

DFID PAKISTAN EVALUATION STRATEGY



Image credit: Vicki Francis/Department for International Development

July 2011

CONTENTS

- I. Summary**
- II. The Pakistan Context**
- III. Objective of the Evaluation Strategy**
- IV. Deciding which Programmes to Evaluate**
- V. Strengthening DFID Pakistan Evaluation Systems**
- VI. Financing Evaluations**
- VII. Building DFID Pakistan evaluation capacity**
- VIII. Working with Partners**
- IX. Communicating Results**

DFID Pakistan Evaluation Strategy

I. SUMMARY

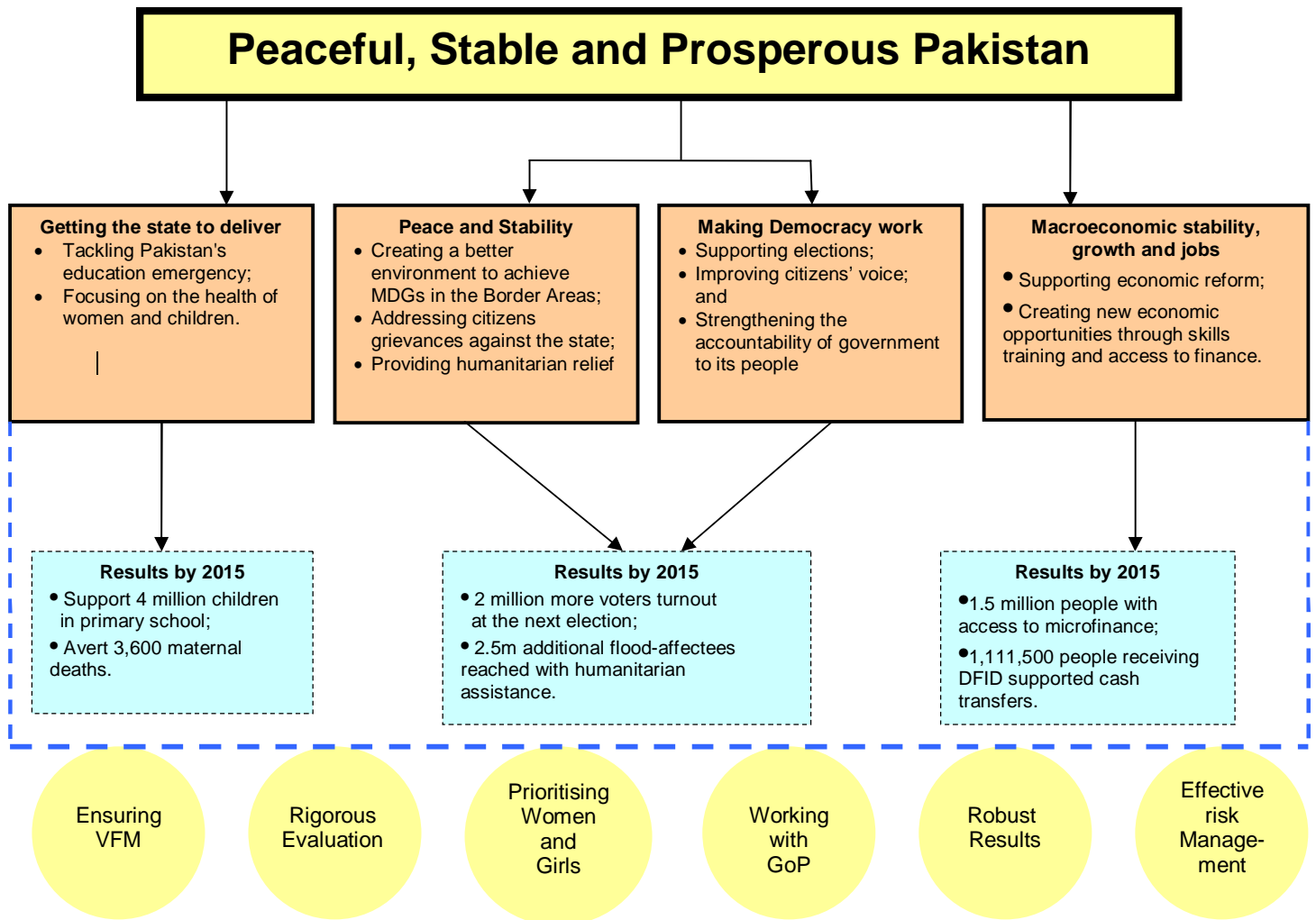
Pakistan is off-track on most of the Millennium Development Goals and, following the bilateral aid review, could become the largest recipient of UK aid by 2015. Reflecting the complexity and importance of the country, the potential scale of investment as well as the level of innovation in the DFID Pakistan (DFID P) programme pipeline, we will:

- Set the ambitious goal of evaluating 50% of our programmes, covering 75% of our spend. All of these evaluations will be started by 2015, with several completed by then;
- Use the criteria set out in this Strategy to determine which programmes to evaluate;
- Integrate the principles of gender sensitivity, statebuilding/peacebuilding and making markets work for the poor in the design of these evaluations;
- Begin three evaluations in 2011 of large DFIDP investments: the Punjab Economic Opportunities Programme, the Improve Citizens Engagement through Devolution programme (ICED) and the Floods response programme. This implements the evaluation commitment set out in the DFIDP Operational Plan;
- Embed evaluation more widely across the office and at every stage of the project cycle: from design to completion;
- Share the lessons learnt from these evaluations across the global DFID network, government counterparts and donor partners. We will use these lessons to inform future decisions, including on the DFIDP programme pipeline;
- Build evaluation capacity across the DFIDP team, with a particular focus on developing Staff-Appointed In-Country (SAIC) skills; and
- Adhere to Development Assistance Committee (DAC) standards across all of our work on evaluation.

II. THE PAKISTAN CONTEXT

1. At least 36 million Pakistanis, one in four, live in poverty. Pakistan is off-track on many of the Millennium Development Goals. Half of all adults, and two out of every three women, are illiterate. One in ten children die before their fifth birthday and 14,000 mothers die during pregnancy. This entrenched poverty leads to suffering, lost opportunity and a sense of grievance; all of which undermine Pakistan's long-term stability and prosperity. Pakistan is also critical to regional and global security. And it retains long, deep and close links to the UK.
2. For all of these reasons, Pakistan is a priority country for the UK and could become the UK's largest aid programme by 2015. As set out in the DFID Operational Plan (<http://www.dfid.gov.uk/pakistan>), and Figure 1 below, the programme will focus on four areas:
 - Peace and Stability;
 - Making Democracy Work;
 - Macroeconomic stability, growth and jobs; and
 - Getting the State to Deliver.
3. Delivering tangible results in these areas, as well as wider policy reform, in a country as complex, large and important as Pakistan will be challenging. That places a premium on robust evaluation; clarity about what we are trying to achieve and how, drawing on the best possible evidence to inform design; understanding what works and what doesn't and being prepared to innovate and take risks to build new evidence of what can be done.
4. The majority of the programmes that will deliver through to 2015 are new and currently under design. That provides a unique window of opportunity to embed evaluation into them *from the beginning* and to inform the DFID programme pipeline ahead of scale up in 2013/14.
5. In improving the impact of the DFID programme, a robust and ambitious approach to evaluation will also give increased confidence to the UK public over the aid programme. The new Independent Commission on Aid Impact (ICAI), reporting directly to the International Development Committee (IDC), will potentially evaluate Pakistan in the third year of its work plan. The IDC is likely to undertake its own enquiry into Pakistan later this year. In addition, public and Parliamentary scrutiny of the programme will remain high.

Figure 1



III. OBJECTIVE OF THE EVALUATION STRATEGY

- The objective of this strategy is, ultimately, to improve the impact of the DFIDP programme by ensuring our investments and decisions are based on the best possible evidence of what works. This will strengthen confidence in the UK aid programme while also helping to progress the global debate on international development.
- This strategy therefore sets out how DFIDP will put evaluation at the heart of its portfolio development, by:
 - Setting an ambitious goal for the number of evaluations we will undertake;
 - Establishing clear criteria to determine which programmes should be evaluated;
 - Developing systems to embed evaluation across the DFIDP office;
 - Building evaluation capacity within the office and with partners;
 - Sharing lessons learnt with key partners;

- Using these evaluations to strengthen our drive for better results and value for money.
8. DFID uses the OECD definition of evaluation as “the systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation, and results in relation to specified evaluation criteria”. The differences between monitoring and evaluation are described in Annex 1

IV. DECIDING WHICH PROGRAMMES TO EVALUATE

9. It is neither feasible nor desirable to evaluate every single DFIDP programme. In some cases, the evidence base for our intervention might already be very strong. In others, the programme may not be evaluable. In order to determine which programmes to evaluate, the following criteria will be used:
- (i) **Strategic importance for DFIDP.** Does the programme make a significant contribution to the objectives and results set out in DFIDP’s Operational Plan?
 - (ii) **Strategic importance for DFID, HMG and/or the international development debate.** Will the programme deliver on DFID-wide priorities (e.g. women and girls), is it critical to wider National Security Council objectives for Pakistan (e.g. on conflict) and/or will it help to inform approaches elsewhere (e.g. on low-cost private schools)?
 - (iii) **Innovation.** Is this a new approach where the evidence-base is, by definition, weak?
 - (iv) **Risk.** Is an evaluation needed to demonstrate how the programme rolled-out and to capture any lessons?
 - (v) **Size.** Is it a significant investment of financial resources?
 - (vi) **Demand.** Do key partners want an evaluation?
10. Every Business Case for a new DFIDP investment will now include an appraisal against these criteria to determine whether an evaluation is necessary. An initial assessment of the entire DFIDP portfolio suggests that DFIDP will aim to evaluate 50% of its programmes, covering 75% of spend. This represents 22 programmes for which evaluations will be started by 2015 (with several complete by then). Approximately 4-6 of these will be full impact evaluations. The full list of these programmes is set out in Table 1 below. The process will begin in 2011 with evaluations commencing on three large DFID Pakistan investments: the Punjab Economic Opportunities Programme (PEOP), Improving Citizens Engagement through Devolution (ICED) and the Floods response programme.

Table 1

Planned evaluations				
		Evaluation	New/Existing/ Scale-up	Programme Approval Date
EDUCATION				
BSG	Punjab Education Sector Programme I	Yes	Existing	Approved
BSG	<i>Punjab Education Sector Programme II*</i>	Yes	New	Jun-12
BSG	KPK Education Sector Programme	Yes	New	Mar-11
BSG	<i>Education Fund for Sindh*</i>	Yes	New	May-11
BSG	Innovative fund for education	Yes	Existing	Approved
BSG	Education Sector Voice and Accountability Programme (ESVAP)	Yes	Existing	Approved
BSG	Transforming education in Pakistan (TEP)	Yes	New	Jun-11
	What's being evaluated	7	4 new 3 existing	
HEALTH				
BSG	<i>Delivering reproductive health results programme*</i>	Yes	New	Jul-11
BSG	<i>Provincial reproductive health programme for Punjab*</i>	Yes	New	Sep-11
BSG	<i>Provincial reproductive health programme for KP*</i>	Yes	New	Sep-11
	What's being evaluated	3	3 new	
GOVERNANCE AND SECURITY				
GG	Aawaz	Yes	New	Nov-11
GG	Supporting Transparent & Accountable Govt	Yes	Existing	Approved
GG	Improving Citizens Engagement through Devolution	Yes	Existing	Approved
GG	<i>Subnational Governance*</i>	Yes	New	May-11
GG	<i>Border Areas*</i>	Yes	New	Jun-11
GG	<i>Multi-Donor Trust Fund*</i>	Yes	Existing	Approved
GG	Floods programme	Yes	Existing	Approved
	What's being evaluated	7	3 new 4 existing	
Economic Growth				
EGG	Pakistan Financial Inclusion Programme	Yes	Existing	Approved
EGG	Punjab Economic Opportunities Programme	Yes	Existing	Approved
EGG	<i>SME Finance Development Programme*</i>	Yes	New	Nov-11
EGG	Citizen Damage Compensation Fund (CDCP)	Yes	New	May-11
EGG	<i>Social Protection (Cash Transfer/BISP/TA)*</i>	Yes	New	Nov-11
	What's being evaluated	5	3 new 2 existing	
OVERALL TOTALS				
	What's being evaluated	22	13 new 9 existing	

**Italicised programmes are yet to be designed*

- BSG Basic Services Group
- GG Governance Group
- EGG Economic Growth Group

V. STRENGTHENING DFIDP EVALUATION SYSTEMS

Thinking about Evaluation from the Outset

11. All new DFIDP Business Cases will adopt the following approach:

- a. Identify the theory of change behind the intervention, clearly identifying the causal chain that links inputs to outputs to outcomes to impacts and highlighting the assumptions along that chain;
- b. Appraise the existing evidence, and evidence gaps, behind a particular intervention. As per the criteria set out above, for innovative or risky programmes, where the evidence is weak, there will be a presumption of evaluation;
- c. Related, assess the programme against all of the criteria set out in section IV to determine whether an evaluation is necessary;
- d. If an evaluation is necessary, consider all the options, including qualitative and quantitative evaluations (also known as impact evaluations) to provide the most robust evidence of the impact and outcome of the programme. Such evaluations can be complex, expensive and require heavy upfront investment in baselines, so will not be suitable for every programme. They will therefore be considered for more innovative programmes where the impact has yet to be proven;
- e. Formative or process evaluations will also be considered where real time data on outcomes, and likely impacts, is critical. Such evaluations are useful in illuminating how programmes are working;
- f. Programme teams will work closely with the DFIDP statistician, evaluation adviser and economist to monitor progress on impact evaluations, including the cost effectiveness of this process;
- g. To accelerate implementation and enable rapid deployment of evaluation expertise, where possible all new business cases should include the terms of reference for any planned evaluations;
- h. Draw on evaluation expertise from within the Results and Programme Planning Team of DFIDP as well as Evaluation Department in Abercrombie House to discuss evaluation questions and design.

12. This process of embedding evaluation considerations systematically into each new programme design will strengthen significantly our overall approach to results, evidence and value-for-money. It will also allow for baseline data to be collected, for the findings of evaluations to inform potential scale-up of programmes and, ultimately, identify the impact of DFIDP investments.

Key Principles during Implementation

13. In implementing this approach to evaluation, DFIDP will adhere to the following key principles:

- For priority programmes, use independent evaluators. These will be contracted at the start of new programmes and include a combination of local and international universities, academics, research organisations, individuals and/or consulting teams. This will strengthen the validity and legitimacy of findings;
- Where internal DFIDP expertise is used, teams will follow evaluation best practice (see below). This includes using Development Assistance Committee (DAC) criteria, well developed log frames and staff independent of programme teams from across DFID;
- Retrospective evaluations (conducted after implementation) will be avoided wherever possible as the findings have less internal value but may help inform wider debate;
- All evaluations (and reviews) will be conducted to Development Assistance Committee norms and standards as well as global best practice. This includes: establishing the theory of change from the outset; using experienced evaluators; having a clear evaluation design; aiming for replicability; objectivity; transparency, including full publication of methods and data, and peer review by colleagues and third parties;
- Evaluations will integrate thinking on gender sensitivity, statebuilding/peacebuilding and making markets work for the poor to assess the contribution of DFIDP programmes in these areas. We will work on the presumption that all findings should be sex disaggregated;
- All evaluations will carefully consider fully the ethical dimensions of the research.

Building DFIDP-wide Systems to Monitor Implementation

14. In order to ensure this Strategy is implemented, DFIDP will put in place the following mechanisms:

- An office wide timetable setting out planned Annual Reviews/Project Completion Reports/Evaluations to enable efficient planning. This will also allow staff from different teams to evaluate each others' programmes and strengthen their own evaluation skills;
- Evaluation plans in Business Cases will be scrutinised by members of the virtual evaluation network (see below) prior to submission to the Programme Board (the monthly cross-office peer-review panel for programmes and delivery). The Programme Board will also scrutinise evaluation results and embed lesson learning from evaluations across the office;
- Each group will identify an evaluation lead. He/she will identify team evaluation needs, including training and technical assistance, and report on team progress against strategy to the Results and Programme Planning Team;
- All staff will be encouraged by Line Managers to build evaluation objectives into their personal objectives. This will be recognised as a corporate contribution;
- To ensure implementation of this strategy, the Deputy Head of Office will be the DFIDP evaluation champion.

VI. FINANCING EVALUATIONS

15. The evaluation of any programme currently under design will be funded from within the budget allocated to that programme. As a benchmark programme teams should expect to allocate up to 2-5% of programme funds for a robust evaluation. For existing programmes that require independent evaluations, programme teams should explore funding options with the Results and Programme Planning team. Large scale evaluations of smaller programmes might not be financially feasible.
16. Evaluations sometimes go on for longer than the programme itself and might therefore need resourcing over and above that available from within the programme. In addition there could be a number of cross cutting evaluations that do not fall neatly into any individual programme. DFIDP will therefore consider setting up a separate Evaluation Fund to support such activity¹. Programme teams would bid for resources from this Fund which would be managed by the Results and Programme Planning team.
17. The Results and Programme Planning team will develop further the themes for cross cutting evaluations in consultation with colleagues in the Research and Evidence Division and Policy Division. The initial themes are likely to include:
- Democracy, security, governance and cross-Whitehall working in fragile states; and
 - Influencing and building political will
18. DFIDP will develop tools for procurement and commissioning of evaluations that can be shared across the office. This would include a mapping of capable evaluation organisations in Pakistan, agreeing framework agreements to enable rapid commissioning and developing standard Terms of Reference and contracts.

VII. BUILDING INTERNAL CAPACITY

19. This strategy sets out a high level of ambition on evaluation. DFIDP will need to ensure it has adequate capacity to deliver the large number of evaluations now planned.
20. This challenge is particularly acute over the next few months given the large number of new programmes currently under design. Where impact evaluations are being considered the challenge is further multiplied given the need to establish a counter-factual case etc. Asking the right evaluation questions, establishing baselines and managing independent evaluators all require specialist expertise. To help manage this transition phase, DFIDP will source an international evaluation expert to support programme teams

¹ DFID India has developed an Impact Evaluation Fund available to all programmes for this purpose.

21. However, in the medium-term DFIDP plans to build this expertise in-house. Evaluation leads (see para 14 above) will be embedded in programme teams and accredited to level 2 evaluation specialists. All advisers are now also required to be accredited to this level. Programme staff will also be encouraged to do this. This will be built into Performance Management Form (PMFs). DFIDP will prioritise building the capacity of Staff-Appointed In-Country (SAIC) to ensure ownership and sustainability of this expertise in-country.
22. Evaluation leads will report on progress to the Programme Board biannually. The Results and Programme Planning team will help monitor progress through the programme board and will also offer cross-office support, including through the Afghanistan/Pakistan Evaluation Advisor (expected in post by October 2011).
23. A cross office introduction to evaluation course has been delivered to all staff. We will also give everyone access to a number of web tools and resources, including training materials and courses from 3ie based in Delhi. Further training will be rolled out as necessary.
24. In addition, a virtual evaluation community will be established by the regional evaluation advisor including a network of accredited staff to share lessons, build evaluation capacity and (through the programme board) peer review proposals.
25. DFIDP can also draw on the expertise of willing colleagues. Evaluation Department, Research and Evidence Division, and DFID India have all indicated their willingness to support this strategy.

VIII. WORKING WITH PARTNERS

26. Evaluating partner government programmes can be sensitive. It is therefore very important to work closely with government partners at the Federal and Provincial level, as well as with like-minded donors where appropriate. This will allow us to understand their M&E capabilities and commitment to lesson learning, create a network of M&E specialists amongst donors and build national and international capacity on evaluation in Pakistan.
 - **Government of Pakistan (GoP)**: The entry point to discussing evaluation with GoP is the Planning and Development Commission (P&D) but, given the 18th Amendment and devolution of increased responsibility to the provinces, progress will also need to be made at provincial level. Programme teams have already begun this process. We will build on this.

Three departments must be consulted on any evaluation: Finance, P&D and the relevant line ministry (e.g. education). In engaging with GoP on this agenda we will:

 - a. Aim to push evaluation up the GoP agenda;
 - b. Encourage GoP to evaluate and undertake joint evaluations;
 - c. Build a common understanding of evaluation across DFID teams and explore opportunities to build capacity at government level.

We will aim to incorporate evaluation considerations into our regular interaction with GoP on fiduciary risk.

Where possible we will seek to use existing quantitative data sources collected by GoP for evaluations rather than commissioning new data collections. This will reduce costs and bolster the use of government statistics. However the data needs to be fit for purpose and the long time lags in data availability need to be explored with the GoP.

- **Local capacity:** Evaluation capacity in Pakistan varies from province to province. NGOs in-country have the capacity to do monitoring and data collection but not impact evaluations. We will seek to build non-government evaluation capacity and partner with national and local institutions. The Results and Programme Planning team will map out existing capacity. There are few private sector bodies that operate nationally. The World Bank recommends that donors aim to build capacity in-country (this is an ongoing need) through models such as the *open policy institute* or *CLEAR initiative*. According to USAID, there are 17 firms who can conduct evaluations. USAID has worked with 6-7 of them. They will share their market analysis of local supply.
- **Other donors:** DFID Pakistan will work closely with other donors including USAID, the World Bank, Asian Development Bank and the European Commission on embedding evaluation into programmes and building capacity.

IX. EVALUATION COMMUNICATIONS

27. DFID Pakistan is committed to communicating widely the findings of its evaluations and any lessons learned. Our starting presumption is that all evaluation reports will be published, including interim reports for those evaluations that are ongoing. We will provide summaries of key findings and, where appropriate, support dissemination e.g through workshops.
28. In addition, evaluation data will be shared with our partners. We will ensure that lessons learned are incorporated into new programmes across DFID and shared more broadly with relevant stakeholders. This will also encourage further research into DFID priority areas.
29. Additional workshops or events will also focus on lesson learning from evaluation results. These will be arranged as necessary by the Results and Programme Planning team. The Programme Board will also be used to monitor progress of implementing the evaluation strategy, including number of evaluations commissioned and progress in building evaluation capacity.

DFID PAKISTAN JULY 2011

ANNEX 1: Monitoring Compared to Evaluating (Source: FCPD)

	Monitoring	Evaluation
• Definition	<ul style="list-style-type: none"> The continuous (regular) gathering of data and information about the progress of project implementation and management 	<ul style="list-style-type: none"> The periodic analysis of project information regarding achievements, effects and impact
• Main purpose	<ul style="list-style-type: none"> To keep track of progress To provide information on actual performance to enable recognition of best practice and corrective action/decision-making To provide early warning signals about change 	<ul style="list-style-type: none"> To assess overall outputs, effect and impact To determine the relevance of project design To identify and explain lessons learnt and provide information for strategic planning and policy options
• Main focus	<ul style="list-style-type: none"> To monitor the use of resources of which the project has control Regular tracking across a range of significant Indicators to provide immediate feedback Records Inputs, tasks, Activities and Outputs during implementation 	<ul style="list-style-type: none"> Assess the project relevance, design, effectiveness, efficiency, sustainability, impact and achievement of the project's Purpose Concludes relationships between means (process) and ends (results and impacts)
• Timing	<ul style="list-style-type: none"> Continuous / ongoing during project implementation 	<ul style="list-style-type: none"> Time-bound. At specified times for specified purposes – before implementation (baseline), during implementation and post-completion (impact assessment) Retrospective analysis
• Implementation	<ul style="list-style-type: none"> Often staffed by the implementing project and primary partners for the duration of the project 	<ul style="list-style-type: none"> Often involves selecting internal or external specialists and may be commissioned internally, by donors or special interest groups