LEARNING SYNTHESIS METHODOLOGY

The Office of Development Effectiveness (ODE) in Australia's Department of Foreign Affairs (DFAT) undertook a Review of 2017 Program Evaluations. Phase 1 assessed the quality of all 37 program evaluations completed by DFAT in 2017 (ODE 2018). Phase 2 was a synthesis of the learnings from the evaluation reports, designed to see what could be learned from them that would be useful for DFAT across the following topic areas:

- promoting gender equality
- policy influence.

This paper summarises the methodology for Phase 2 and should be read alongside the two learning papers.

DFAT has long understood that context is critical in terms of aid effectiveness but has been less able to say which aspects of context are important and in what ways. The previous ODE syntheses of evaluation reports (ODE 2014) adopted a narrative synthesis method that analysed similar findings from across evaluation reports, to answer the questions 'What works and why?' or 'What doesn't work and why?'. It was hoped in 2018 to take this a step further to find additional insight in terms of the mechanisms[1] underlying successful or unsuccessful interventions and the particular aspects of context or implementation that allowed the mechanisms to 'fire' successfully. As a result, a realist synthesis evaluation methodology was chosen (Box 1).

Box 1: Realist evaluation and realist synthesis

Pawson and Tilley's analogy of gunpowder is useful to understand mechanisms and context. The chemical mechanism is clearly understood, but it will only explode if the context or conditions (such as the gunpowder being sufficiently compacted and dry) are right (Pawson and Tilley 1997). This understanding requires the evaluation or evaluation synthesis to identify both the mechanisms by which programs operate and the contexts that allow their mechanisms to operate. It makes explicit how the aspects of context and intervention interact in real, open systems. The emphasis is on identifying context-mechanism-outcome configurations to answer the question: 'What types of aid work for whom in what circumstances and in what respects, and how?'

The initial questions guiding the synthesis were:

1. What are the underlying mechanisms by which policy influence or promoting gender equality work in theory?

2. How do different contexts affect different DFAT-supported mechanisms for influencing policy and promoting gender equality in terms of producing desired outcomes?

3. How successfully does DFAT adapt its policy influence and promote gender equality work to different contexts?

---

[1] A mechanism is the invisible force that explains what it is that makes an intervention work. In the classic example, it is not the CCTV camera (the intervention) that causes less crime, but one or more mechanisms, such as deterrence.

4. In what ways is DFAT aid capability affecting program outcomes?

5. What other learnings about mechanisms and how context affects outcomes can be gleaned from these documents?

1.1. SOURCES OF INFORMATION

The primary sources of information for the review were the 37 evaluation reports completed by program areas in 2017 and 14 interviews with program evaluators and DFAT staff. The 37 evaluations were identified in the 2017 DFAT Aid Evaluation Plan and completed and published in 2017. Table 1 lists all the evaluation reports reviewed. Other sources of information included the academic literature and grey literature around policy influence and promoting gender equality as well as some design documents for the evaluated initiatives.

1.2. METHOD

1.2.1. Choosing initial codes for explanatory text

The researcher provided an initial list of 16 codes for policy influencing outcomes, promoting gender equality and aid capability, based on the literature on policy influence, DFAT's gender policy (2016) and DFAT's Aid Health Check Update (2018).

1.2.2. Coding of reports by the team

The review team comprised a researcher (Jo Hall) and six ODE team members.[2] Explanatory text, that is 'text in the evaluation report that explains the effect—either positive or negative—that something is having on intermediary or final outcomes', was highlighted and one or more relevant codes were selected by review team members as they quality reviewed each of their assigned reports (Phase 1). This process was moderated three times over several weeks. Inconsistencies between team members in selecting the same text and choosing the same codes improved over the course of the moderations, but there were still some inconsistencies after three moderation exercises. To address this, the researcher checked all coding on entering the data into Nvivo 11 Pro software.

1.2.2.1. Interviews and coding of interview transcripts

Fourteen interviews were held with 12 program evaluators, one former initiative manager and one former DFAT employee. All interviewees were highly specialised in their fields or very experienced evaluators. Half were women. Three of the interviewees currently work for DFAT and five interviewees worked previously for DFAT in long-term, senior positions.

The ethical aspects of this research followed the Australian National University's ethical guidelines, including obtaining informed consent to participate from all interviewees.

Interviews took a semi-structured form, seeking detail about context and mechanisms identified in the evaluation report, or seeking more information or clarification of details of the evaluation that did not make it into the report. Transcriptions were entered into Nvivo and coded in the same way as for the evaluation reports.

---

[2]   There was a pool of six ODE reviewers (Gina De Pretto, Gabriella Tauni, Fiona Meehan, Irene Wettenhall, Jacinta Overs and Leo Carroll), but not all reviewers were able to participate in all the moderation exercises or team meetings. Some reviewed more reports than others. This improved consistency in coding over time. In any future exercise a much smaller team would be recommended.

1.3. ANALYSIS AND SUB-CODING

Analysis, sub-coding and further literature review was an iterative process undertaken by the researcher. Relevant explanatory text references were extracted using Nvivo and reviewed, looking for particular mechanisms that seemed to be operating. As potential mechanisms and themes emerged, sub-codes were added in Nvivo. Sub-coded text was analysed for the strength of evidence and detail around the mechanism, outcome and context. Any corresponding literature was explored. This process continued progressively until analysis was complete.

1.4. PEER REVIEW

For each of the learning papers, a preliminary working paper was prepared by the researcher, drawing together the main evidence and tentative areas of findings. These rough working papers were shared with ODE for preliminary feedback on contesting the strength of the evidence and indicating which aspects might be more useful to emphasise.

Following feedback from ODE a more refined draft product was prepared and shared with ODE for each paper. ODE reviewed these with the Independent Evaluation Committee and with relevant areas in DFAT (Gender Equality, Governance, Fragility and Water Branches) and comments were provided back to the researcher until a final product was agreed.

1.4.1. Limitations

The evaluation reports are not representative of the whole aid program, nor are they a random sample. However, they do represent a large volume of aid (nearly $2.5 billion in total program value), and they cover a range of DFAT's target countries (17 in the Pacific, 15 in South East and East Asia and 5 in South and West Asia). There were no evaluation reports of initiatives in Sub-Saharan Africa and the Middle East or Latin America and the Caribbean. Broadly, all DFAT program sectors were covered. However, as this exercise was seeking to learn from the evaluations and not to be an exhaustive review, this limitation did not threaten the validity of the synthesis. Close iteration with the literature and peer review with DFAT helped ensure the learnings were broadly applicable and valid.

A second limitation was the degree to which the evaluation reports themselves were explicit about mechanisms or explained the links between context and outcomes. Commonly, evaluation reports do not explain the detail of what an intervention actually did or test the theory by which an intervention is operating. The researcher had to deduce the mechanisms that seemed to be operating from the text of the reports. As per their mandate, the evaluation reports focused largely on performance of the aid initiative and looked for aspects of implementation rather than context to explain successful or unsuccessful outcomes. This tended to overshadow any relevant contextual factors (such as political stability or volatility or fragmentation among different policy making institutions). At interview, people sometimes found it difficult to explain what it was about the intervention (mechanism) or the context that was important in relation to the outcomes being successful or otherwise.

A related limitation was around the quality of the evidence contained in the evaluation reports. For the gender equality questions, for example, some evaluation reports considered women and gender interchangeably so provided less useful information for the synthesis. In other evaluation reports it was not always clear on what basis an evaluator had made a judgment. Each piece of evidence was considered on its merit for the synthesis.

The limitations made it difficult to answer the original research questions in full. But the synthesis was useful in identifying families of mechanisms and in generating relevant learnings for DFAT.

Table 1 Characteristics of evaluation reports: Initiative name, approximate value (over the life of the initiative) and sector.

Initiative name	Approximate value/$million	Sector
Australia's Education Partnership with Indonesia*	369	Education
Eastern Indonesia National Road Improvement Project* ‡	346	Infrastructure
Indonesia Infrastructure Initiative Phase 2	330	Infrastructure
Pacific Women Shaping Pacific Development*	320	Gender
Australian support for improved nutrition in Timor-Leste*	132	Health
Australia Indonesia Partnership Rural Economic Development Program*	112	Economic Development
PNG Health and HIV Multilateral Partnership*	112	Health
Solomon Islands: Health Sector Support Program Phase 3 §	66	Health
Solomon Islands: Education Sector Program 2: Joint Annual Review on Co-funding §	63	Education
Cambodia Agricultural Value Chain Phase 1* ‡	60	Agriculture
Australia's humanitarian assistance to Myanmar	59	Humanitarian
Pacific Women Shaping Pacific Development: activities in PNG*	58	Gender
Vanuatu: Governance for Growth* ‡	52	Governance
Vietnam Mekong Delta Transport Infrastructure Program §	47	Infrastructure
Cambodia: 3i—Investing in Infrastructure*	46	Infrastructure
Vanuatu: Roads for Development	29	Infrastructure
Fiji: Pacific Women's Country Plan Review	26	Gender
Australia Afghanistan Community Resilience Scheme	24	Community Engagement
Timor-Leste: National Program for Village Development Support*	22	Community Engagement
Fiji Community Development program	19	Community Engagement

Initiative name	Approximate value/$million	Sector
Vietnam: Integrated Coastal Management Program	17	Climate Change Management
Pacific Financial Inclusion Program	16	Economic Development
Independent Evaluation of the Pacific Leadership Program Phase 3*	16	Governance
Mekong Business Initiative	11	Economic Development
Pakistan Trade, Investment and Policy Reform Program	10	Economic Development
Australia response to El Nino in PNG	8	Climate Change Management
Pakistan Challenging Gender Based Violence Program*	8	Gender
Philippines: Australia's Support for Peace in Mindanao	7	Community Engagement
Samoa Inclusive Education Demonstration Program	6	Education
Greater Mekong Water Resources Program	6	Water Resource Management
Build Back Safer Schools for All: Australia's response to Nepal earthquakes	5	Humanitarian
Sri Lanka: Community Forestry Program	5	Agriculture
Bougainville Youth Initiative	5	Community Engagement
Tropical Cyclone Winston Education Response	4	Humanitarian
Strengthening Pre-service Teacher Education in Myanmar	3	Education
Cleared Ground De-Mining Project in Palau	3	Demining
Vanuatu: Strengthening Early Childhood Care and Education*	1	Education
Total	2,422	

Source: DFAT documents

Notes

* The evaluation team leader or project manager (in one case) was interviewed for these initiatives.

± Vanuatu Governance for Growth included two evaluation reports commissioned by DFAT. One was a review by the Overseas Development Institute.

‡ These evaluations were conducted by ODE.

§ These evaluation reports were excluded from the ODE evaluation quality review (Phase I) as not constituting regular evaluations.

1.5. REFERENCES

Department of Foreign Affairs and Trade. 2016. DFAT aid evaluation policy.

Office of Development Effectiveness. 2014. Learning from Australian aid operational evaluations.

Office of Development Effectiveness. 2018. Review of 2017 Program Evaluations.

Pawson, R. & Tilley, N. 1997. *Realistic evaluation*. Thousand Oaks, Sage.