



NOTE 1: Why Evaluate?

Success depends on knowing what works.

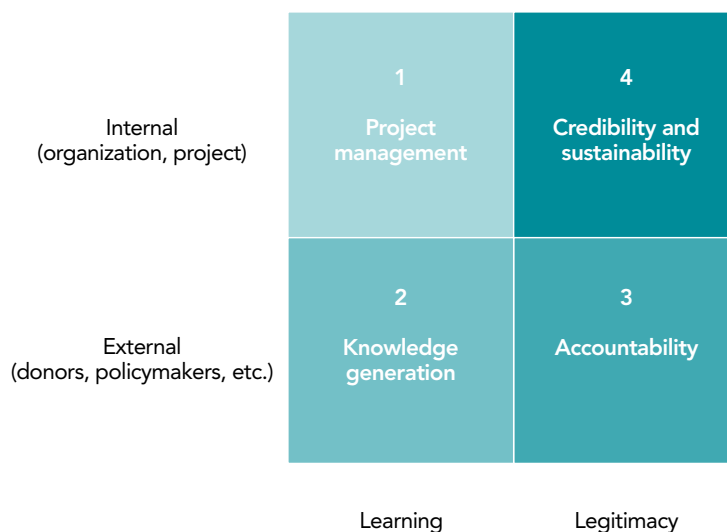
— Bill Gates

The objective of this note is to provide an overview of how individual organizations and the field as a whole benefit from the knowledge acquired from formal evaluation, particularly through impact evaluation. We argue that there are two major purposes of evaluation: learning and establishing legitimacy. For each purpose, there are internal and external audiences (see figure 1.1). Together, they yield four good reasons to conduct evaluations:

- To manage projects
- To generate knowledge
- To ensure accountability
- To strengthen our organization's credibility and sustainability

These are discussed below.

FIGURE 1.1 Benefits of evaluation



Project Management

Youth-focused interventions are inherently complex. Because we are dealing with a dynamic target group in transition biologically, socially, and legally, the interventions we put in place are highly diverse in nature and have outcomes across a range of sectors. Properly evaluating these interventions, albeit challenging, is a crucial ingredient in the recipe for success.

Evaluations allow us to see the true value of our work. Most of us want to know what difference our programs are making in the lives of the young people we serve. Did our project achieve the desired results? Who benefitted more, who less? Evaluations help to answer these and other questions by assessing the relevance, effectiveness, efficiency, impact, and sustainability of an intervention.

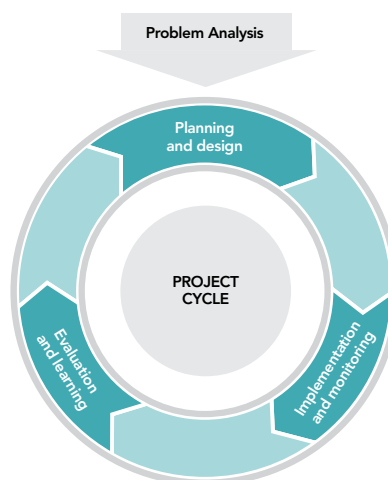
Evaluations foster learning. By assessing the design, implementation, or results of an intervention, evaluations enhance organizational learning. They allow us to identify which parts of our intervention were successful and which ones may not be working as intended. For example, an evaluation may reveal that the amount of training provided during an intervention was insufficient, resulting in low learning, or, on the contrary, was too intensive, overwhelming the students and leading to dropout. Similarly, an evaluation may help us understand unintended consequences of our project, such as an increase in parents' alcohol consumption associated with providing girls with income opportunities.

Evaluations support new planning. Evaluations provide program managers with the information we need to make strategic decisions about necessary changes in project design, planning, or implementation. Although evaluations in general (and impact evaluations in particular) produce information periodically rather than continuously, they are nevertheless valuable parts of the project cycle. Even retrospective evaluations are essentially forward looking with regard to the next generation of programming (UNICEF 1991). As illustrated in figure 1.2, evaluation applies the lessons from ongoing or terminated interventions to the planning and design of current and future programs. A well-designed evaluation helps practitioners make the necessary funding cuts to those youth programs that are not achieving their objectives, while sustaining programs that are, or could be, achieving good results. Without data from a good evaluation, the risk of reaching wrong conclusions about whether programs should continue and how resources should be allocated becomes much more significant (World Bank 2009).

“Having good data on why youth dropped out of training enabled us to justify providing stipends. This allowed us to bring the dropout rate from 35 percent to 9 percent.”

— Program Manager,
Caribbean NGO

FIGURE 1.2 The project cycle



Knowledge Generation

The youth livelihood field is characterized by a severe lack of sound evidence.

Even if our institutions do well with regular data collection for monitoring and standard performance assessments (such as by conducting simple before-and-after comparisons or focus groups), we often fail to build generalizable knowledge that would benefit the entire field (Savedoff, Levine, and Birdsall 2006). Acquiring this knowledge typically demands impact studies that use specific methodologies to provide reliable estimates of the success of a specific intervention.

Despite the billions of dollars spent implementing youth livelihood programs, relatively few impact studies exist. For example, in a global review of the evidence of youth employment interventions, Betcherman and colleagues (2007) found only three quality evaluations of youth entrepreneurship programs. Similarly, little is known about other livelihood promotion strategies, such as second chance education, public works programming, or financial education and services for young people (see box 1.1). Even though there have been increasing efforts to build sound evidence in recent years, much more knowledge is needed.

BOX 1.1 Existing evidence on youth employment

Betcherman et al. (2007) conducted a global review of youth employment interventions and found that “only one in ten programs have evaluations which measure both net impact and cost.” The types of interventions with the most severe knowledge gaps were found to be subsidized employment schemes, youth entrepreneurship, employment services, and regulatory reforms. On a regional level, evidence was particularly scarce in Asia, the Middle East and North Africa, and sub-Saharan Africa. An updated database of youth employment interventions and evaluations is available on the Youth Employment Inventory Web site (<http://www.youth-employment-inventory.org>).

Card, Kluve, and Weber (2009) conducted a meta-analysis of active labor market programs in the OECD. Comparing program types, subsidized public sector employment programs were found to have the least favorable impact estimates. Job search assistance programs had relatively favorable short-run impacts, whereas classroom and on-the-job training programs tended to show better outcomes in the medium-run than in the short run. The authors found that programs for youths in the OECD were less likely to yield positive impacts than untargeted programs.

Ibarrarán and Rosas Shady (2009) summarize the findings from rigorous evaluations of job-training programs in Latin America. In contrast to the evidence for developed countries, the results suggest positive effects on employment and the quality of jobs for the trainees, especially among women and the younger participants. The review acknowledges that there is still a major knowledge gap on long-term impacts of such interventions in Latin America.

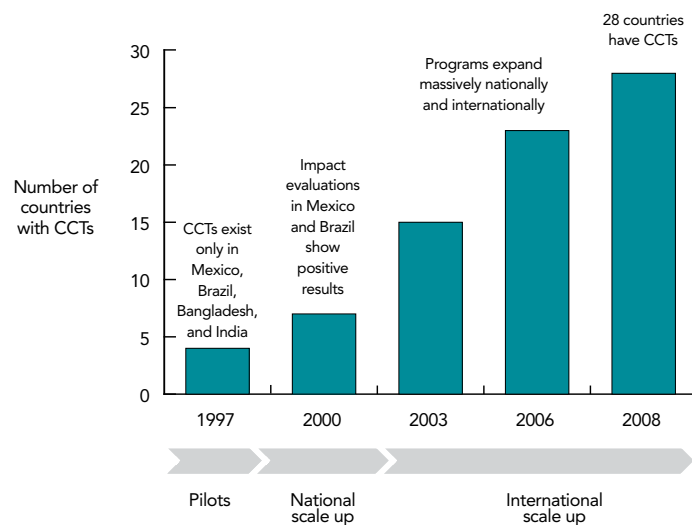
Cunningham, Sanchez-Puerta, and Wuermlí (2010) summarize the state of evidence on active labor market programs for youth, classified by the constraint they are trying to address. Calling for rigorous learning and evaluation across all types of programs, knowledge gaps were found to be particularly severe for interventions such as second chance education, behavioral skills training, entrepreneurial training, public works and public service programs, technology-based job search assistance, skills certification, and microfinance, among others.

Note: See the [resources](#) section at the end of this guide for a list of completed and on-going impact evaluations in the youth livelihood field.

The dearth of rigorous studies—despite huge demand—severely limits large-scale investments in the sector. The lack of evidence is a constraint to winning public support for youth livelihood interventions. Government officials typically want impact and cost-benefit estimates before investing in large programs. As a result of the lack of such evidence in the youth livelihood field, it is often difficult to make a convincing case in comparison to other interventions, such as infrastructure development, where much more evidence is available. Improving the evidence base would therefore also facilitate scale up and replication.

This potential has become obvious in other policy areas. The growing evidence on conditional cash transfer (CCT) programs, for instance, has enabled the international community to promote large-scale interventions in this area across the globe. Mexico and Brazil were two of only four countries worldwide with CCTs in 1997, but the evidence from their impact evaluations has resulted in a massive expansion of the model to twenty-eight countries in 2008 (see figure 1.3).

FIGURE 1.3 From evidence to policy: Conditional cash transfer programs



Sources: Fiszbein and Schady (2009); Rawlings and Rubio (2005).

The publicly funded job-training initiatives in Latin America, the Jóvenes programs, combined technical training with soft skills training, internships, and other services. Impact evaluations demonstrated measurable effects, such as an increase in employment rates and wages that reached more than 10 percent for younger and female cohorts in some countries. As a result, the Jóvenes programs were quickly replicated across the continent. Since the newer programs were also evaluated, the case illustrated the diversity of impacts that could occur in different countries and settings, highlighting that a critical mass of evaluations is always needed to be able to generalize results. Although the Jóvenes experience is still an exception in the youth livelihood sector with respect to the systematic evaluation of impacts, it gives a flavor for the possibilities for expanding the field if only we could distill better knowledge and evidence from the hundreds of interventions we are implementing every year.

Systematically building evidence about what works in strengthening young people's economic opportunities would make it possible to improve the effectiveness and efficiency of our work by bringing vital knowledge into the service of practitioners and policymakers and ultimately strengthen the entire field.

Accountability

In addition to enhancing internal and sector-wide learning, evaluation strengthens the legitimacy of our operations. Funding agencies and society are increasingly demanding accountability from development programs, and evaluations—impact evaluations in particular—can provide the needed evidence on whether a particular program achieved its desired results.

Our resources should not be taken for granted. In almost all instances, our projects are financed with public or private funds, such as official development assistance

(essentially taxpayer money) or private donations. In both cases, someone entrusted us, directly or indirectly, to use this money in the best possible way to help young people achieve a better life. The fact that we are entrusted to develop and implement a youth livelihood project means that this money is not going to be used to build rural roads, enhance an HIV/AIDS prevention project, or buy school materials. Given the scarcity of resources, it seems only natural to use evaluations to provide an honest account of our work: how the money has been used, the activities that were financed, and the results we have achieved.

We have a responsibility to ensure the best possible use of funds. Development interventions are inherently complex, and it would be illusive to expect a 100 percent success rate. On the other hand, a project does not automatically increase people's well-being simply because it is well intended. In order to make sure that a specific program is doing more good than harm and that the benefits of the investment exceed the benefits under alternative uses of the resources, practitioners should always make it a priority to carefully assess the effectiveness of that intervention ([Jones et al. 2009](#)).

Credibility and Sustainability

Evaluations help increase the legitimacy of the project and the reputation of the implementing organization. This argument is not often mentioned in the literature, but in practice it may be among the most compelling reasons to conduct an impact evaluation, as it directly benefits the program and implementing organization.

Impact evaluations can enhance the credibility and reputation of our organization. Because quality evaluations are rare, they receive special attention. As a result, the simple fact that an organization or project agrees to carry out an impact evaluation already indicates good standards in programming. If the evaluation shows good results, then the payoff for the organization and program can be immense. Imagine that among the hundreds of players in the field, *you* are the one who is able to demonstrate that your method is working, that your program is successfully providing young men and women with income opportunities clearly superior to those that would have been available to them had they not participated in your program. The difference is that now you are not only able to *claim* that your intervention is effective, you are able to *prove* it. This makes a big difference in the eyes of donors and policymakers, who, prior to the evaluation, were unable to differentiate the impacts of your intervention from the alleged impacts of numerous other programs.

The ability to stand out can provide a series of benefits for both the project and the organization. Positive evaluation results can be used in advocacy and fundraising efforts to obtain greater support from donors, governments, and the general public. With greater public and political support, our project and organization can quickly become a reference in the field. This, in turn, often leads to an increase in the demand for services, and we may be expected to expand our services nationally and across borders.

Take, for example, the case of Colectivo Integral de Desarrollo in Peru. In 2003, the organization was among the very first to provide rigorous evidence that their model to promote young low-income entrepreneurs was increasing business size, improving business survival, and boosting incomes. As a result of that evidence, they received multiple awards and had no more difficulties securing funding for their programming. In fact, Colectivo is now supported by a grant program of the Inter-American Development Bank and is expanding its model to Central America and the Caribbean (see box 1.2).

.....
International donors are increasingly looking at rigorous impact evaluations to measure the success of the programs they fund. In 2011, both the UK's Department for International Development (DFID) and the U.S. Agency for International Development (USAID) have strengthened their focus on results. See:

DFID's Business Plan 2011–2015, Section 2 "Make British aid more effective by improving transparency and value for money"
<http://www.dfid.gov.uk/Documents/DFID-business-plan.pdf>

USAID's Evaluation Policy 2011
<http://www.usaid.gov/evaluation/USAIDEvaluationPolicy.pdf>

“Because we could prove how youth employment improved, the government invited us to co-design an employability program under the new president.”

— Program Director,
Chilean NGO

.....
A program in the Middle East provided an innovative approach to training young women for jobs as executive assistants to women entrepreneurs. The program leveraged substantial connections to the business sector and in particular to women entrepreneurs, who were interested in supporting and empowering disadvantaged young women. The initial evaluation, however, found fairly high levels of dropouts for the young women trainees once they had been placed in jobs. Further investigation as to the cause of these dropouts found that women entrepreneurs had very high expectations of these young women, but offered insufficient mentorship to them to support their success in demanding work environments. Through open dialogue with the local implementing partner, the program reframed life skills modules to better prepare youth for the demands of these jobs, and also re-oriented business owners to ensure they were providing sufficient mentorship for new employees. The implementing agency was able to achieve better outcomes as a result of these mid-course corrections in the program strategy, shared these lessons with other donors, and in turn, secured additional funding to expand the program.

BOX 1.2 Benefits of conducting an impact evaluation

Response by the president of the Peruvian NGO Colectivo Integral de Desarrollo to the question “How do you think your organization has benefited from conducting an impact evaluation?”

“It improved the quality of our intervention.”

“It improved the program’s credibility.”

“It improved the value of our brand in the eyes of donors.”

“It increased demand for our services.”

“We earned national and international recognition.”

“We are a model institution for the replication in other contexts and countries.”

Source: Dino Linares, Colectivo Integral de Desarrollo president, personal communication (January 28, 2011).

Negative Evaluation Results Are Not Necessarily Bad

We may be afraid that negative evaluation results will lead to funding cuts from our donors. Yet, evaluations that fail to confirm positive results of an intervention can be put to good use.

Negative results are unavoidable in innovative programming. Innovation and creativity are crucial to helping young people master their transition to work. Such innovation will by definition involve failures. As in any other field such as medicine, chemistry, or physics, building successful products and services requires testing, prototyping, refining, and adapting to local circumstances. Failures are a necessary step toward state-of-the-art programming.

Negative results can help improve operations. If, early on, we are able to understand the problems that may reduce the effectiveness of our intervention, then we are in good shape to build successful projects in the long run. Bad news from negative evaluation results points us toward ways of improving our programming.

Addressing negative results proactively fosters credibility. No donor or policymaker expects or believes that every project will be a great success. Disseminating findings, whether favorable or not, signals our ability to be self-critical and our commitment to continuous learning and evidence-based programming. Granted, the pressure to show results and to justify budgets can create strong incentives to report positive findings above the negative ones. But in the long term, an honest discussion of what worked and what did not is likely to yield the biggest payoff.

Key Points

1. Evaluations are first and foremost about learning for the benefit of our own project and organization. Evaluations allow us to show the true value of our work and inform the design and planning of other interventions.
2. Evaluations create a much-needed evidence base for the youth livelihood field. More and better knowledge about what works and what doesn’t will help practitioners design successful interventions and convince policymakers to provide public support.

3. Evaluations provide legitimacy by holding ourselves accountable to donors and the public. Evaluations ensure the good use of taxpayer money and donations.
4. Evaluations enhance our credibility and reputation. In a sector in which robust evidence is scarce, conducting evaluations can have significant payoffs in terms of boosting demand for our services, strengthening our organization's brand, and ultimately securing sustainable financial support.
5. Evaluations don't have to show good results to be useful. On the contrary, failures foster learning. Proactively addressing and disseminating negative evaluation results will likely enhance our credibility and reputation.

NUSAF Case Study: Why Evaluate?

The NUSAF impact evaluation was initiated by the Government of Uganda with support from the World Bank. The primary reason for the impact evaluation was to improve program management. Seeking to estimate the impact of the Youth Opportunities Program on the livelihoods and wellbeing of youth in Northern Uganda, the evaluation was intended to inform future rounds of programming and potential scale up.

The impact evaluation was also intended to fill an important gap in understanding the effectiveness of employment and entrepreneurial skills training programs, particularly in the African context. By providing grants to obtain skills training and start-up capital for establishing productive enterprises, the Youth Opportunities Program is a hybrid of two of the most common types of employment programs. Since little is known about the effectiveness of such an approach, the evaluation would generate knowledge that could inform the entire youth livelihoods field.

Source: Blattman, Fiala, and Martinez (2011).

The Power of Measuring Results

- If you do not measure results, you cannot tell success from failure.
- If you cannot see success, you cannot learn from it.
- If you cannot see failure, you cannot correct it.
- If you can demonstrate results, you can win public support and funding.

Source: Adapted from Osborn and Gaebler (1992).

Key Reading

Savedoff, W., Levine, R., and Birdsall, N. 2006. *When Will We Ever Learn? Improving Lives through Impact Evaluation*. Washington, DC: Center for Global Development.
<http://www.cgdev.org/content/publications/detail/7973>.

Notes