluation. In parliament, training succeeded in building self-efficacy (M) which resulted in improved use of evidence (O) in a context where training was delivered to a newly-recruited researchers during their induction periods (I), helping them quickly meet the specific demands of their jobs (C). Skills application was supported by follow-up support from BCURE to produce concrete evidence products, tools and templates (I), which facilitated (M) trainees to more efficiently meet the high volume of evidence products required of them (O), reinforced by senior managers in the unit (M), who were already proactively engaged in an evidence agenda (C), by providing hands on support and feedback and encouraged trainees to improve the quality of their work (C), which in turn generates recognition and career rewards (C) that increase motivation for evidence use (M).' In the Ministry of Health, training had also succeeded in building self-efficacy (M), and the searching, synthesising and presentation approaches and tools provided by BCURE helped facilitate (M) trainees to present evidence more effectively. These contributed to improved evidence use (O) particularly among trainees with opportunities to apply evidence in the development of specific policies, standards and guidelines (C). The most sustained gains in individual capacities and ongoing use of evidence (O) was among motivated individuals (C) based in divisions and units with well-resourced donor programmes that offer trainees opportunities to apply and further develop EIPM skills, which in turn generates recognition and career rewards (C) that increase motivation for evidence use (M). Evidence use was less sustained in divisions where officials take a more administrative role in policy development and there are fewer opportunities to apply skills (C)
Sierra Leone	Medium scale training (964 training days),	Strong evidence through Stage 2 interviews that	Some evidence at Stage 2 (from interviews with participants and programme staff) that line ministry Cabinet Focal Persons trained to support line ministries to apply new cabinet

	Training approach and scale	Improvements in individuals' capacity *Kirkpatrick Level 1*	Changes in individuals' behaviour *Kirkpatrick Level 2*
Evidence from Stage 2	delivered to technical staff in line ministries to help support implementation of organisational reforms to cabinet procedures. Training also delivered to cabinet staff to help them support implementation	training and regular meetings had helped civil servants understand the new procedures established in the new cabinet manual. Participants appreciated the participatory nature of the training, the use of practical case studies and the opportunity to learn and share from colleagues	Procedures were able to perform their new roles to some extent, but that there was a 'long way to go.' Training, combined with ongoing support from the new (BCURE supported) cabinet research unit (I) helped increased cabinet focal person self-efficacy (M) to perform the duties of their new role and support the implementation of the procedures (O). The presence of the Cabinet Secretary in the training helped ensure full participation, especially of senior civil servant staff (I). However, the need for support from other ministry staff for CFPs to perform their role (C), potentially undermines their ability to apply their learning
South Africa *Evidence from Stage 2*	Small scale workshops, aiming to provide an introduction to EIPM to participants who might become mentees, raise awareness about EIPM and its potential value, and provide spaces for dialogue	Some evidence through Stage 2 interviews that workshops had introduced participants to relevant terminology and methods and reinforced their understanding of the importance of evidence. Participants appreciated the opportunities to share challenges and solutions in the workshops, although there was limited time for practical skills application	Limited evidence of instrumental changes in practice as a result of the workshops. Where there was strong prior interest and enthusiasm for EIPM (C), and where it was relevant to existing work (C), this seemed to enable workshop attendees to connect to the concepts and understand the immediate usefulness to their work (M), assisting in conceptualising their work and offering potential solutions to work challenges (O) Where prior interest and understanding were absent (C), the workshops increased interest and awareness in EIPM and the workshop content was regarded as potentially useful (M), but participants were not actively applying the concepts and methods (O)

10 Communications framework

This section summarises the evaluation communications framework, developed during the inception phase and updated during the mid-point of the evaluation.

Summary

The primary aim of the BCURE evaluation is to strengthen the global evidence base on whether capacity-building approaches to support evidence-informed policy can be a cost effective way to reduce poverty, and if so how they can be implemented to achieve the greatest impact. This framework describes how communications will support the evaluation in achieving this outcome.

The role of the BCURE evaluation communication function is to carry out activities that not only raise awareness of the evaluation with target audiences but also help best position the learning and findings for uptake. By ensuring that lessons learned from the evaluation on what works and doesn't are strategically shaped and shared, we hope that funders, designers and implementers of evidence-informed policymaking (EIPM) initiatives can make better choices when it comes to supporting similar initiatives.

The communications function is also responsible for sharing what is understood about the effectiveness of the BCURE programme with DFID and its implementing partners.

The BCURE evaluation communication strategy contains the following objectives:

1. To communicate where and how, and in what circumstances, decision makers can better access, use and understand evidence
2. To provide support and assistance to the BCURE programmes on the most effective ways of communicating evaluation findings to partners and key audiences in the countries in which they work
3. To reach, engage and inform the emerging Community of Practice around EIPM about how and why capacity-building for evidence use is important and effective in improving development outcomes.

This **communications framework** supports these objectives by identifying and analysing the evaluation's target audiences, identifying the opportunities and spaces for engagement and planning the specific activities and channels that will be used.

The latest **situation analysis** conducted as part of this framework highlights a number of difficulties in identifying both 'new' audiences and the spaces for sharing the evaluation findings. For example, although the EIPM 'community' is a key entry point for engagement, it is by and large led by 'supply-side' actors such as researchers, donors and knowledge brokers rather than important 'demand-side' actors such as the high and mid-level government policymakers who are at the heart of BCURE CToC. The analysis also finds that while there are a number of health- and development research-focused EPIM and capacity development networks and initiatives, there are relatively few that focus specifically on governance and public sector reform in developing countries– a key field for the uptake of the evaluation findings.

Since many of the EIPM platforms and spaces for engagement are largely driven by a core number of actors (including BCURE implementing partners INASP and AFIDEP), we conclude that the most effective way to reach the majority of our stakeholders is by utilising existing channels (such as knowledge platforms and networks) and through the BCURE implementing partners and DFID's Evidence into Action team who we have identified as **'amplifiers'** as well as recipients of the evaluation findings.
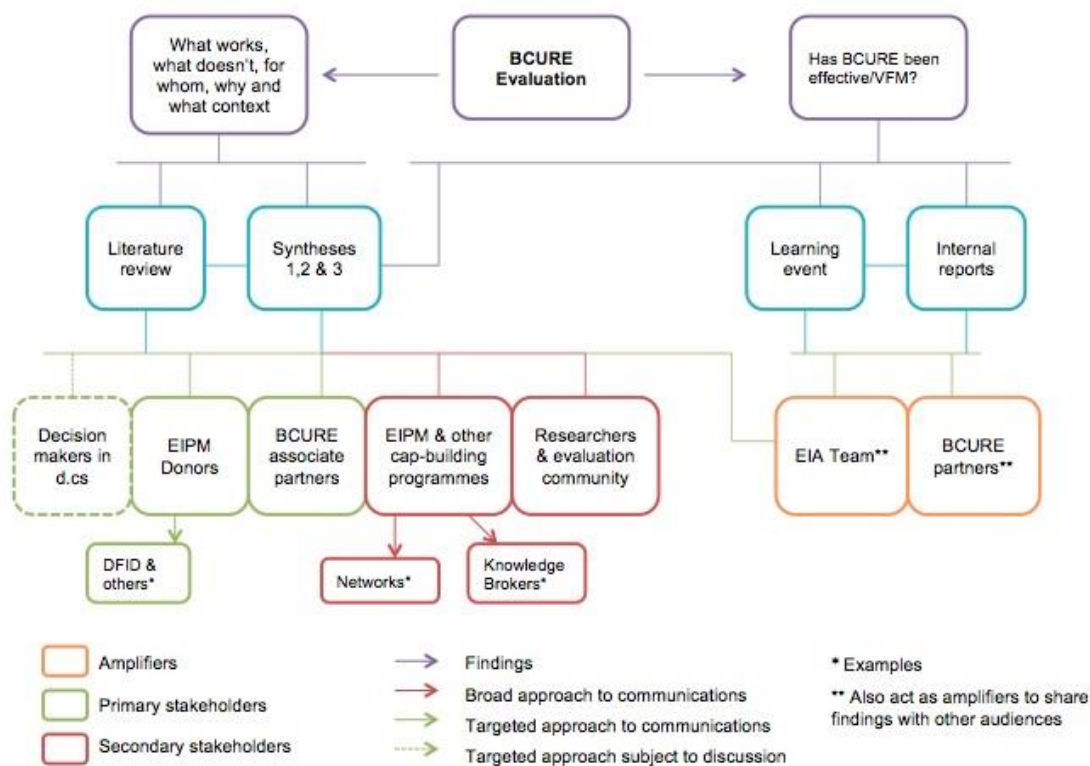
The latest round of **stakeholder identification and analysis** has begun to fill the gaps in our knowledge of key target groups, such as governance programmes, additional EIPM donors and their portfolios. In the process we have found that while the BCURE evaluation is well equipped to share top-line messages with a general set of audiences, a more nuanced understanding of our primary target stakeholders is needed so that findings can be packaged and channelled in ways that are appropriate to their needs.

To address these issues, we propose taking a **two-pronged approach** to communicate with three tiers of target audiences (Figure 1). To maximise effectiveness, we suggest implementing our outreach in close collaboration with a group of **amplifiers**. This will be a group of selected individuals who will help create demand for findings from the evaluation among primary and secondary target audiences by profiling the evaluation findings and encouraging debate on the findings.

Our most intensive engagement will be with a set of **primary target audiences** who require a more **differentiated approach**. For these, we plan to use a range of well-tailored communications products and channels to regularly communicate the evaluation findings including via our amplifier intermediaries. For **secondary audiences**, our engagement will take a much **broader approach** by utilising multiple communication channels (including intermediaries such as networks and knowledge platforms) but less frequently and with less tailoring to meet specific informational needs.

Figure 1 illustrates our **two-pronged approach** to BCURE communications that focuses on both broad and targeted approaches according to the different stakeholder groups. The following sections then discuss aspects of the communications framework in more detail.

Figure 5: BCURE evaluation communications approach

Situation analysis for BCURE evaluation communications

This area of work seeks to establish:

1. The resources and communication channels the BCURE evaluation has for capturing learning and undertaking dissemination during the programme
2. The likely sources of information and opportunities to share learning
3. Dissemination opportunities (audiences/spaces) during the life of the programme relating to the evaluation (e.g. online platforms, conferences, online discussions)

The situation analysis for this framework builds on work already done as part of the communications strategy and looks at the challenges and opportunities presented to BCURE evaluation communications specifically rather than to the programme per se.

Information in relation to this was gathered during the inception period and supplemented with further desk-based research for the purposes of this document. Sources used include BCURE and BCURE evaluation documentation (including the common theory of change); a review of external barriers and opportunities (see communications strategy), the stakeholder analysis to date (also see communications strategy; websites of organisations and networks working in EIPM; and records of discussions with members of the BCURE evaluation team and DFID. A full list of channels and opportunities is available on request.

Table 6: Challenges and opportunities for BCURE evaluation communications

Challenge	Opportunity
The concept of EIPM in target countries is relatively new and therefore relatively low profile	Use amplifiers to help create demand for findings from the evaluation among primary and secondary target audiences by profiling the evaluation findings and encouraging debate on the findings. For example, the evaluation senior adviser Professor Fanie Cloete, at Stellenbosch who has a high profile in EIDM/PM in Africa. Spot opportunities beyond these spaces for sharing evaluation findings e.g. DFID Governance Advisers retreat, What Works Global Summit etc. Keep informed of relevant debates and discussions on EIPM by plugging into and engaging with knowledge platforms and discussion fora such EPBDN, the Pelican Initiative, Policy & Ideas, Knowledge Brokers Forum and the BCURE DGroups.
There are few EIPM initiatives tackling the issue beyond health and research uptake	Broaden outreach by identifying programmes and networks in which capacity building towards EIPM is a component. Engage with stakeholders moving into the government space through research-policy networks such as the UKCDS Research Strengthening Group and the Think Tank Initiative. Encourage DFID to showcase BCURE evaluation as an example of accountability and transparency, and building effective government institutions.
Capacity development networks are often generic and too broad in terms of areas of interest	Target networks that focus on building government and public sector effectiveness, where there is potential to position EIPM as relevant to public sector reform. For example, LenCD's 'Effective Institutions' working group that aims to prepare members' governments for meeting the Busan Aid Effectiveness targets; the capacity-building section of the GSDRC website; and accountability and transparency initiatives such as the international Open Government Partnership.
There are a number of assumptions underpinning the concept of 'evidence-informed policymaking'	Use existing EIPM spaces and places to convene discussions and showcase new findings such as the Alliance for Useful Evidence, Capacity Alliance and Research 2 Action. Raise awareness of the BCURE literature review insights and learning with a wide range of audiences to stimulate interest and debate. Share internal learning through external blogs and professional networking sites. For example, blogging on LEN CD, R2A, better evaluation etc.
BCURE evaluation has to build reputation/credibility with target audiences	Develop language across all communications that positions the BCURE evaluation as a unique opportunity to gain insight into the effectiveness of EIPM capacity-building initiatives. Share information about BCURE evaluation findings through the communication channels of key partners to mutually shared stakeholder groups. Utilise entry points into key EIPM initiatives via reputation and convening power of amplifiers e.g. INASP and DFID.

Research uptake by decision makers is often regarded as a simple, linear process	Interact with the growing body of supply-side organisations and networks raising awareness of the complexity of EIPM such as UKCDS. Promote the broader view of how research findings make their way into government decision making. For example the literature review briefing
Communication activities and ambition are restricted by modest budget allocation	Use low and no-cost online communication tools to share the findings including Itad website, Twitter, Tumblr, Mailchimp etc. Look for efficient ways in which harnessing capacity of BCURE evaluation team and partners to roll out certain tasks.

Stakeholder mapping and analysis

The BCURE evaluation will only realise its intended value if the findings from the programme are effectively communicated to identified audiences who then act on the new knowledge. The aim of our work in this area has been to establish the needs, interest and contexts of the stakeholders of the communication activities (e.g. the BCURE team, BCURE implementing partners DFID, other donors and organisations with an interest in EIPM).

We have identified the following three broad audience categories. Depending to their role, some stakeholders fall under more than one category.

Amplifiers

The role of amplifiers will be to:

1. Assist in creating demand for findings from the evaluation among primary and secondary target audiences by profiling the evaluation findings and what they will offer;
2. Amplify findings from the evaluation among primary and secondary target audiences by profiling communications products and events and encouraging debate on the findings.

The role of these amplifiers is critical to the outreach of the evaluation findings for a number of reasons. Firstly, it will open up windows and spaces for the communication of findings that the evaluation team are either not aware of or will struggle to reach, so extending the reach of the communications strategy. Secondly, it would ensure that communications products are well tailored to stakeholders, maximising their utility and therefore value for money.

Activities might include:

- BCURE programme implementers convening meetings to present and discuss the findings with government partners as part of their sustainability planning.
- DFID facilitating discussions with DFID country offices where BCURE interventions happen and in which the evaluation case studies are carried out.
- BCURE programme implementers encouraging partners to sign up to the BCURE evaluation newsletter.
- DFID engaging with DFID country offices where there is no BCURE programme but where there is interest in demand-side work e.g. Nepal.
- A BCURE evaluation presentation at the BCURE learning event to explore this role in greater detail.
- BCURE programme implementers cross-posting evaluation blogs on their own websites.

Regular communication with this set of stakeholders is essential. Amplifiers will be reached via BCURE DGroups and will be kept up-to-date on the work of the evaluation via a newsletter three times a year. The literature review and synthesis briefings will act as a tool for the amplifiers to take forward their own informing and influencing. Each time an evaluation product is published, amplifiers will receive a 'What's new?' e-alert with a specific call to action depending on the nature of the product. For example, these 'calls' could include asking them to share the case studies with their networks, reflect on their own learning as part

of their project blog or to host on- or off-line discussions with their own partners that explore the implications of the findings.

Primary stakeholders

The **primary target audiences** are BCURE implementing partners, BCURE implementing associates, EIPM donors and funders and government ministries and organisations. As such, they are perceived to be the most receptive to the evaluation findings and among the most influential in terms of using these in their own policy and programming. Our primary stakeholders include stakeholders who are harder to reach but whose uptake of the findings would make a significant contribution to the theory of change.

We would like these stakeholders to respond by:

▪ Formally registering an expression of interest. For example, emailing with questions on specific issues

▪ Sending invitations to the evaluation team to come and give a detailed presentation of the learning and findings

▪ Extending the reach of the evaluation findings to their colleagues and networks. For example, by incorporating the briefings into their EIPM resources. E.g The YakoViko Evidence-Informed Policymaking Toolkit and the LenCD learning package.

These stakeholders require a deeper level of understanding and engagement through a more targeted approach via individuals i.e. our selected amplifiers and specific programmes or portfolios of work. A mix of written products (e.g. the literature review and synthesis briefings) and digital communications (e.g. the evaluation newsletter, blogs and Twitter) will be used to communicate and position the evaluations findings for uptake. We will ask amplifiers to share these and facilitate virtual and face-to-face discussions to explore them further. For example, the DFID BCURE team might hold a discussion with country offices via Yammer or host sessions at Advisers' Professional Development Conferences.

Secondary stakeholders

Secondary target audiences are wide-ranging and easy to reach en masse via knowledge brokers and platforms. They include:

▪ EIPM project implementers and networks: these are a diverse set of organisations but with a common focus. Priority will be given to EIPM programmes that promote governance and accountability as well as health and research uptake. As far as possible, we will use existing EIPM fora such as WHO's EVIPNET, Health Information for All, the Alliance for Useful Evidence and capacity development networks such as LEN CD and Capacity Alliance Feeds for the communication of the evaluation findings.

▪ Other development partners who actively support EIPM more broadly: these include governance programmes in which EIPM is a component including health promotion, environmental governance, voice and accountability and governance reform. This is a disparate group of stakeholders. To reach them, the evaluation team will rely heavily on the amplifiers who engage directly with specific organisations.

▪ Research organisations, programmes and think tanks involved in consortia responsible for delivering research on EIPM and capacity development. While their influence on policy and programming is important, this influence tends to be indirect and makes itself felt over a medium-term timeframe.

▪ Evaluation community: this includes evaluation focused organisations and evaluation specialists. Evaluation focused organisations working with DFID will be targeted through our social media work. More direct engagement with the evaluation community will be done through focal points within the evaluation community such as the European Evaluation Society and What Works Global Summit.

We expect the responses of secondary stakeholders to be 'light touch' and could include:

- ▪ Signing up to the BCURE evaluation newsletter
- ▪ Following the BCURE evaluation on social media and sharing posts/tweets with followers
- ▪ Citing the evaluation findings in literature reviews and policy briefings

Our outreach to secondary stakeholders target will focus on communicating the breadth of the findings emerging from the evaluation and encouraging discussion about these. Activities and channels will include conducting a social media campaign via Twitter, posting regular blogs, and using other knowledge platforms such as the LenCD library and the Africa Evidence Network database to raise awareness about the evaluation and draw target audiences to the evaluations' portal, hosted on Itad's website. The evaluation methodology and results will be presented in at least three public discussions including the European Evaluation Society Annual Conference, the UKCDS Research Capacity Strengthening Group and at the Centre for Development Impact.

Target audiences

Target audience Categories	Specific stakeholders
Amplifiers	DFID Evidence into Action Team BCURE implementing partners BCURE evaluation team
Primary audience	BCURE implementing partners BCURE implementing partner associates EIPM donors and funders such as • Multilateral organisations: World Bank, NEPAD, UNDP and WHO • Bilateral organisations: SIDA, IDRC, USAID, DFID, DSIG Netherlands • Philanthropics: Hewlett Foundation, Wellcome Foundation, Gates Foundation and Open Society Institute Government ministries and organisations such as the Performance Monitoring and Evaluation Department of the South African Government. Participants (interview respondents) in the BCURE evaluation
Secondary audience	EIPM project implementers, such as the Knowledge Sector Initiative (Indonesia), Supporting the Use of Research Evidence (SURE) (WHO Worldwide), NEPAD Capacity Development Programme and DFID Nepal's Evidence for Development and NICE International. Capacity development networks such as LEN CD, Capacity4Dev.eu, APDEV, Africa Cabinet Government Network and ACBF. Research capacity strengthening organisations and networks such as UKCDS, EBPDM, iDSI and the Alliance for Useful Evidence. Other development programmes that focus on EIPM for example, the Public Sector Accountability and Governance programme, ESPINN and PATHS2 in Nigeria, ESP Nepal, FLEGT and ACT in Tanzania. Research organisations, programmes and think tanks including 3ie, the Centre for Development Impact (IDS), Overseas Development Institute, Institute for Government, the Alliance Health Policy Systems Research and the Centre for Evidence and Social Innovation. Evaluation community e.g. UK, European, African and Asian Evaluation Society Members

Summary of BCURE evaluation communications activities

- • Development of a knowledge page on the Itad website, containing evaluation outputs http://www.itad.com/knowledge-and-resources/bcure/
- • Publication of the literature review and Stage 1 and 2 synthesis reports along with briefing notes and blogs on the Itad website, and dissemination to key audiences listed above.
- • Academic publication: Punton, M., Vogel, I. and Lloyd, R. (2016b). Reflections from a Realist Evaluation in Progress: Scaling Ladders and Stitching Theory. CDI Practice Paper, 18.

- Presentations at the European Evaluation Society Conference 2016, the What Works Conference 2016, the American Evaluation Conference 2017, and the UK Realist Evaluation Conference 2015 and 2016.
- Presentations and practical sessions with implementing partners at the BCURE Learning Events 2014, 2015 and 2016.

Following the completion of the final evaluation report, the following activities are planned:

- Full design of synthesis report to maximise readability
- Blogs on the Itad website, where possible cross-posted to reach further audiences in Table 2
- Face-to-face presentations with DFID staff
- Targeted dissemination of synthesis report with primary and secondary stakeholders in Table 2
- Follow-up calls and webinar for BCURE implementing partners
- Presentation of findings at key conferences in 2018

11 RAMESES standards for realist evaluation

In 2016, a set of reporting standards were developed for realist evaluations as part of the RAMESES II Project.[13] 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 of the main report
4. Programme theory	Describe the initial programme theory (or theories) that underpin the programme, policy or initiative	Annex 4 details the programme theory and how it has evolved over time
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 Annex 3.1
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 3.10
Methods		
7. Rationale for using realist evaluation	Explain why a realist evaluation approach was chosen and (if relevant) adapted	See Annex 3.2
8. Environment surrounding the evaluation	Describe the environment in which the evaluation took place	See Section 3.1 of the main report
9. Describe programme policy, initiative or	Provide relevant details on the programme, policy or initiative evaluated	See Section 3.1 of the main report

No.	Standard	Relevant section of report or annexes
product evaluated		
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 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 Annex 3.4
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 Annex 3.4 and Annex 8
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 Annex 3.4, 3.7 and 4.
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 Annex 3.4
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 Sections 5-7 of the report, with further detail in Annex 4
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 the report in tables and summary boxes, and in the overall conclusions
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 2.4 of the main report
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 the report in 'insights from the literature' boxes
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 conclusions

No.	Standard	Relevant section of report or annexes
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 3.10

12. References

Befani, B. and Mayne, J. (2014). Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation. *IDS Bulletin, 45(*6): 17–36. Link.

DFID. (2009). Political Economy Analysis: How to Note. A DFID Practice Paper. Link.

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 Policymaking. Working Paper 21. Cape Town: Institute for Poverty, Land and Agrarian Studies. Link.

Funnell, S.C. and Rogers, P.J. (2011). Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. San Francisco, CA: Jossey-Bass

Hudson, D., Marquette, H. and Waldock, S. (2016). Everyday Political Analysis, Developmental Leadership Programme, University of Birmingham. Link.

IDRC. (2014). Lasting Change – Strengthening Capacity to Improve Health Systems. Ottawa: IDRC. Link.

Lemire, S.T., Nielsen, S.B. and Dybdal, L. (2012). Making contribution analysis work: A practical framework for handling influencing factors and alternative explanations. *Evaluation, 18*(3).

Mayne, J. (2008). Contribution analysis: An approach to exploring cause and effect. International Learning and Change (ILAC) Brief, 16.

Mayne, J. (2011). Addressing cause and effect in simple and complex settings through contribution analysis. In R. Schwartz, K. Forss and M. Marra (eds), *Evaluating the Complex*. Piscataway, NJ: Transaction.

Mayne, J. (2012). Contribution analysis: Coming of age? Evaluation, 18(3): 270–80.

Michaelis, C., & Westhorp, G. (2016). International Climate Fund Evaluation Evidence Framework, Final Draft.

Moncrieffe, J. and Luttrell, C. (2005). An Analytical Framework for Understanding the Political Economy of Sectors and Policy Arenas. Overseas Development Institute, 1–36.

Noblit, G. and Hare, R. (1988). *Meta-Ethnography: Synthesizing Qualitative Studies*. London: Sage.

Nutley, S. M., Davies, H. and 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. (2006). Evidence-based Policy: A Realist Perspective. London: SAGE Publications.

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

Pawson, R. and Tilley, N. (2004). *Realist Evaluation*. Link.

Poole, A. (2011) 'Political Economy Assessments at Sector and Project Levels', How-To Note. Washington DC: World Bank. Link.

Pope, C., Mays, N. and Popay, J. (2007). Synthesising Qualitative And Quantitative Health Evidence: A Guide To Methods. McGraw-Hill International.

Punton, M., Brown, C., Echt, L., Fisher, C., Lloyd, R., Morton, S. and Vogel, I. (2016a). How Can Capacity Development Promote Evidence-Informed Policymaking? Literature Review for the BCURE programme. Brighton. Link.

Punton, M., Vogel, I. and Lloyd, R. (2016b). Reflections from a realist evaluation in progress: scaling ladders and stitching theory. *CDI Practice Paper*, (18). Link.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). Broadening the Range of Designs and Methods for Impact Evaluations. DFID Working Paper, 38(April 2012).

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

Waldman, T. (2014). The Use of Statebuilding Research in Fragile Contexts: Evidence from British Policymaking in Afghanistan, Nepal and Sierra Leone. Journal of Intervention and Statebuilding, 8(2–3): 149–72. Link.

Westhorp, G. (2014). Realist Impact Evaluation: An Introduction. Methods Lab. London: ODI. Link.

Wong, G., Westhorp, G., Manzano, A., Greenhalgh, J., Jagosh, J. and Greenhalgh, T. (2016). RAMESES II reporting standards for realist evaluations. *BMC Medicine, 14*(1), 96. Link.

Wimbush, E., Montague, S. and Mulherin, T. (2012). Applications of contribution analysis to outcome planning and impact evaluation. Evaluation, 18(3), 310–29. Link.

Wong, G., Westhorp, G., Pawson, R. and Greenhalgh, T. (2013). *Realist Synthesis: Rameses Training Materials* (misc). London: Health Research Health Services and Delivery Research Program. Link.

Wong, G., Westhorp, G., Manzano, A., Greenhalgh, J., Jagosh, J. and Greenhalgh, T. (2016). RAMESES II Reporting Standards for Realist Evaluations. *BMC Medicine, 14*(1), 96. Link.

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