



Evaluating Australian aid evaluations

By Terence Wood

It is aid's lot to be a strange amalgam of certainty and doubt. Certainty in the claims splashed across the websites of aid donors, in the brochures of NGOs, and in the speeches of politicians. Doubt in the minds of actual aid workers. It's not that aid doesn't work (sometimes it does, remarkably well), it's that too often, for any individual aid activity, too little will be known about how well it worked and why. With absolute disasters and spectacular successes the answers are clear enough. But a lot of aid work lies between these extremes. And not knowing

means not learning, and not improving.

There's an obvious solution to all this: evaluation. Spending money to systematically glean whether individual aid projects have worked, and why they have or have not. Like all obvious solutions in the world of aid, this one's been thought of. And almost all credible aid agencies and NGOs invest in evaluation. The Australian Government aid program does. The trouble is that good evaluations themselves aren't easy. Useful evaluations require skill, resources and effort.

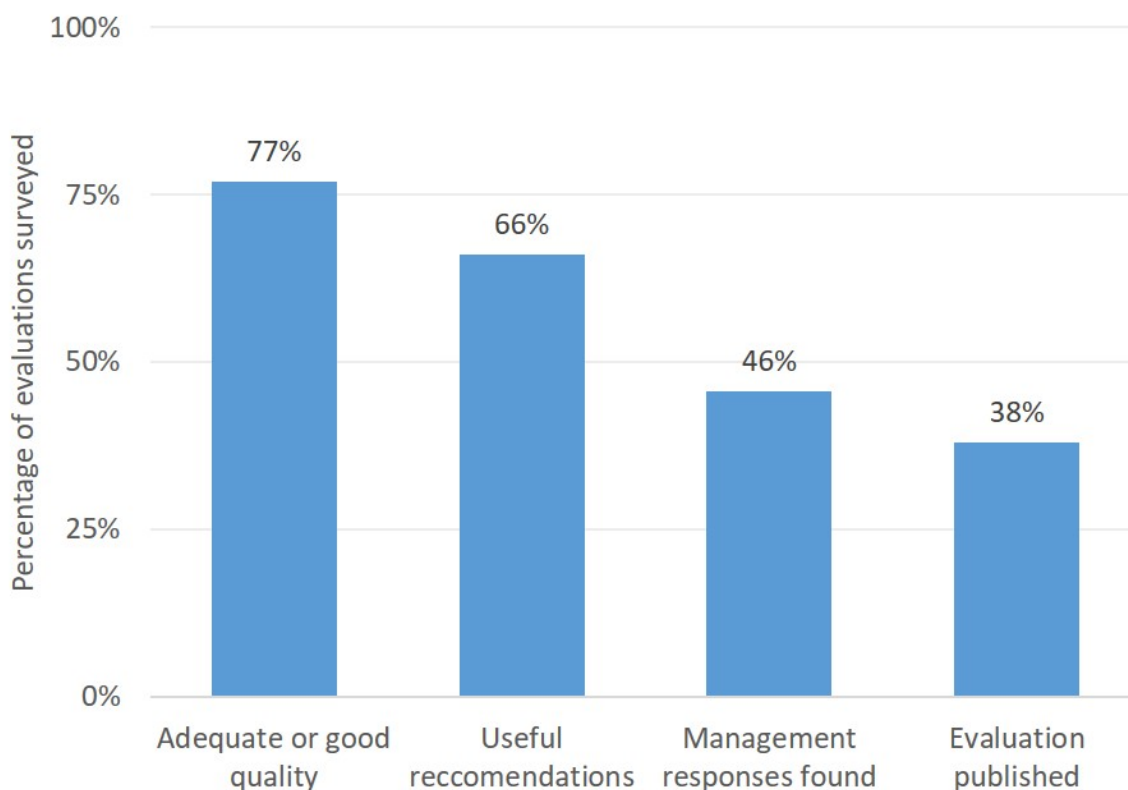
Last month I gave [a presentation](#) on the Office of Development Effectiveness's [review](#) of the quality of Australian Government aid program evaluations conducted in 2014. (This is the most recent review; the review process lags the production of evaluations themselves.) The Office of Development Effectiveness (ODE) is a quasi-independent entity within DFAT tasked with strengthening the Aid Program's ability to evaluate its work. The mere fact that this review exists was good news. The ODE has long been a strong point of Australian aid and thankfully it survived the end of AusAID. And as the quality of the review showed, ODE is still in good intellectual shape. (As you can see in slides 7 and 8 [here](#) [pdf], I had a few criticisms of the review but it was a solid, well-thought-through piece of work.)

The good news in the review itself was that, by the reviewers' assessment, 77 per cent of the Australian aid program's evaluations in 2014 were of adequate or better quality. Given the shock therapy inflicted on the aid program in 2013 and 2014, this is something worth celebrating. The not-so-good news is that adequate evaluations, and evaluations that contribute substantively to learning, are two different things. One problem, as an audience member at the ODE forum pointed out, was that most of the evaluations reviewed were simply evaluations of processes and outputs, and didn't focus on actually trying to capture impact-aid's actual contribution to human welfare. So 'adequate or good' doesn't mean gold

standard but it's a lot better than bad.

Other issues can be seen in this chart:

Key numbers from evaluation review



It shows us that, although 77 per cent of evaluations were “adequate or good”, only two thirds of evaluations were judged by the reviewers to have useful recommendations. Needless to say, useful recommendations are important if evaluations are actually going to help improve aid, so the fact a third of evaluations lacked these is a concern. Worse still though, fewer than half of the evaluations the ODE team looked at had management responses. As the name suggests, management responses are documents in which aid managers demonstrate they’ve engaged with an evaluation, detail their take on its findings, and outline changes that may occur. When management responses are missing it

is hard to be confident that the people with the power to change things have taken the time to learn what needs to be changed. Such a low management response rate suggests an aid program that isn't taking its evaluations seriously enough.

What's more, only 38 per cent of evaluations from 2014 ended up being made public via publication on DFAT's website. Sometimes there were good reasons for this. But the publication rate was still much lower than it should be. A 38 per cent publication rate is a transparency failure.

What explains this state of affairs?

The good news—the quite good evaluations—can be explained by the fact that there is still some development knowledge as well as a commitment to learning in DFAT. As for the bad news, the causes are perennial ones: time and money.

Really good evaluations (particularly ones of impact) come at a price. And although getting good staff engagement is partially an issue of organisational culture, a lot comes down to the more tangible constraint of actually having enough staff. You need enough aid staff to allow people the time to engage and learn from them evaluations. And this was clearly an issue in 2014. You also need sufficient money to be budgeted for evaluations. The review itself did not state that money was an issue in 2014, but given aid cuts had started, and given evaluations are easy targets in lean times, my guess is that this was a constraint.

People often talk about value for money in aid. That's nice in theory. But it's often putting the horse in front of the cart. When it comes to evaluations, the catch-cry ought to be 'money for value'. Because money spent now on evaluations, and giving staff space to actively engage with evaluations, is an integral part of increasing the value of aid.

Terence Wood is a Research Fellow at the Development Policy Centre. Terence's

research interests include aid policy, the politics of aid, and governance in developing countries.

About the author/s

Terence Wood

Terence Wood is a Research Fellow at the Development Policy Centre. His research focuses on political governance in Western Melanesia, and Australian and New Zealand aid.