Evidence and scale

Notes for remarks to the CIFF Board dinner, 16 May 2014

I'm going to say a few words this evening about the importance of evidence and rigour, and the challenges of taking development programmes to scale. I'm going to back up my argument with some new evidence about what it takes for development programmes to succeed, which I think has implications for the way that CIFF works.

But let me begin with a confession: **I am not a great fan of philanthropic foundations**.

I have nothing against foundations in principle. I was utterly persuaded by Mike Green and Matt Bishop's book, *Philanthrocapitalism*, that foundations have a unique contribution to make. Because of the way they are funded and governed, they can do things that official aid agencies struggle with. They can bring business-like project management and partnership skills to development projects. They can be guided by rigorous evidence rather than politics. They can invest in unpopular and un-photogenic but important interventions, such as access to safe abortions, or funding statistical agencies or public policy. They can take risks, admit failure and learn.

So why am I not a great fan? Because in my experience few foundations do any of these things. Some foundations (you know who I mean!) obsess far more about their brand and public relations than any public donor I have ever encountered, even though they don't depend on public support for any of their funding. Foundations typically employ staff from the same talent pool as donor agencies, so they don't actually bring any particular business experience or expertise. They appear to me to be just as likely to make decisions based on hunch and emotion as any public donor. They seem just as risk averse, often more so, and just as unwilling to identify and learn from failure. On the whole, I just don't see foundations making the unique contribution to the development ecosystem that Mike Green and Matt Bishop rightly say they could.

That's why I am so glad to be here with the leadership of CIFF this evening. I see in CIFF, in its values, its leadership and its staff, a foundation that is willing to be different, and live up to the promise.

I have been struck by the occasional criticism of CIFF for being slow to get going with spending its money, which is in part because of your insistence on a more rigorous, evidence-based approach. In my view that's an interesting insight not into CIFF but into the attitudes of the aid industry more generally, which still tends to judge organisations not by what they achieve but by how quickly they get money out of the door. I am glad to see CIFF continuing to invest its assets and build its future capacity to spend more money well, rather than fritter it away on ill-thought out projects out of some misplaced desire to spend money quickly.

I very much hope that CIFF will stick to its values, continue to insist on rigorous evidence as it scales up, and so help drive change through the development system.

One reason I am confident CIFF will do a brilliant job is Michael Anderson. I want to tell you all a story about him.

About ten years ago, before Michal and I both had gray hairs, when we were youngish DFID officials, we were both at a retreat for the DFID leadership. These were pretty ghastly affairs, and this was no exception, but at this particular retreat we had quite fun 'buzz talks' between the formal agenda items. These were a kind of "open mic" session at which anyone could grab the microphone and speak for two minutes about anything that was on their mind. What was good about this was that you could say what you wanted and it was all off the record. So half way through the second day, Michael grabs the microphone. I am sure he did not expect that someone would remember what he said ten years later and quote it back to him.

So Michael stands up in front of all these DFID leaders and says that the National Institute for Health and Clinical Excellence, NICE, had been established to test rigorously the value for money of different interventions in the health service and make sure that scarce resources were being used in the NHS to have the maximum impact, and it was proving a huge success. So why on earth, Michael asked, was there no National Institute for Development Effectiveness doing the same for development – indeed, why wasn't there an International Institute?

I didn't know then that Michael would also turn out to be an outstanding diplomat, manager and leader, but I do remember knowing then that this was someone with a commitment deep in his bones to the need to understand evidence about effectiveness and value for money.

It is this idea of using rigorous evidence to take success to scale – the idea which Michael was talking about ten years ago – that I want to develop a little this evening. I am as convinced now as I was then that this is the right approach, but I think I understand it a little more now and I understand more about what it means for an organization like CIFF.

NICE primarily evaluates medicines or procedures. These are relatively simple, reproducible inputs in healthcare. We can be reasonably sure that if a drug, on average, provides benefits cost effectively for sample of people with a particular condition, then it will do so if we roll it out at scale.

The analogue in development is reproducible inputs such as vaccines, Insecticide Treated Bednets, wells, textbooks, microfinance programme, fortified food or contraceptives. We can and we should use rigorous impact evaluations to judge whether these inputs are effective at improving the quality of people's lives and accelerating development; and we should look carefully at the cost effectiveness of these interventions to decide which are worth pursuing and which are too expensive or ineffective.

But there is an important difference between what NICE is doing and what we are trying to do in development. When NICE approves a new drug, or advises against a surgical procedure, there is an institutional mechanism in place, namely the NHS, to deliver those inputs to patients. There are GPs who react to guidelines, who change their prescribing behaviour; and surgeons to change their repertoire of surgical procedures.

In development, those corresponding institutions are largely absent. Roughly speaking, what it is to be a developed country is to have institutions, whether government or private sector, that can deliver medicines or fresh water or food or textbooks to people who need them. If developing countries could do these things, they wouldn't be developing countries.

And so in development our challenge is not just to know which simple inputs are the most cost effective, but also to know how to reach people with those inputs.

Consider the efforts to eradicate Polio. In one sense it is easy: we have an effective and very cheap oral vaccine. It is clearly, demonstrably fantastic value for money to eradicate Polio by vaccinating every child against Polio, so savings millions of lives and preventing disability for millions of people in the future.

But as we all know, eradicating Polio isn't so easy. The scaling up problem is not buying more vaccine but reaching everyone with it. Part of that is logistical – the supply networks and cold chains; keeping track of which communities have been vaccinated; preventing fraud and corruption. Part of that is to do with public attitudes to vaccination and awareness of the health benefits. Part of that is to do with the problem of free riding – it is perfectly rational to want everyone else to be vaccinated against Polio, but not to get your own children vaccinated which may cost you time and you fear it might put them at risk. The Taliban killed three aid workers in Pakistan earlier this year, because they believe that vaccination programmes are an international cover for spying. I was reading today about the progress that is being made on vaccination in Afar region in Ethiopia after the regional Islamic Affairs office convened a meeting of 40 Islamic scholars there last October, and they gave their support to vaccination. Vaccinating hundreds of millions of people turns out to be a complex problem.

And when you think about it, almost everything we are trying to do in development has many of these characteristics. If you think about how to increase access to family planning – well you have exactly the same challenges not only of logistics but of education, trust and power that inhibit the eradication of Polio.

So the question is, can we use rigorous evidence to take successful models to scale when we are trying to solve complex problems of this kind. I suggest we can, but we have to adapt what we mean by taking to scale.

There was a study last year which illustrated this problem superbly in a paper by Bold, Kimenyi, Mwabu, Ng'ang'a and Sandefur at Oxford. Rigorous randomised controlled trials by NGOs had demonstrated that using contract teachers had improved test scores in trials in Western Kenya and parts of India. In the study, this intervention was replicated by NGOs and the government across all Kenyan provinces. What this randomized study found was that when the government hired more contract teachers, this had no effect at all on test scores, whereas when the NGOs hired them, this had a big and statistically significant positive effect. So what we learn from this is that intervention is not, as we thought, the provision of a contract teacher, but an intervention in a particular delivery chain. If this delivery chain involves a government Ministry with limited implementation capacity or which is subject to various political pressures, the system will respond differently than it would if apparently the same input is delivered by an NGO. So when we think about scaling up, we have understand that the intervention we are taking to scale is not only a particular drug, or the provision of supply teacher, but possibly a whole delivery chain which is very difficult to reproduce.

When we think about scaling up entire delivery chains, and not just the supply of particular commodities, you start to wonder about whether it makes sense to talk of evidence-based best practice. Perhaps these kinds of problem cannot be solved by rolling out solutions but by finding solutions locally. Back in the last century, when I moved from the Prime Minister's office in Downing Street to DFID, I arrived with a lot of scepticism about the idea that prevailed in DFID – and which still does - that development solutions had to be locally-owned and country-led. I could see of course that it was going to be easier to make change happen if existing institutions were not actively resisting what you were doing, but it seemed to me that there is enough knowledge and evidence about the value of markets, or of property rights, or civil liberties, or vaccination, or sending girls to school – that these were good approaches to pursue irrespective of whether everyone in a country was signed up to them. I wondered if all the talk about country ownership was just a form of political correctness. Fortunately I didn't voice this at the time, because I now think I was largely wrong.

There still hasn't been a lot of proper work or evidence specifically about what is meant by country ownership. But there is an important strand of research which helps us to understand better how institutional change, and development, happens in practice. There is a 2009 paper by Adler, Sage and Woolcock which discusses Indonesia and Cambodia and introduces the idea of a development being a series of 'good struggles'. The idea here is successful institutions emerge from a process of internal political contestation that cannot be bypassed by transplanting best practice from elsewhere

There are at least four reasons why this process of struggle might be important.

- First, the struggle leads to subtle changes in the solutions so that they fit the context; and these changes enable solutions to fit their environment, without which they do not function properly.
- Second, people take time to learn. Just because you have an MBA doesn't mean you can run a company. Having Roger Federer show you how to serve a tennis ball wouldn't mean you can do so right away. Other than for simple tasks, most of us have to learn by doing. Establishing habits may require repetition and practice, both for individuals and for organisations.
- Third, the struggle confers legitimacy. Michael Woolcock points out that a careful lawyer could have drafted the Good Friday Agreement (which brought peace to Northern Ireland) in a few hours: so why did there have to be so much bloodshed and anguish? Why were all-night negotiations needed to get an agreement? Perhaps the process of compromising – of give and take, of testing limits and building trust – is a pre-requisite for all parties to accept the compromise as the best available.
- Fourth and finally, systems co-evolve. This is an idea from complexity theory, which I'd be happy to talk about over a drink some other time; but the point is this: individual institutions do not operate in a vacuum. Each organisation is in a process of evolution, shaped by an external environment which includes other institutions which are themselves evolving. Particular organisations cannot jump ahead of this if the environment they need to authorise and support them is not also evolving.

Lant Pritchett, one of my colleagues at CGD and a Professor at Harvard, concludes that we have our development paradigm backwards. Instead of thinking that creating capable organisations will deliver results effectively, perhaps successful organisations are the *consequence*, not the *cause*, of success. Capability in formal organisations is what happens when successful folk practices, which evolve out of years of struggle and adaptation, are consolidated into formal processes.

All that makes the idea of 'scaling up' successful, evidence-based interventions rather more elusive than we might first think. According to Lant we should not be trying to build successful organisations but rather to create instances of success from which effective organisations can emerge.

And that is where a foundation like CIFF comes in.

I said at the beginning that I would talk about two recent empirical studies that might shed light on how CIFF can manage development programmes effectively, so let me tell you about them now. As far as I know, these are the only studies of their kind looking at this question, partly because of lack of data. Neither of them has yet been published in a peer-reviewed journal, but they are both rigorous, well done pieces of work.

The first is a study of the evaluations of 10,000 aid projects over the last ten years from nine different development organizations. In this paper Dan Honig, from Harvard University, looks at whether different kinds of projects have been successful in different contexts, and he looks at the impact of organizational devolution within aid agencies. He takes all this data and does some regression analysis to try to determine the factors that affect the success of aid projects.

Dan finds, as you would expect, that aid projects are much less likely to succeed in complex or fragile environments, such as in post-conflict countries, and that more complex projects are less likely to be successful than simple projects. So far, you don't really need a Harvard academic to tell you this. The interesting part of his findings is that aid agencies that allow a large degree of autonomy to their people on the ground and in implementing agencies see a much lower decline in performance for projects in a complex environment than do agencies which exert a higher degree of monitoring and control.

According to Dan's numbers, USAID projects scores are about 20% worse in fragile countries than in more simple environments. But if USAID had the same amount of organisational autonomy as DFID, Dan's results suggest their project success would fall by only about 2% when they work in fragile countries.

So Dan Honig's paper confirms what you might expect: the more likely it is that things will change in unexpected ways, the more important it is to have power and decision-making sit with the people who can see that change coming and respond to it.

The second paper I want to tell you about was written by Imran Rasul and Dan Rogger from University College London. They have assembled an extraordinary dataset of 4,700 public sector projects in Nigeria. They have hand coded independent assessments of the projects' completion rate and quality, and the complexity of the project; and they have conducted a rigorous survey to quantify the management practices of the 63 different organisations responsible for those projects.

Rasul and Rogger also find that more complex projects have lower success rates than simple ones, which is what you would expect. And they also find a strong statistically significant effect from schemes to create incentives for the bureaucrats and to measure performance – but these effects are *negative*, not positive, and much more negative for complex projects than for simple ones.

Rasul and Rogger have some quite detailed information about organizational incentives, and so they are able to provide some evidence about what matters. In summary, they find that freedom to adapt and respond improves results for most programmes but most especially for complex projects. They find that incentives schemes make very little difference either way – presumably because the organisations carrying out the work are motivated by intrinsic motivations. And they find that **performance monitoring has a significant negative effect on results**.

Rasul and Rogger show that there is a modest complementarity between incentives and autonomy. They get a correlation coefficient of about 24%. One interpretation of this is that where organisations are able to measure results, they are more likely to be willing or able to grant autonomy to the implementing agents. The important thing in their data, however, is that it is the autonomy, not the results measurement, which is bringing about the improvement.

It follows that the results agenda is likely to improve project effectiveness when it is used *instead of* micromanagement of inputs and processes, but likely to make little or no difference if it is used *on top of* that micromanagement, as has often been the case with official aid agencies in recent years.

Now these are only two studies, and I am sure I am telling you about them because they support my view of the world, so you have to aim off a bit for confirmation bias.

I hope I would be the last person to suggest that you should base your entire organizational strategy on two studies which haven't yet even appeared in a journal. But they do give some empirical support to a set of ideas that are gaining ground in development thinking. I think they are also consistent with the way we would think about investment and scaling up in business.

So what can we conclude from all this?

First, it is essential that we **continue to examine rigorously the impact and cost effectiveness of different interventions**, so that we make well informed resource allocation choices.

Second, in development the challenge is usually not the absence of finance for these inputs, but the **absence of effective delivery mechanisms**. These are much more complex problems to solve than simply rolling out more products through existing systems.

Third, there are good reasons to think that these complex problems **have to be solved through a process of local struggle**. You cannot transplant the answers to these complex challenges in the same way as you can import the drugs or bed-nets you want to distribute.

Fourth, solving complex problems in complex environments is **more likely to succeed when aid agencies and implementing organisations have more autonomy**. Conversely, high levels of project monitoring and control make success much less likely.

Fifth, **results measurement can and should be the basis of greater autonomy**. It is perfectly reasonable for funding organisations to allow autonomy only to the extent that they can track results. There are some organisations – and I fear DFID is one of them – that have layered results management on top of all the existing monitoring and micromanagement schemes, and so reduced rather than increased effectiveness, especially when it comes to solving complex problems in complex situations.

I know more about development than I do about managing a hedge fund, but I would hazard a guess that most of this is common sense to a business investor. As I understand it, a large part of being an excellent venture capitalist is identifying great teams with great strategies, and giving them a sufficient degree of latitude to get on to deliver. Activist investors will, when necessary, push through a change of leadership or a change of strategic direction, and sometimes pull the plug, but even the best investors are unlikely to get good results overall if they try to micromanage every decision.

People often make parallels between development and business. But it is important that we draw the right parallels. **Too often we think of scaling up in development like rolling a new product line across an existing series of shops. That's the wrong model. Scaling up in development is more like building a series of separate businesses from scratch, each in a different market**. There is a skill and expertise in providing the investment and support, autonomy, incentives and measurement to build new businesses. This is the know-how that CIFF can bring to the development sector, so helping to solve a set of problems which has largely defied conventional aid agencies that have, for a variety of reasons, tried to deal with complexity through control, which is for a variety of reasons likely to be self-defeating.

From where I am sitting, it seems to me that CIFF has the right priorities, attitudes, leadership and staff to be relentlessly rigorous about making evidence-based decision, to focus on value for money, and to lead the way in the development industry to ensure that we have the right combination of focus on results and autonomy to enable governments and implementing agencies to find solutions to the complex, but hugely important, challenges they face.

Owen Barder
Center for Global Development
May 2014.