CENTER FOR GLOBAL DEVELOPMENT **ESSAY**

How Feedback Loops Can Improve Aid (and Maybe Governance)

Dennis Whittle

August 2013

www.cgdev.org/publication/how-feedback-loops-can-improve-aid

ABSTRACT

If private markets can produce the iPhone, why can't aid organizations create and implement development initiatives that are equally innovative and sought after by people around the world? The key difference is feedback loops. Well-functioning private markets excel at providing consumers with a constantly improving stream of high-quality products and services. Why? Because consumers give companies constant feedback on what they like and what they don't. Companies that listen to their consumers by modifying existing products and launching new ones have a chance of increasing their revenues and profits; companies that don't are at risk of going out of business. Is it possible to create analogous mechanisms that require aid organizations to listen to what regular citizens want—and then act on what they hear?

This essay provides a set of principles that aid practitioners can use to design feedback loops with a higher probability of success.

The author thanks his colleagues at CGD for their comments on an earlier presentation of these concepts and is particularly indebted to William Savedoff for key insights. Charles Kenny sponsored the work with the center, and early conversations with him sparked many of the ideas presented here. Lawrence MacDonald suggested the form of this essay. All errors of fact, logic, and judgment are solely those of the author.

The Center for Global Development is an independent, nonprofit policy research organization that is dedicated to reducing global poverty and inequality and to making globalization work for the poor. CGD is grateful for support of this work from its funders and board of directors.

Feedback loops:

a. "The section of a control system that allows for feedback and self-correction and that adjusts its operation according to differences between the actual output and the desired output" (American Heritage Dictionary, 4th edition, updated 2009).

b. "Mechanisms that allow regular citizens to influence both the selection and quality of development initiatives implemented by aid providers and government agencies" (author's definition).

"AH, NON—PAS COMME ÇA!" my new colleague Jean-Luc said sharply as he reached down and yanked out the rubber seedling. He held it up in front of the trembling farmer's face. "Bapak, tiga meter!" he growled. Three meters—that was the optimal spacing for planting rubber trees. Not the one and a half meters that he had just measured. Jean-Luc marched down the row of new plantings, yanking out every other one, and I watched as the farmer's face grew more and more afraid. Finally, Jean-Luc said to me and the Indonesian official present, "Let's go," and we got into the jeep and roared off down the road, back toward the capital of Jambi, a district on the Indonesian island of Sumatra.[1]

The date was late 1987, and this was my first "mission" to Indonesia. I had joined the World Bank only a year before and spent my first six months helping negotiate a structural adjustment credit in Niger, where my main job was to tell government officials how to improve their water and sanitation systems. As I prepared for that earlier job, I was initially terrified and overwhelmed: I knew almost nothing about water and sanitation, and I had never been to Niger. Fortunately, I found that the bank had hired a well-known engineering firm from France, and it had left behind numerous reports filled with all sorts of analyses. It turned out to be pretty straightforward just to read the reports, harvest the recommendations, and tell the government what to do in order to receive our low-interest structural adjustment credit worth tens of millions of dollars.

Having learned the ropes in Niger, I felt well prepared on that first trip to a farmer's field in Indonesia. I had read lots of reports from rubber experts, many of whom had spent the previous decades running plantations in places like Malaysia and Thailand and even Vietnam. It was clear that plantations could optimize the rubber yield per hectare by spacing the rubber trees at specific distances and applying fertilizer and pesticides at well-established

[1] Names in the stories in this essay have been changed.

rates at certain times of year. Building on this knowledge, in the 1970s the World Bank lent hundreds of millions of dollars to finance the planting of huge areas of Indonesia with tree crops, including rubber. The early projects had been implemented through state-owned plantation companies, which typically owned plots in the thousands of hectares. The Indonesian government was so pleased that it had asked the bank to begin designing projects to help small farmers grow rubber, and the bank responded enthusiastically, with several million more dollars of financing.[2]

These new smallholder projects recruited a lot of farmers who had never grown rubber trees. The objective was to get them to convert their fields of food and other crops primarily to rubber. They would be given a bank credit to cover their costs and were expected to repay according to a certain schedule. Our models showed that even after making their loan payments, farmers would have substantially higher incomes than before.

Many farmers, I later learned, were reluctant to be recruited, so they had to be convinced by local government officials, who needed to meet certain implementation targets. Others, like the farmer I had just met, embraced the opportunity to try something new that might increase their income. Yet, since he had no experience with rubber, this farmer was wary of taking too much risk by converting all of his precious land. He had been told that he was required to plant 555 trees on his hectare of land,[3] so he did a reasonable thing: he planted all of those trees on half of his land, and he reserved the other half for his usual food crops and to graze a few small livestock.

I was so outraged by what Jean-Luc did to the farmer that I complained about it to my boss, but he just shook his head and sighed. (My complaint later got back to Jean-Luc, who was furious.) So I went back to my office in Jakarta, put my head down, and burrowed into my spreadsheets analyzing debt repayment schedules, future rubber yields by year, and the anticipated needs for rubber processing plants.

One day a few months later, I got a call from a young English guy who was an advisor to the Indonesian Ministry of Planning. He said that he had just returned from a different province in Indonesia where the World Bank was financing a similar project, though this one to plant coconut.

---

[2] See, for example, the staff appraisal report for Smallholder Rubber Development Project II, January 30, 1985, available at http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/1985/01/01/000009265_3970818102212/Rendered/PDF/multi_page.pdf.

[3] See "Spacing of Plantings," in *Arid Zone Forestry: A Guide for Field Technicians,* FAO Corporate Document Repository, http://www.fao.org/docrep/T0122E/t0122e08.htm.

"You know, the way you are going about this is all wrong," he told me. "Farmers don't want to plant rubber and coconut so intensively because they need more of their land to plant crops that will give them immediate income rather than having to wait several years."

"But we give them credit to hold them over in those intervening years," I replied. "What's the problem?"

"Well, many things," he told me, "including the fact that the credit would not cover all their losses and, even worse, local officials skim off a lot of that money. Plus, sometimes the farmers just like to plant certain things that are familiar to them and that they cannot easily buy in the market. Remember that they are not originally tree crops farmers, and they don't really understand whether the whole thing will work."

"Listen, why don't you come over to my office and I will show you the spreadsheets so you can see what's going on. I am surprised you weren't able to teach the farmer what to do," I told him.

"Well for heaven's sake—I'm not a tree crops expert," he replied. "I know very little about rubber and coconut planting."

"No wonder!" I exclaimed. "What are you, anyway?"

"I'm an anthropologist by training. I try to figure out about people's cultures and practices and how they like to live life."

I shook my head and did what every respectable World Bank economist at that time would do: I made excuses about not being able to meet him and hung up and continued with my spreadsheets and analysis. My colleagues and I later had vigorous and extended debates—including one very public showdown in a large meeting—with the British anthropologist and his fellow advisors[4] about the best approach to smallholder tree crops and other agriculture-sector issues. I was worried at first, because their arguments had a ring of truth, but eventually my colleagues and I prevailed because we held the power of the purse. We were able to lend the Indonesian government hundreds of millions more dollars, and our anthropologist friend was not.

* * *

That encounter between Jean-Luc and the farmer haunted me for years. First and foremost, I could not shake the image of the farmer's face after his seedlings were ripped from the ground. He had been so proud when we arrived, and he looked so confused and scared as we pulled away in our jeep down the heavily rutted roads. And the battle we had with the government advisors had left a bad taste in my mouth. It had initially felt good to be on the

[4] These colleagues included two of the world's leading agricultural economists, from Harvard and Stanford, respectively.

winning side, but many of their arguments rang true, even at the time. My bank colleagues and I had amassed enough sheer firepower to prevail, and I felt like a bully.

For the next several years, I continued to feel ill at ease with the work I was doing at the World Bank, where I participated in or led many projects and studies in several countries. I was keenly aware that I had the luxury of working with the largest aggregation of top development experts in the world. But the way we worked did not seem right. It took me more than a decade to begin to articulate the problem and to understand that *poor information* and *perverse incentives* were the main causes.

Over time, I have grown to believe that regular people should have the most say in selecting which aid initiatives are chosen to help them. In this sense, I have become a proponent of "bottom-up" aid. I am unaware of any rigorous analysis that has been done on this topic, but my own estimate is that aid is currently 20 percent driven by regular citizens and 80 percent driven by experts (of which I used to be one). Even if not exact, my estimate underscores the inherent bias I see in aid planning, and I believe this ratio needs to be reversed. There is of a role for experts, especially relating to the design and provision of public goods at the national and international levels. But experts should drive 20 percent of the agenda, not 80 percent.

The good news is that recent technological breakthroughs are enabling us to dramatically increase our ability to find out what people like the Indonesian rubber farmer really want—and whether they are getting it. Even for top-down projects driven by experts, new technologies can help determine whether the projects are working well or not, and allow for midcourse corrections to help the projects achieve the desired impact. The challenges now are to (a) enable the right flows of information and (b) figure out how to alter the incentives within the system so that we get more of the outcomes we are hoping for.

* * *

The mental model in the development community when I began working in the field in 1984[5] was very top-down and expert driven, and went as follows: The developing world suffers from a severe shortage of both know-how and money. Official aid agencies such as the World Bank, the UN, and bilateral aid agencies need to aggregate the world's best expertise and money—and then deliver them to poorer countries. At that time, there was also a growing number of nongovernmental organizations (NGOs). Though they get money from different sources (from donors rather than taxpayers), in practice many NGOs operate with a similar mind-set: "We know what people need, we know how to deliver it, and we are here to give it to them."

There are two major flaws in this approach. First, we in the development community are not very good at knowing what people need or want. (In addition to my experience at the World

[5] Initially at the Asian Development Bank and USAID in Manila before joining the World Bank in 1986.

Bank, a recent experiment run by GlobalGiving in Kenya showed that even most local experts and implementing agencies were not good at judging what mattered most to the people they were ostensibly serving—see below.) Second, we do not know how to deliver aid initiatives effectively on a consistent basis. Work by Bill Easterly and others has shown that the trillions of dollars spent over the last six decades have had low returns.[6] Even the more optimistic studies conclude that aid has had satisfactory returns only under limited circumstances.

The shortcomings of the top-down aid system reflect a broader problem with relying solely or primarily on experts. Philip Tetlock's celebrated recent work suggests that experts' predictions about the outcomes of complex situations and initiatives are poor—barely better than would be achieved by flipping a coin, and worse than would be achieved by applying various simple rules of thumb.[7] While most of Tetlock's work does not involve development experts, the predictions that experts in his studies were trying to make were analogous to those that development experts make: if we do X, Y, and Z, then the outcomes A, B, and C will be realized.

During my time at the World Bank, most of my colleagues and I realized that many of our projects did not achieve the outcomes we had predicted. Unfortunately, we often had inadequate information to remedy the situation during implementation—not to mention strong incentives to avoid emphasizing or surfacing any shortcomings. The prevailing approach was to spend about a year or eighteen months designing and appraising a project, then to spend five years implementing it, and then finally to evaluate the project. Although we did "supervise" projects during implementation, we could typically spend only a day or two at a sample of project sites. So we had to depend heavily on government implementing agencies for information.

For example, in the case of the rubber projects in Indonesia, the government agency not only reported back to us on farmers' attitudes toward the initiative ("The famers like planting rubber—it makes them better off!") but they also decided which sites we should visit. (And when we asked to see certain sites, they frequently demurred, saying the road was washed out or the site manager was at a training course.) Only later, during the formal evaluation, did we discover that many tracts of land that we had financed had never been planted at all[8]—and that many unhappy farmers had failed to pay back their loans, saddling the government and banking system with big losses.

---

6 William Easterly, The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics (Cambridge, MA: MIT Press, 2001); and William Easterly, The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good (New York: Penguin, 2006).

7 Philip E. Tetlock, Expert Political Judgment: How Good Is It? How Can We Know? (Princeton, NJ: Princeton University Press, 2005).

8 Note that we at the World Bank were not the only ones unhappy about the "ghost plantings." The government's own planning bureau was furious at the estate crop ministry for allowing this to happen through

The problem was made worse by the institutional constraints we faced. World Bank budgets for project preparation, supervision, and evaluation were increasingly standardized across projects and across countries. As project managers, we faced clear incentives—namely, to get as many projects approved by the board as possible, and then to get them implemented in the allotted time frame; within the allotted budget; and without protests by the intended beneficiaries, public relations problems from local or international NGOs, or complaints from the government.

As long as we responded to these clear incentives, we got promotions and steady pay raises. But if we spent too much money monitoring the projects and slowed implementation down to try to correct for problems, our projects could be downgraded, and this would generally be reflected on our performance evaluations. If for some reason we were put in charge of a project with egregious and unconcealable problems, the rational thing to do was to ask for a transfer to a different project or even to a different country. Otherwise, we might be forced to negotiate with the government to formally redesign or even close the project prematurely, which was messy and time consuming, and (most damaging) would reduce the number of new projects we could prepare and send to the board.[9] Further, the incentives were highly asymmetrical; there was no upside to making a project go better than anticipated.[10] If problems surfaced later, during the formal evaluations, it was not a huge concern because project managers had often moved on to a different country.[11]

The enemy of smooth project implementation was often citizen voice. It turned out that the poor rubber farmer I described above had relatively minor complaints compared with some others affected by aid projects in Indonesia. Far more serious were the complaints that thousands of villagers had about being forcibly relocated to make way for a dam that was being built. Some villagers were forcibly "transmigrated" to outlying islands with no physical or social infrastructure, in theory to create a better life for themselves. A large number of

---

either oversight or fraud. But by then, enough time had elapsed that it was hard to establish who within the ministry was to blame, and no one was held accountable.

9 For a nuanced and comprehensive description of organizational incentives inside the World Bank, see David A. Phillips, Reforming the World Bank: Twenty Years of Trial —and Error (New York: Cambridge University Press, 2009).

10 In fact, I was criticized once because my project achieved its primary objectives too soon in the project cycle! For an excellent discussion of the asymmetry in incentives within aid-funded initiatives in general, see Lant Pritchett, "It Pays to Be Ignorant: A Simple Political Economy of Rigorous Program Evaluation," The Journal of Policy Reform 5, no. 4 (2002): 251–69.

11 For a classic analysis of incentives within a typical bilateral agency, see Elinor Ostrom, Aid, Incentives, and Sustainability: An Institutional Analysis of Development Cooperation: Summary Report (Stockholm: Sida, 2002).

these people were very unhappy about the lack of roads and functioning markets, schools, and health clinics.

In each case, the bank assigned seasoned veterans to work together with the government to keep the situation from getting out of hand and to keep projects on track, with some refinements but rarely major modifications. The goal was, in the words of my colleague who had been assigned to manage the dam project, "to keep people's complaints down to a dull roar, and to avoid big stories, especially in the local newspapers, that might rile people up further."

Over the ensuing years, I noted a gradual change in the type of people hired by the bank and other aid agencies. Increasingly I met new staff who had spent lots of time in the field and came to the agencies with a genuine understanding of and concern for real people. They came into their jobs with great enthusiasm, and whenever I met them it gave me glimmers of hope.

Alas, with few exceptions,[12] these promising new staff members soon learned that it was impossible to maintain enthusiasm and commitment in the face of the nearly overwhelming effort required to process projects through the World Bank system. Ironically, the bureaucratic load typical of the 1980s was ratcheted sharply upward in the 1990s by a number of well-intended "safeguards" related to environmental, social, gender, and even procurement matters. These safeguards generated numerous requirements and led to even more studies, paperwork, and internal clearances—reducing even further the time that staff were able to spend in the field talking to real people.

The tragedy of this system is that most aid workers (certainly most of my colleagues at the World Bank) start out wanting to improve people's lives. They do the best they can, relying on the best information they can gather, to make a positive difference. Yet over time, they get ground down psychologically (and even physically) by the organizational constraints and incentives they face, and after a few years many of them lose touch with why they started this line of work in the first place.

Just before I left the World Bank in late 2000, I did an informal poll of many of my colleagues there, asking them, "What proportion of your energy do you feel you are able to use in service of actually helping make the world a better place?" Their answers clustered tightly around 25 percent. A typical response was "You know, I have not even thought about that question for so long. Most of my life and attention here are taken up by the need

---

12 The World Bank's Scott Guggenheim, who helped design the Kecamatan Development Program in Indonesia, was a rare exception. To help escape the normal pressures, he raised external funding and moved his entire team out of the World Bank's main offices in the Indonesian stock exchange building and into separate space in a different part of town. This move required not only exceptional entrepreneurial skill on his part but also the support of an unusually empowering director, Mark Baird. See http://www.denniswhittle.com/2006/11/when-official-aid-works.html.

to write reports, attend meetings, get clearances, and book travel. I rarely get a chance to pause and ask whether it all makes sense, or whether I am making a real difference."

* * *

So how can we move forward beyond this outdated and ineffective mental model for development aid? There are no silver bullets, but some key principles are coming into focus.

In the last few years, there has been great excitement about the use of randomized controlled trials (RCTs) in development.[13] RCTs were hailed as a way of overcoming an "inadequate understanding of poverty" and as "radical rethinking of the way we fight poverty."[14] The idea is that by adopting the same standards of rigor applied in science, we would be able to find out "what works" and apply it around the world.

While the RCT movement is probably helping improve the rigor of ex post evaluations of aid projects, its value in helping improve development outcomes on a wide scale is only modest. Despite the claims that RCTs drive medical advances, recent studies have questioned whether this is the case, for two reasons. First, medical companies have had exceptional difficulties replicating (in standard laboratory conditions) the results of preclinical RCT-based trials for interventions like cancer drugs. Even under laboratory conditions, scientists at the drug companies Amgen and Bayer, for example, were able to reproduce the results of only 11 percent and 21 percent, respectively, of the RCT-based trials they studied.[15]

Second, the drugs are administered under varying clinical conditions and to patients whose body chemistry differs significantly. According to a paper by Margaret Eppstein and colleagues, though "many consider multicenter randomized controlled trials to be the gold standard of evidence-based medicine … results are often inconclusive or may not be generally applicable due to differences in the contexts within which care is provided."[16]

These drawbacks to RCTs in medicine echo the criticisms leveled at RCTs for development. Angus Deaton provided an early technical analysis of how the results from RCTs were unlikely to be transferable to different contexts.[17] Anyone who has managed aid projects

---

13 See, for example, Abhijit V. Banerjee and Esther Duflo, Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty (New York: PublicAffairs, 2011).

14 "Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty," Poor Economics, accessed May 31, 2013, http://www.pooreconomics.com/about-book.

15 See Florian Prinz, Thomas Schlange, and Khusru Asadullah, "Believe It or Not: How Much Can We Rely on Published Data on Potential Drug Targets?," Nature Reviews Drug Discovery 10, no. 9 (2011): 712–13; and C. Glenn Begley and Lee M. Ellis, "Drug Development: Raise Standards for Preclinical Cancer Research," Nature 483 (2012): 531–33, http://www.nature.com/nature/journal/v483/n7391/full/483531a.html.

16 Margaret J. Eppstein, J. D. Horbar, J. S. Buzas, and S. A. Kauffman, "Searching the Clinical Fitness Landscape," PLoS ONE 7, no. 11 (2012): E49901, doi:10.1371/journal.pone.0049901.

17 Angus Deaton, "Instruments, Randomization, and Learning about Development," Journal of Economic Literature 48, no. 2 (2010): 424–55.

realizes that there is a huge number of design and implementation parameters—and that it is maddeningly difficult to know which of these makes the difference between success and failure. In the preparation phase, we tend to give a lot of weight to the salience of certain factors, such as eligibility criteria, prices, technical features, and so on. But during implementation, we realize that a thousand different factors affect outcomes—the personality of the project director, internal dynamics within the project team, political changes in the local administration, how well the project is explained to local people, and even bad weather can have major effects.

Development initiatives are not necessarily *complicated,* but they are *complex*.[18] Complicated systems have many parts and aspects, but the outcomes can be predicted accurately if the initial conditions are known, since the parts themselves interact in a consistent and often linear way. Building a bridge over a wide river is complicated, as is building an airplane and most other engineering challenges, but experts are able every day to build bridges and airplanes that work reliably. In contrast, *complex* systems such as health care are very often completely unpredictable, and "infinitesimal causes can have enormous consequences"—a phenomenon known popularly as the "butterfly effect."[19] These complex systems—often involving human behaviors and interactions—are notoriously difficult to predict, much less control.

Lant Pritchett and Justin Sandefur have extended Deaton's analysis to show that RCTs can be expected to have little "external validity" in development projects (i.e., the results are not transferrable beyond the initial context of the study).[20] Along with his colleagues Salimah Samji and Jeffrey Hammer, Pritchett goes on to propose a new method of rapid and ongoing iteration of project design during implementation.[21] The idea is to start with a design considered reasonable, but to focus a much greater proportion of available resources on finding out how well the design is working in practice and then refining it as the project proceeds.

Michael Woolcock argues that the more complex the initiative, the less likely RCT results are to be applicable across different contexts. The elusive idea of "best practices" is valid only in projects with "low causal density"—generally those whose outcome does not depend heavily on human behavior. He argues that using case study methodologies is critical to finding out

---

18 For a good overview of complexity and its implications for emergent systems, see John Gribbin, Deep Simplicity: Bringing Order to Chaos and Complexity (New York: Random House, 2004).

19 Horgan, John, "Can Engineers and Scientists Ever Master 'Complexity'?," Cross-Check (blog), Scientific American, December 10, 2012, http://blogs.scientificamerican.com/cross-check/2012/12/10/can-engineers-and-scientists-ever-master-complexity/.

20 See Lant Pritchett, "Development as Experimentation (and How Experiments Can Play Some Role)" (unpublished manuscript, July 2011); and Lant Pritchett and Justin Sandefur, "Context Matters for Size: Why External Validity Claims and Development Practice Don't Mix" (unpublished manuscript, July 18, 2013).

21 Lant Pritchett, Salimah Samji, and Jeffrey Hammer, "It's All about MeE: Using Structured Experiential Learning ('e') to Crawl the Design Space," WIDER Working Paper No. 2012/104, United Nations University–World Institute for Development Economics Research, Helsinki, 2012.

what works in different situations. He suggests an approach similar to the one Eppstein and colleagues[22] endorse for better medical research, namely "making it up as you go along: you work with others and learn from collective experience to iterate your way to a customized best fit."[23]

It is true that RCTs can provide rigorous evidence of the impact of certain interventions under certain circumstances. But in the end, randomized trials don't provide the tools for us to "radically rethink the way we fight poverty." Excessive faith in RCTs is in fact likely to reinforce the same top-down approaches that have had such poor results so far. Such approaches lend themselves heavily to initiatives in which experts determine the desired outcomes ex ante, marginalizing the voices of the people they are supposed to help. Instead of giving an incentive to local people to work together to forge solutions, they risk creating global "best practices" that ignore local knowledge. The tendency toward cookie-cutter prescriptions inhibits the emergence of local networks of problem solvers who experiment repeatedly until they find approaches that work.

* * *

Fortunately, there are new models that emphasize participation, accountability, and feedback. Though these models are still in their formative stages, what many of them have in common is an assumption of rich and timely feedback loops that allow—or even require—implementing agencies to iterate constantly. Daron Acemoglu and James A. Robinson argue in their recent book *Why Nations Fail*[24] that developing countries remain poor because their institutions are exclusionary and do not reflect the voices and needs of much of society. They argue that as long as elites in these countries maintain near-monopolistic control over public institutions, there will be only slow improvements in the quality of life for most people. Commenting on this book, Owen Barder notes, "If we think of politics as an endogenous characteristic of a complex system, then perhaps we have more hope of accelerating development by trying to tweak the internal feedback loops, and so shaping future system dynamics, than by offering exogenous solutions from the outside."[25]

Acemoglu, Robinson, and Barder together suggest that helping citizens in developing countries have better voice through effective feedback loops would have high returns by increasing political inclusion and reducing the political monopoly of the elites. A key question is how aid projects can encourage the formation of these feedback loops without eliciting a backlash from the very elites whose power will be reduced by them.

---

22 Eppstein et al., "Searching the Clinical Fitness Landscape."

23 See Michael Woolcock, Using Case Studies to Explore the External Validity of "Complex" Development Interventions (Washington, DC: World Bank, forthcoming).

24 Daron Acemoglu and James A. Robinson, Why Nations Fail: The Origins of Power, Prosperity, and Poverty (New York: Crown, 2012).

25 Owen Barder, "It's the Politics, Stupid," Owen Abroad (blog), June 1, 2013, http://www.owen.org/blog/6752.

* * *

The good news is that a number of experiments have been launched over the past few years to pilot new ways to collect and use feedback. The International Aid Transparency Initiative[26] and Publish What You Fund[27] have made information on official aid available on a much wider scale. The same is true for NGOs, thanks to the BRIDGE (Basic Registry of Uniquely Identified Global Entities) project, launched in mid-2013 by GlobalGiving, GuideStar, the Foundation Center, and TechSoup Global.[28] This project, partly funded by the Bill & Melinda Gates Foundation and the William and Flora Hewlett Foundation, is creating common data standards similar to IP addresses on the Internet that will enable people to compare data and information for about three million NGOs worldwide.[29]

Together with my colleagues David Boyd and Anna Sosdian (with key inputs from Marc Maxson), I recently reviewed a number of these experiments. Some were launched by nonprofits, some by local governments, and others by official aid agencies. Some pilots failed completely, a few were very successful, and most had promise but lacked one or more of the elements required to either gather enough data or create pressures on implementing agencies to remedy the situation.

A small sample of these includes the following:[30]

- **World Vision (Uganda).** Thirty school districts in Uganda used the Participatory Community Scorecard (PCS), developed in collaboration with World Vision. The PCS enabled communities themselves to develop the schools' performance criteria, which the communities would monitor. For thirty other schools, experts defined the performance criteria, which communities would monitor. The schools for which the *community* defined performance criteria showed a .19 standard deviation increase in test scores, moving the average student from the 50th to the 58th percentile in performance; increased pupil attendance by 8–10 percent; reduced teacher absenteeism by 13 percent; and cost a total of $1.50 per student. The schools for which *experts* developed the criteria showed *no increase* in student test scores.[31]

---

[26] http://www.aidtransparency.net/.

[27] http://www.publishwhatyoufund.org/.

[28] Internet-based aid platforms such as GlobalGiving.org are growing and multiplying. GlobalGiving alone has already facilitated about $120 million in funding for eight thousand projects in 130 countries, and its creation has spawned the launch of additional platforms such as Kiva.org, which facilitates hundreds of millions of dollars of micro-loans, and the newer GiveDirectly.org, which enables funders to give money to individual people rather than organizations.

[29] Victoria Vrana, "BRIDGE to Somewhere: A Conversation with GlobalGiving, GuideStar, the Foundation Center, and TechSoup Global," Markets for Good, June 3, 2013, http://www.marketsforgood.org/bridge-to-somewhere-a-conversation-with-global-giving-guidestar-the-foundation-center-and-techsoup-global/.

[30] More analysis of these and other examples is forthcoming at http://www.feedbacklabs.org.

[31] Abigail Barr, Frederick Mugisha, Pieter Serneels, and Andrew Zeitlin, "Information and Collective Action in the Community Monitoring of Schools: Field and Lab Experimental Evidence from Uganda" (unpublished

- **Exposing Corrupt Politicians (Brazil).** In this experiment, the results of audits of city finances were released to the public before elections. Compared with a control group without audits, mayors were seven percentage points less likely to get reelected when these audits showed corruption violations. The effect was more than doubled in towns where the audit results were broadcast on a radio station.[32]
- **Rapid SMS (Malawi).** UNICEF trained local health workers to use an SMS (text message)–based tool that allowed them to report data on each child's health measurements. These data previously took up to three months to be compiled on paper and sent to headquarters, and they were mostly used for reporting and research purposes. Under the new system, it takes only two minutes to enter the data into the phone and transmit them, and the platform responds immediately with tailored advice on nutritional needs for each specific child.
- **CheckMySchool.org (Philippines).** This tool allows anyone—parents, students, teachers, administrators, or NGOs—to report problems at schools, ranging from absent teachers to missing textbooks to broken toilets. Comments and complaints can be channeled through e-mail, SMS, Facebook, or a website and are viewable by the general public. The Department of Education has committed to taking quick action on complaints.
- **Crisis Response Map (Haiti).** After the 2010 earthquake in Haiti, this online platform created by Ushahidi allowed anyone to provide updates or request help using SMS, e-mail, phone, or Twitter. Thousands of reports were submitted, allowing rescue operations to reduce duplication of effort and focus on the areas most affected.
- **GlobalGiving Storytelling Project (Kenya).** With support from the Rockefeller Foundation, GlobalGiving has collected stories from tens of thousands of people in Kenya about what they care most about. In one iteration involving four communities, GlobalGiving asked a panel of sixty-five aid experts and implementing agencies to guess the top priorities for the 2,500 respondents. The expert panel guessed only half of the top six priorities of the community; and only one of the sixty-five experts correctly guessed the single most pressing issue (social relations).[33]

The number of effective feedback loop experiments in aid is still small, and though there are a priori reasons to expect them to improve aid projects, there have been few rigorous statistical evaluations of how well they are working. Especially given our caution

---

draft manuscript, August 2012); and Andrew Zeitlin, Lawrence Bategeka, Madina Guloba, Ibrahim Kasirye, and Frederick Mugisha, "Management and Motivation in Ugandan Primary Schools: Impact Evaluation Final Report" (unpublished draft manuscript, Centre for the Study of African Economies, University of Oxford, UK, October 2011).

32 "Exposing Corrupt Politicians," J-Pal Policy Briefcase, December 2011, http://www.povertyactionlab.org/publication/exposing-corrupt-politicians.

33 Dennis Whittle, "If You Flip a Coin, Can You Be an Expert?," Pulling for the Underdog (blog), September 2, 2010, http://www.denniswhittle.com/2010/09/if-you-can-flip-coin-can-you-be-expert.html.

above about generalizing the results of RCTs, it would be foolhardy to try to draw any conclusions or "best practices." It may be more useful, at least at this stage, to identify some general principles and questions that are likely to inform the design of feedback loops that result in better outcomes.

Our review suggests that asking the following questions during the design stage is likely to increase the chance that the feedback loop will actually result in better outcomes:

1. What information is the feedback loop soliciting?
2. Who is most qualified to provide that information?
3. What incentives do those people have to provide the information? What are the costs and benefits that they perceive?
4. How will people provide the information? In person? Using certain technologies? Will the information be confidential or public?
5. Who are the intended recipients of the information and how will they get it?
6. What specific actions do we want the recipients of the information to take?
7. What incentives (carrot, stick, or both) and capacity do the recipients have to take action? And how will we know action was taken?

As Pritchett and Woolcock argue, the best way to approach these questions is probably just to start with a reasonable hypothesis, and then iterate based on experience. The answer to each question will depend critically on the context, for example the type of project, the sector, and even the type of implementing agency. A few studies that suggest reasonable places to start are beginning to emerge.

With respect to Question 1 above, perhaps the most neglected feedback loop information is simply what people themselves care about most.[34] As the GlobalGiving Storytelling Project illustrated, experts and implementing agencies are often out of touch with what people want. Further, as Ben Olken has shown, direct participation by people simply in *choosing* projects can have a huge effect on improving their satisfaction with outcomes and their political engagement. In some contexts, the impact of giving people a say in the *choice* of which projects are implemented can be even greater than the impact of allowing people to monitor implementation.[35]

The World Vision initiative in Uganda described above suggests further insights into how to answer Question 1. In particular, allowing communities to define the "scorecard" for what gets measured even within a preexisting program (in this case primary education) might be critical to improving outcomes. In contrast, the study found that allowing experts to define the scorecard produced no improvement in outcomes.

---

34 This point is also evident in Deepa Narayan and Raj Patel's book Voices of the Poor (New York: Oxford University Press for the World Bank, 2000).

35 Benjamin A. Olken, "Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia," American Political Science Review 104, no. 2 (2010): 243–67.

With respect to Question 2 above, other work by Ben Olken suggests that community participation in monitoring may work best for projects that do not require technical knowledge. Communities may be good at monitoring whether a school is working well (by observing, for example, whether the teachers are present and whether their children seem to be engaged and learning), but they are less good at monitoring expenditures on road construction, for which top-down audits are better able to evaluate construction techniques and raw material costs.[36]

On the other hand, the work of Gray-Molina and others on hospitals in Bolivia suggests that for more complex delivery systems such as health care, communities may be better at monitoring certain services than rules-based audits or evaluations related to proxies such as competitive hiring and staff supervision practices.[37]

In the Exposing Corrupt Politicians example in Brazil cited above, expert auditors were needed to actually find the information about corruption; regular citizens would have had a hard time doing that on their own (Question 2). In this case, the auditors' incentive to provide the information (Question 3) was their normal salaries and operating budgets—nothing else was needed, although we might anticipate that local politicians will try to exert pressure on them not to report the information in the future. But the feedback loop would not have worked if the information had not been provided to citizen voters using local radio stations as the most powerful dissemination mechanism (Questions 4 and 5). And it was clear what actions the recipients of the information were going to take (Question 6)—they voted against the mayors associated with corruption.

In our initial review of feedback experiments, Question 3—incentives to provide the information—was often a point of failure. In general, there is a cost to information providers, including their time and money (for example, the cost of sending an SMS[38]), and sometimes risk of retribution. In Tanzania, where half the public water points don't work, the organization Daraja created an initiative called Maji Matone to allow citizens to report on their water points. Over six months, only fifty-three reports were received (compared with an initial target of three thousand). Although a formal analysis has not been done, it appears that there were at least three problems. First, men control the mobile phones, while women collect the water. Second, women apparently felt that there was a risk of retribution from local water officials. And third, women felt that their reports were unlikely to generate any remedial action.

---

36 Benjamin A. Olken, "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," Journal of Political Economy 115 (2007): 200–49.

37 George Gray-Molina, Ernesto Pérez de Rada, and Ernesto Yañez, "Does Voice Matter? Participation and Controlling Corruption in Bolivian Hospitals." In Diagnosis Corruption: Fraud in Latin America's Public Hospitals, ed. Rafael Di Tella and William D. Savedoff (Washington, DC: Inter-American Development Bank, 2001), 27–55.

38 Some initiatives have addressed this cost by providing SMS credits to feedback providers.

With respect to action by feedback recipients (Question 7), Tessa Bold and colleagues found that feedback loops worked very differently depending on whether implementation was being done by a government or an NGO.[39] An NGO-led initiative had improved educational outcomes in India and Western Kenya by identifying lagging students and assigning contract teachers to tutor them. Bold and her colleagues replicated the approach throughout all provinces of Kenya, randomizing whether the program was implemented by an NGO or the government. Test scores rose significantly in the NGO-implemented programs but not at all in government-implemented programs. It was not clear whether this difference was due to incentives or capacity constraints, although one can imagine a combination of both factors at work.

There are some reasons to believe that the new online aid intermediation platforms may have greater incentives and ability than traditional aid agencies to promote accountability through feedback loops,[40] but many of the seven questions raised above remain to be addressed with these new mechanisms as well.

* * *

As noted, there is no single answer to any of the above questions; the answers will depend on the context and will result from repeated iteration starting from reasonable hypotheses. One objective of future research should be to help generate these starting hypotheses. To this end, the research could fruitfully address five broad issues:[41]

1. **How do we provide incentives for broad-based feedback?** At a minimum, people need to be technically able to provide feedback, they need to be able to afford it, they need to feel that it will make a difference, and they need to feel that they will not suffer retribution.
2. **How do we know that feedback is representative of the entire population?** In many places there is differential access to cell phones and even in-person meetings. The phenomenon of *elite capture*, whereby powerful local interests exert heavy pressure on elections and other decisions, is equally prevalent. Yet our work to date suggests that these problems are generally lessened by the introduction of new

39 Tessa Bold, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur, "Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education," Working Paper 321, Center for Global Development, Washington, DC, 2013.

40 Devesh Kapur and Dennis Whittle, "Can the Privatization of Foreign Aid Enhance Accountability?" Journal of International Law and Politics 42 (2010): 1143–80.

41 These questions will guide the work of the newly formed FeedbackLabs.org, which was born after an earlier blog post for the Center for Global Development on this issue (Dennis Whittle, "Make a Consumer Reports for Aid," Global Development: Views from the Center (blog), January 15, 2013, http://www.cgdev.org/blog/make-consumer-reports-aid). Founding members of the Labs include Development Gateway, GlobalGiving, Ashoka, Ushahidi, Frontline SMS, GroundTruth, Keystone, and Twaweza East Africa. Feedback Labs will frame the conceptual issues, catalyze experimentation, support initiatives to help make existing feedback loops more effective, and help governments and official and private aid agencies mainstream promising approaches.

forms of feedback loops and that more can be done to ameliorate if not eliminate bias.

3. **How do we combine the wisdom of the crowds with the broad perspective and experience of experts?** Effective feedback loops promise to increase the power of regular citizens in decision making about aid projects that affect them. But these local decisions can often be improved by hearing the perspectives and advice of experts, who frequently have deeper knowledge of specific topics and of how specific approaches have worked in different countries. How can we use new feedback loops to create better conversations between citizens and experts about what investments and services would have the biggest impact on well-being?
4. **How do we ensure there are strong incentives for aid providers, governments, and implementing agencies to adopt and act on feedback mechanisms?** Despite the growing number of feedback pilots underway, these experiments still affect only a miniscule percentage of all government initiatives and aid projects. And a large proportion of the pilots to date have not led to significant changes that improve outcomes. Given the inherent incentives against accountability by big aid agencies, what combination of carrots and sticks will it take to bring about widespread adoption of feedback loops?
5. **What is the relationship between effective feedback loops in aid and democratic governance?** My research started out as an inquiry into how feedback loops could make aid agencies more accountable. But is it possible or desirable to separate feedback in aid from feedback more broadly in governance? Can or will promoting more effective citizen feedback in specific aid or government programs lead to greater citizen voice more broadly?

Conclusion

Promoting strong and timely feedback loops is key to making aid, philanthropy, and government initiatives more effective. Even in top-down programs, benevolent experts and government officials have an interest in knowing how well implementation is proceeding so that they can make midcourse corrections instead of relying on costly (expert) evaluations that come too late. But more broadly, feedback loops can also help us rebalance the way that development programs are formulated and conducted. Though progress has been made in listening to the voices of regular citizens, it is reasonable to guess that development assistance is still 80 percent determined by experts and only 20 percent by citizens. Good feedback loops could reverse this ratio and put the bulk of the decision-making power in the hands of regular people.

In the future, the default model will be that aid officials need to demonstrate (a) why they believe regular citizens actually want each proposed project and (b) how citizen voice will be used to ensure high-quality implementation. Of course, there may be exceptions, for example certain types of policy projects or public goods with free-rider problems, but the burden will be shifted to the aid official to make the case for why citizen voice should not play a major role.

What is the fastest way to bring about this future? As discussed above, there are inherent disincentives for aid agencies and aid workers to seek out and act on citizen feedback. How to overcome these disincentives is a topic for another paper. But one thing I have learned in my nearly three decades of aid experience is that immediate and sweeping mandates rarely work—and sometimes they even backfire by creating new compliance burdens that reduce the time available for staff to address the real issues.

The best approach in the near term is to provide carrots rather than sticks, at both the inter- and the intra-institutional level. Boards of governors of the different aid agencies should increase the resources made available to grantees that demonstrate a commitment to effective feedback loops. And within institutions, senior management should significantly increase the resources dedicated to experimentation and research, using the principles and addressing the conceptual issues sketched out above. Once successful approaches are well established, boards of governors as well as senior management should phase in mandates while providing a virtuous cycle of additional resources and incentives for agencies and staff that implement effective feedback loops.

To hasten this transformation, think tanks and citizen groups can rally public support for agencies that listen to people, and they can shine light on agencies that persist with the old mental model of "experts know best." The Center for Global Development's Commitment to Development Index, the World Bank's Doing Business survey, Transparency International's Corruption Perceptions Index, and Publish What You Fund's Aid Transparency Index are all examples of what might be emulated for feedback loops.

My colleagues and I will be continuing our work to support research, experimentation, and mainstreaming at FeedbackLabs.org. Readers' participation, ideas, and feedback are all welcome on the website.