

ODE on whole of government approaches to law and justice: where's the evidence?

Author : Tracey Blunck

Date : March 19, 2014

In December 2012 AusAID's Office of Development Effectiveness published a thematic evaluation of Australia's law and justice assistance titled [Building on Local Strengths: Evaluation of Australian Law and Justice Assistance](#). Law and justice assistance (e.g. technical assistance to courts and support for police service and corrections systems) accounted for almost 15% of Australia's bilateral aid program in 2010–11. In this post, I'm going to review the evaluation with a focus on whether the arguments made by the authors, and the evidence they rely on, hold up under scrutiny. An evaluation of the evaluation, so to speak! Since the evaluation is quite wide-ranging I'm just going to focus on the authors' findings relating to whole-of-government delivery.

What is whole-of-government delivery and what did the authors find?

Whole-of-government (WoG) delivery is a defining feature of Australian law and justice assistance. Rather than projects being delivered solely by AusAID and its contractors, many are the responsibility of various federal and state government departments. The Australian Federal Police (AFP) is Australia's second largest aid agency (after AusAID). Several other departments are also involved.

The authors' findings in relation to WoG aid are set out on pages 31 and 32 of the evaluation. They find that there are advantages such as increased willingness from recipients to engage with peers rather than contracted advisors and the development of mutually beneficial long-term relationships (the positive findings). The authors also identify problems with the approach (the negative findings), including a risk that fragmented programming will lead 'to less effective support and poor value for money' and a 'proliferation of small-scale assistance', which may also decrease effectiveness.

How convincing is the evidence?

To prepare the evaluation, the authors undertook three in-depth country case studies in [Cambodia](#), [Indonesia](#) and the [Solomon Islands](#). The cases chosen for the sample were not random, and this is potentially the first issue for the evaluation. It is interesting to note that despite the purposive sampling, Indonesia was chosen for the study, where law and justice assistance made up less than 5% of bilateral aid between 2005 and 2010, but Papua New Guinea, Australia's longest running law and justice program, was excluded. I'm not in a position to analyse whether a selection of other countries would have produced different results, but I think that it's important to bear in mind that the sample of countries selected for this survey might not have been the most useful.

After looking closely at the evidence used to support the authors' findings, I was left with the overall impression that the quality of evidence was poor. The best evidence in support of the positive findings is in

the Indonesian case study, which contains a section devoted to "twinning" and the assistance provided by a variety of Australian government agencies (pp. 32-5). The evidence presented largely supports the authors' positive findings on the benefits of peer engagement with reference to literature reviews and interviews.

For me, the real problem lies with the authors' negative findings. I could find no evidence in the evaluation or the case studies to support them. In terms of fragmented programming, the authors state on page 31 of the evaluation:

... we found more instances of parallel support by different agencies, with poor coordination and elements of interagency rivalry. While mutual support between AusAID and the AFP is clearly better than it has been in the past, the level of collaboration in the design and delivery of assistance is still low.

I reviewed all three case studies and could find no evidence to support this statement. The authors seem to have made no attempts to highlight the programs where interagency rivalry occurred and where AusAID and the AFP provide mutual support for one another.

I found the same lack of evidence when I examined the claim of proliferation of small-scale assistance. The authors suggest on page 31 of the evaluation that due to a lack of coordination there have been 'many reports of duplication and overlap' with a 'common complaint' being that so much training is offered it 'amounts to a significant drain on [officials'] capacity'. I couldn't identify any references to such reports or complaints in the case studies. The use of direct quotes could have been an easy solution to this problem.

Another issue I identified was that nearly all of the supporting evidence comes from the Indonesian case study. There is no mention of WoG aid (as it relates to Australia) in the Solomon Islands case study, and only very limited discussion in the Cambodian case study. In my view, this raises questions of whether the authors' findings hold true outside of Indonesia. This criticism is particularly relevant given that the evaluation sought to draw generalisations to improve Australian law and justice assistance as a whole.

How convincing is the argumentation?

A general criticism I have is of the authors' failure to identify the sources of their findings. The authors may have found many examples of agencies competing for funds, but they give no indication where these examples can be found. This makes reviewing the evidence difficult and frustrating, if not impossible.

A recent [report](#) prepared for DFID argues that demonstrating causality lies at the 'heart' of an impact evaluation (p. 2). Ultimately, the question the authors were aiming to answer with their findings and evidence was: 'how effective is WoG delivery?' Their findings suggest that WoG delivery increases effectiveness in terms of peer-to-peer relationships but decreases effectiveness when programs are fragmented and provided on a smaller-scale.

It appears to me that to make their causal claims, the authors have used an approach akin to process tracing, a qualitative method that [attempts to identify](#) [pdf] the 'specific ways a particular cause produced ... a particular effect'. An important part (as [emphasised](#) by Hughes and Monroe, p. 32) of this process is identifying and dismissing (if possible) alternative explanations for the observed effect to help ensure the non-spuriousness of results. I couldn't identify any attempts by the authors to do this in the case of the positive or negative findings. This may be because the authors felt their findings were logical (I certainly thought they were) and, at least in the case of the negative findings, were supported by external literature (see for example Riddell's [Does Foreign Aid Really Work?](#), p. 216). However, in terms of demonstrating causality, this approach is inadequate. At the very least, the authors ought to have considered whether the variables identified as impacting on effectiveness held across other cases and not just within Indonesia.

Suggestions for improvement

In my view, the evaluation could have been improved if each case study was broken down into smaller parts with one part specifically evaluating the effectiveness of WoG aid. Taking this approach may have revealed more cases of increased/decreased effectiveness that could have helped inform the authors' findings. The evaluation findings would have also been improved by utilising mixed methods research. For example, in the case of the proliferation of small-scale assistance, assuming the finding was valid, quantitative data would presumably be available to demonstrate a rise in funding requests for increasingly smaller projects. Alternatively, a set of structured interviews with informants may have been useful.

Conclusion

To sum up, after my evaluation of the evaluation I was left with the impression that the evidence and arguments presented by the authors in support of their findings were weak. While the authors' findings are plausible, these weaknesses undermine the utility of the evaluation, and raise questions about its credibility.

Tracey Blunck completed her Masters in Public Policy specialising in Development Policy at the Crawford School last year, and won the Raymond Apthorpe prize for the best student. This post is based on an essay she wrote as part of her Aid and Development Policy class, in the second semester of 2013, where students were required to review some aspect of a recent evaluation from the Office of Development Effectiveness. On Friday 21 March, Devpolicy will be hosting a [forum](#) to discuss new ODE evaluations on volunteers and aid quality.