

The Use and Usability of Evaluation: *Staying One Step Ahead* – Using and communicating evaluation results from DFID's Girls' Education Challenge (GEC)

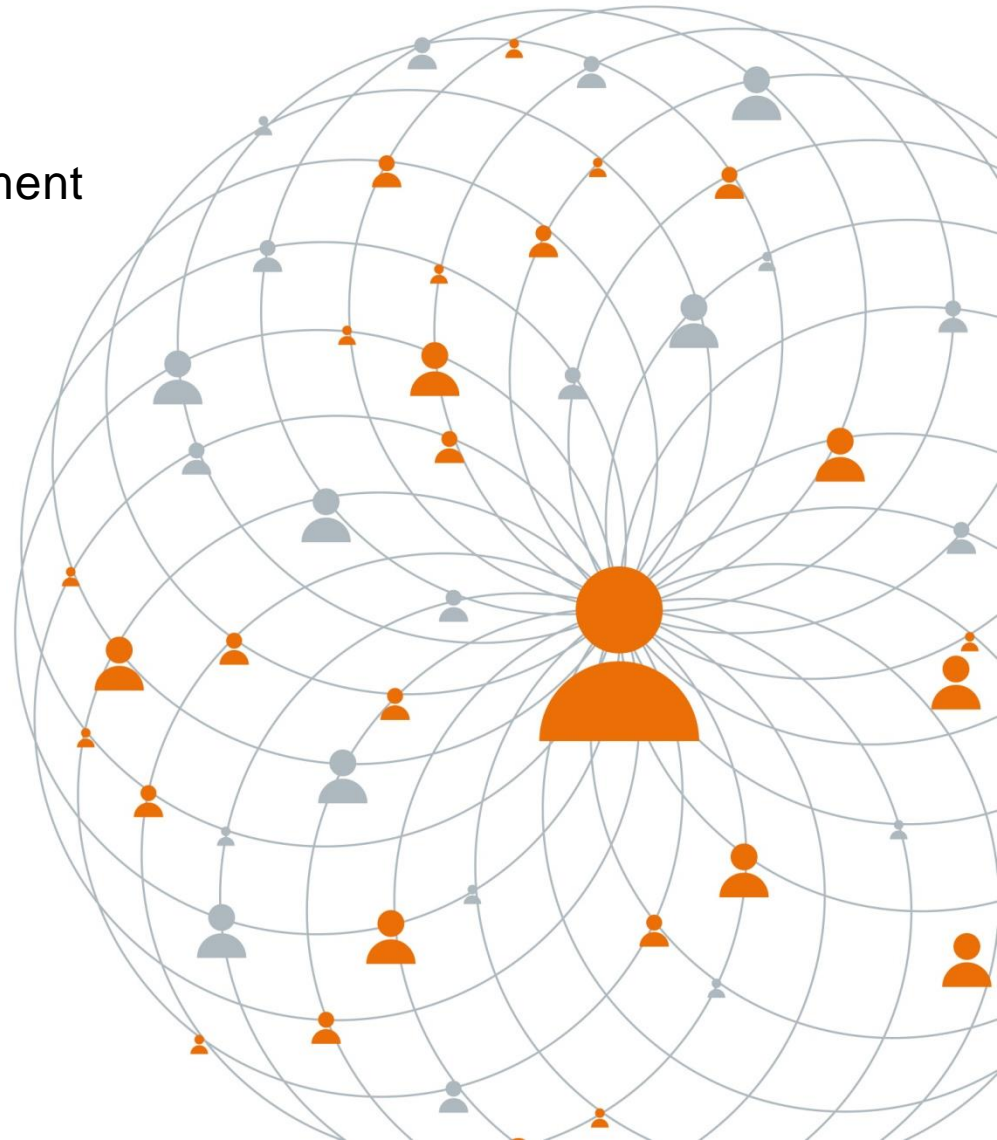
James Bonner
Department for International Development



Simon Griffiths
Coffey International



10 May 2017
UK Evaluation Society Conference

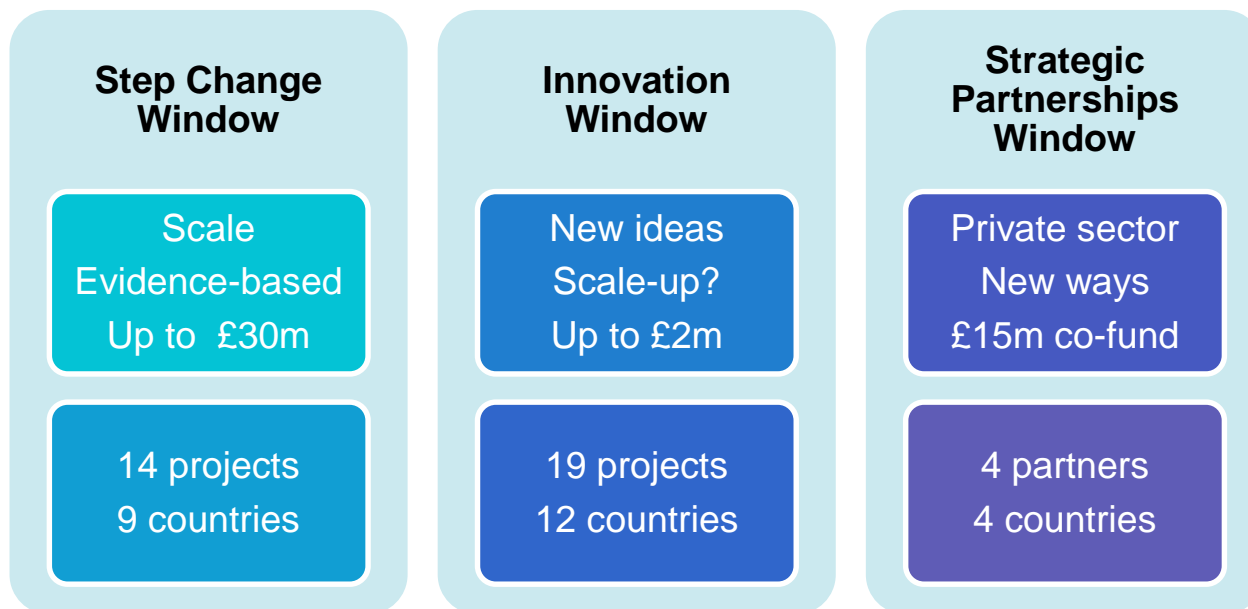


Presentation Objectives

- Share our reflections after 5 years on the GEC
 - Review our decisions, actions and utility
 - Consider alternatives to our approach in hindsight
-

Girls' Education Challenge (GEC) at a glance

- 1m marginalised primary and secondary girls with improved learning outcomes
- £355m global challenge fund – 37 projects / 18 countries – ‘holistic’ approaches
 - 2012 to March 2017
- Fund Manager (FM) – PwC led consortium – programme /performance (M&E) management
- Evaluation Manager (EM) – Coffey led consortium – independent programme evaluation



- Rigorous evaluation – counterfactual longitudinal evaluation + qualitative research – all projects
 - Impact evaluation supports part PbR on literacy and numeracy – SCW & IW

SCW approach – 2 streams of evidence and 2 perspectives

Evaluation research led by the Evaluation Manager for the SCW only:

- Evaluates the **wider impact of the SCW as a whole** on girls in their targeted communities
- Produces **lessons learned** to inform GEC and wider DFID programming and policy
- Sample is **representative of the general SCW target population** (initially girls aged 5-15 at baseline)

Evaluation research led by Projects:

- Projects evaluate the impact of **their interventions** on their **specific target groups**
- Projects produce lessons learned about the effectiveness of **their interventions**
- Projects' samples are larger than the Evaluation Manager's and representative of their **specific target groups**

Evaluation Manager SCW Midline Research ¹			
6000 Household surveys and EGRA & EGMA learning tests	3000 Children traced through school visit surveys	7000 School-based learning tests of boys and girls (3 countries)	750 Qualitative in-depth interviews

Project SCW Midline Research ²		
14 Midline Evaluation Reports	14 Large datasets for household surveys & learning tests	14 Outcome Spreadsheets

Note 1: Research repeated at baseline and endline except school-based learning tests and qualitative in-depth interviews that are not repeated at endline.

Note 2: Research repeated at baseline and endline

Approach to Knowledge Management and dissemination

- FM leads knowledge management and dissemination
 - All communication and dissemination coordinated through **Knowledge Management Working Group** (KMWG)
 - EM and DFID sits on KMWG
 - Global and regional midline and endline learning events
 - EM generates independent evaluation evidence for dissemination
-

KM and dissemination – What's happened?

- **Project evaluation evidence**

- Ad hoc publication by individual projects
- Sharing of lessons between projects and stakeholders (learning events, newsletters)
- Ad hoc sharing lessons with policy-makers in-country (events, meetings)

- **EM evaluation evidence**

- EM reports published on DFID's website with policy briefs ([see additional slides](#))
 - Quantitative learning and household survey data accessible on UK Data Archive (<https://discover.ukdataservice.ac.uk/catalogue/?sn=7736&>) – unclear how much use?
 - Presentations of findings (DFID, FM, projects, stakeholders, policy-makers)
 - Some sharing of lessons with policy-makers (conferences, events, publications)
-

Use of evaluation at the project level

What worked well?

- Rigorous quantitative approach **improved capacity** of some projects /external evaluators and research organisations
 - Quantitative impact learning data of high enough quality to **support PbR**
 - **Large amount of quantitative learning data** available for thousands of marginalised girls living in very challenging environments
 - **Projects more aware** of what other projects are doing, where and for whom
 - Some projects **changed their approaches** as a result of evaluation process and new evidence
 - Projects used evidence to **engage and inform stakeholders** and partners in-country
-

Use of evaluation at the project level

What worked less well?

- **Variable quality** of evidence – coverage, accuracy, reliability – lack of evaluation capacity
 - **Qualitative evidence generally weak** – not always able to fully understand:
 - Context and its influence?
 - How and why effects happened?
 - Process of change?
 - What specifically worked?
 - Gender differences?
 - Specific effects on different sub-groups?
 - Some projects **lack capacity to fully absorb** and use complex quantitative data
 - **Reported lessons learned improved but weak** – difficult to generalise beyond specific context in which projects operate
 - **Unclear extent evidence used to drive improvements** in activities and outcomes
-

Use of evaluation at the programme level

What worked well?

- Made sense of a huge amount of data and evidence that **enhanced accountability** for large programme spend
 - DFID used EM baseline evidence to **shape education policy** – e.g. found learning levels (quality of education) much worse than access at baseline
 - **EM and FM raised awareness and informed other parts of DFID** about the GEC – PbR, M&E, private sector engagement, learning & learning measurement, challenge funds
 - Smaller EM analytical products (process review, thematic research) **quicker to produce and more responsive** to DFID's learning priorities at the right time
 - **Increased global awareness** among policy-makers about DFID's investment in girls' education
-

Use of evaluation at the programme level

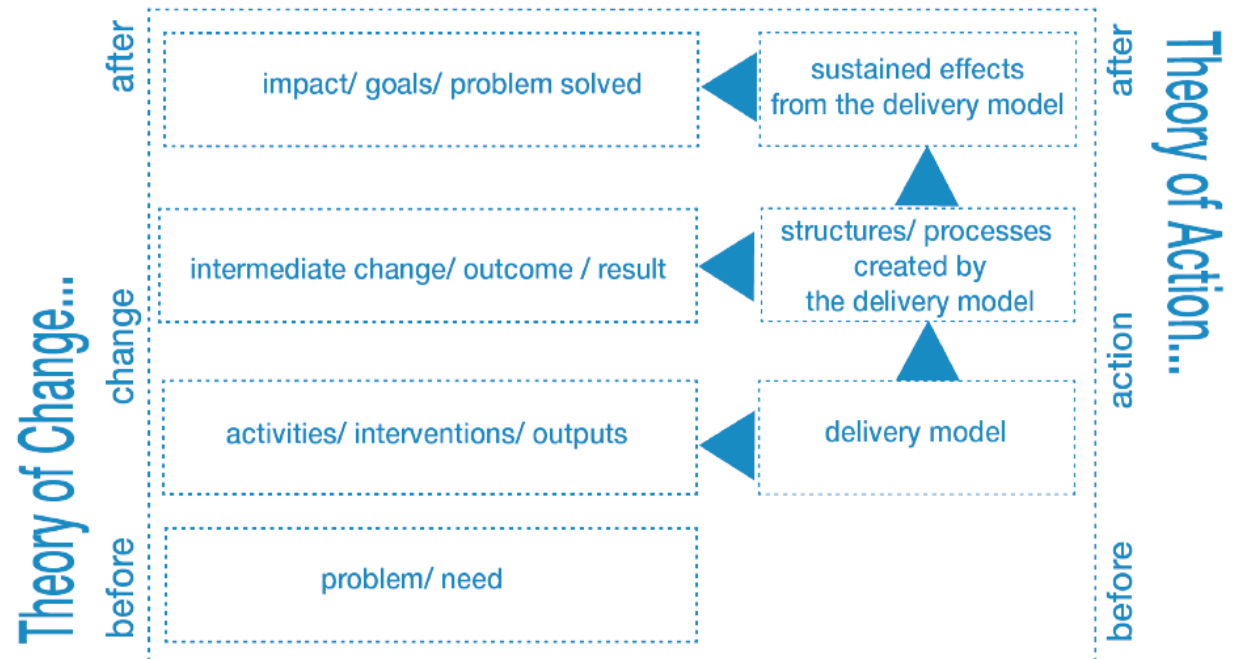
What worked less well?

- Big analytical products (baseline, midline, endline) take **long time to produce** – too late to inform immediate improvements in GEC programme and projects?
 - **Very difficult to generalise** about what works from such diverse target groups, contexts and types of interventions
 - **Difficult targeting** right stakeholders with right evidence at the right time
 - **Difficult managing differences in key learning messages** emerging from project evidence vs EM evidence – inherent differences
-

What would we have done differently?

Incentivise and support projects to conduct much better and more useful qualitative research and analysis

- **Remove or reduce burden on projects** for quantitative impact evaluation – EM do learning tests? Project do learning tests and EM do HH surveys?
- Greater focus on developing **learning and adaptive management** processes
- Greater focus on:
 - **Process** evaluation – design vs delivery effects?
 - **Situation** analysis – conditions for change?
 - **Gender** analysis – perverse effects?
 - **Sub-group** analysis – change for whom?



What would we have done differently?

Allow sufficient time to ensure that key learning priorities are identified at the start and make sure these drive the evaluation design, planning, delivery processes

- **Allow more time** to discuss and plan at start and throughout – revisit evaluation design
- **Engage DFID** – make identifying learning /policy priorities first step in evaluation design
- **More segmented and targeted approach** to identifying and engaging audiences
- **Maintain large analytical products** at baseline and midline (the new endline) – drop endline?
- **Iteratively commission more smaller, more focused, learning-responsive products** in between baseline and endline?



Additional slides – GEC Evaluation Manager publications

EM Evaluation Publications	Publication Webpage (www.gov.uk/)
Step Change Window Baseline Report (January 2015)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/425360/Step-change-window-full2.pdf
Innovation Window Baseline Report (January 2015)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/425338/innovation-window-Baseline-Report-fulla2.pdf
Strategic Partnerships Baseline Report (January 2016)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/494679/Baseline-Strategic-Partnerships-Window-13January2016.pdf
Evaluation Manager Girls' Education Challenge Step Change Window Baseline Quantitative (Household Surveys and Learning Test) Data (2013-2014) including Code Books and User Guides	https://discover.ukdataservice.ac.uk/catalogue/?sn=7736&
GEC Process Review Report (February 2016)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/501596/Process-Review-Report.pdf
Process Review Brief (February 2016)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/501592/Process-Review-Brief.pdf
Thematic Research Report – <i>Narrow Windows, Revolving Doors</i> (March 2016)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/520249/Narrow-Windows-Revolving-Doors-What-affects-adolescent-girls_ability-to-stay-in-school-Research-Report-March-2016.pdf
Thematic Research Report Brief (March 2016)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/520248/Narrow-Windows-Revolving-Doors-What-affects-adolescent-girls_ability-to-stay-in-school-Policy-brief.pdf
Step Change Window Midline Evaluation Report (March 2017)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/609664/Step-Change-Project-Window.pdf
Innovation Window Midline Evaluation Report (March 2017)	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/609665/Innovation-Project-Window.pdf