IE4D

IMPACT EVALUATION

FOR DEVELOPMENT

Principles for Action

WHO WE ARE

THE IE4D GROUP

The authors of this paper come from a variety of perspectives. As scholars, practitioners, and commissioners of evaluation in development, research and philanthropy, our thematic interests, disciplines, geographic locale, and experiences may differ but we share a fundamental belief that evaluative knowledge has the potential to contribute to positive social change.

We know that the full potential of evaluation is not always (or even often) realized in international development and philanthropy. There are many reasons for this – some to do with a lack of capacity, some methodological, some due to power imbalances, and some the result of prevailing incentive structures. Evaluation, like development, needs to be an open and dynamic enterprise. Some of the current trends in evaluation, especially in impact evaluation in international development, limit unnecessarily the range of approaches to assessing the impact of development initiatives.

We believe that impact evaluation needs to draw from a diverse range of approaches if it is to be useful in a wide range of development contexts, rigorous, feasible, credible, and ethical.

Developed with support from the Rockefeller Foundation this article is a contribution to ongoing global and regional discussions about ways of realizing the potential of impact evaluation to improve development and strengthening our commitment to work towards it.

Patricia Rogers is Professor of Public Sector Evaluation at the Royal Melbourne Institute of Technology, Australia. Her work focuses on credible and useful evaluation methods, approaches and systems for complicated and complex programs and policies.

Sanjeev Khagram is a professor of public affairs and international studies at the University of Washington as well as the Lead Steward of Innovations for Scaling Impact (iScale).

David Bonbright is founder and Chief Executive of Keystone (U.K., U.S. and South Africa), which helps organizations develop new ways of planning, measuring and reporting social change. He has also worked for the Aga Khan Foundation, Ford Foundation and Ashoka.

Sarah Earl is Senior Program Specialist in the Evaluation Unit at the International Development Research Centre (Canada). Her interest is ensuring that evaluation and research realize their full potential to contribute to positive social change.

Fred Carden is Director of Evaluation at the International Development Research Centre (Canada). His particular expertise is in the development and adaptation of evaluation methodology for the evaluation of development research.
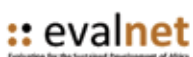
Zenda Ofir is an international evaluation specialist, past President of the African Evaluation Association (AfrEA), former board member of the American Evaluation Association and the NONIE Steering Committee, and evaluation advisor to a variety of international organizations.

Nancy MacPherson is the Managing Director for Evaluation at the Rockefeller Foundation based in New York. The Foundation's Evaluation Office aims to strengthen evaluative practice in philanthropy and development by supporting rigorous, innovative and context appropriate approaches to evaluation and learning.

PRINCIPLES OF IE₄D

I mpact evaluation can make a difference to development. Accurate, feasible and useful evaluation that is aligned with the type of development initiative it is meant to assess can help make development better. It can guide improvements in policies, programs, and practices, identify and explain successes (so they can be emulated) and failures (so they can be avoided), shape investment decisions, and encourage funders and partners to maintain their support. There is, however, another side to impact evaluation. Low quality or misaligned impact evaluation can waste scarce resources and reinforce inequalities by supporting ineffective or inappropriate policies or practices or by failing to demonstrate the benefits of effective programs.

The promise of well-defined and properly implemented evaluation to improve development, as well as the threat of poorly defined and improperly implemented evaluation to derail it, have brought us together. In this paper we set out an agenda for action to ensure impact evaluation fulfills its promise. We argue for a decisive move away from impact evaluation *of* development to impact evaluation *for* development – impact evaluation that not only assesses development but consciously and demonstrably contributes to improving development. This paper sets out a seven-point agenda of rethinking, reshaping and reforming impact evaluation for improving development. We propose actions that can be taken by all those involved in impact evaluation of development - practitioners, thought leaders, agenda-setters, philanthropists, and managers and commissioners of evaluation.

We propose that impact evaluation should:

1 **Contribute to sustainable improvements in development.** Impact evaluation must ultimately be judged by its contribution to actually improving development. This means that technical merit is necessary but not sufficient. Impact evaluation must be judged in terms of its timeliness, relevance, and usefulness to the different actors involved in making development work. Development should also be evaluated on whether it is likely to build the capacities of local people, institutions and systems over the longer term.

2 **Suit the nature of development.** Impact evaluation methods and processes should be appropriate to the context of development and aligned with the type of development initiative under consideration. Impact evaluation should not start with a decision about a particular approach, but remain open to an appropriate mix of methods to suit the nature and context of the specific development initiative being evaluated.

3 **Draw on the full range of methods and designs** for systematic and rigorous empirical investigation, including ethnographic, case study, statistical and experimental/quasi-experimental approaches and methods, as well as those from such fields as complexity science, participatory research, and action learning. No one approach is inherently more rigorous than another. Rigor depends upon both *appropriate* method choice – selecting methods based on evaluative purposes and contexts – and successful implementation that meets accepted standards of quality.

4 **Produce a comprehensive analysis of impacts and outcomes** including intended, unintended, positive and negative effects, as well as the distribution of results, costs and benefits.

5 **Explain how and why impacts occur.** A good understanding of how results have been achieved (what works, under what conditions, for whom, how, why and for how long) is essential for learning from successes and failures. How local actions are affected by national and international systems, strategies and policies and vice versa should also be assessed. Rigorous examination of causal mechanisms and contextual factors is required when transferring initiatives to new settings and wider application.

6 **Be an integral part of robust systems of monitoring, assessment and learning.** Isolated, discrete impact evaluations are of less value and use than impact evaluations that are integrated into robust systems of monitoring, evaluation and learning. The latter can create synergies between adaptation during a project or program and improvements made after periodic evaluation of results. Impact evaluation is one element of the process of understanding what works, under what conditions, for whom, how and why and for how long. Integrating impact evaluation within a sound monitoring and learning system strengthens its contribution to improving development.

7 **Involve fundamental rethinking, reshaping and reform of existing practice.** Immediate steps can be taken to practice appropriate evaluation to improve development. The distribution of power, the pressure for tangible immediate results and impact, and prevailing incentives in development often push agencies, foundations and practitioners toward inappropriate approaches to evaluation. A deliberate reform effort is essential to ensure that impact evaluation contributes to improved development.

RETHINKING IMPACT EVALUATION

1. MAKE A DIFFERENCE

The value of impact evaluation must ultimately be judged by its usefulness in helping to improve development. In order to do this, impact evaluation needs to respond to the challenges produced by the changing nature of development.

Development clearly involves more than donors and national governments. Yet, evaluation practice often tends to consider them the primary audience of evaluations in general and impact evaluation specifically. Citizens, civil society, philanthropic organizations, the private sector, regional and global organizations, partnerships and networks all contribute to development. In particular, there is increasing recognition of the role of community members as primary agents of development, not as passive recipients of aid. Impact evaluation therefore needs to find practical, yet meaningful, ways to engage the range of actors in the process and address their specific information needs.

Impact evaluation should increase and improve accountability not only to funders and decision makers, but also to the primary constituents and other key stakeholders. Mutual accountability provides deeper legitimacy and improved effectiveness. Narrow legitimacy is achieved through accountability towards funders and other partners. A more profound legitimacy is achieved when those who are meant to benefit from development are able to hold to account those who promote development.

Too many development efforts fail because of an insufficient focus on creating conditions for long-term success. Development initiatives are not only about producing results in the short-term, but also about developing capacities in individuals, institutions and systems for ongoing results, thus enhancing the chance that development will be sustained. Impact evaluation needs to both inform short-term decisions about policies and programs and help to build capacity for improved planning, implementation and learning to produce long-term results. Rethinking the role of impact

evaluation includes considering how it can support or undermine self-determination, constructive negotiation, and improved results in the long term.

2. WAG THE TAIL

Many of the methods and approaches used today in impact evaluation have been developed for discrete, standardized interventions, like drugs or seed varieties. These are development initiatives where the intended results, and how to measure them, are clear and agreed, and controlled comparisons of results with and without drugs or seeds are both possible and meaningful. For development interventions like those above, these methods can work well when they have adequate resources and are well implemented.

Starting from current practice, two principles about impact evaluation rise above others: it must contribute to improved development, and it must suit the nature of development.

But many development initiatives are not like this. They are broadly defined approaches or strategies, often implemented in different ways in different situations, and highly influenced by a wide range of contextual factors. Simple with/without comparisons are often not the most meaningful way of assessing impacts because the initiatives (and their underlying change logics) have complicated aspects (multiple components that vary in different settings) or complex aspects (dynamic and emergent) or both.

Development has changed significantly in the 21st century due, in part, to processes of globalization. In particular, development initiatives are more often complicated and complex in the following ways:

- Diverse initiatives are being implemented more often than discrete, standardized projects.

- Multiple development goals associated with a wide range of issues are being promoted simultaneously.

- Development effectiveness involves a wider range of dimensions.

- A broader range of legitimate stakeholders (beyond aid agencies and national governments) are often actively engaged, and diverse sources and types of capital are used to support development initiatives. Primary constituents are often recognized as central agents of development.

- Contextual factors including social, cultural, economic, geopolitical factors are understood as critical conditions that shape the nature of development.

- This broader array of initiatives often involves many interacting influences and non-linear, recursive causality.

- Effective development often requires experimentation, repeated practice, learning and adaptation during a project or program.

The increasingly prevalent emphasis on using evaluative evidence to inform development practice and make funding decisions has particular implications for the practice of impact evaluation. If the only evidence that is seen as credible is evidence from discrete, standardized interventions, there is a risk that multi-faceted or systems-based initiatives will not be supported – even if they are effective. This will have dire consequences for development.

So what are we to do? Avoiding these realities is not an option. Failing to undertake rigorous impact evaluation of initiatives with complicated and complex aspects, and relying on their face validity, or good intentions, or selective descriptions of direct beneficiaries, is not sufficient. But neither is it appropriate to mold these interventions into discrete, standardized programs to make them easier to evaluate – that would be like the tail wagging the dog. Impact evaluation for development has to be able to evaluate the impact of complicated and complex initiatives in a rigorous manner, and provide useful advice to support translating those findings to other situations. And it can do so, by following the principles for reshaping and reforming impact evaluation below.

3. BE MORE SCIENTIFIC

We applaud the increased emphasis on more scientific impact evaluation. However, some efforts to improve the technical merit of impact evaluation have ended up with an unnecessarily narrowly defined set of methods and inappropriate hierarchy of which methods are best, driven by a conceptualization of science that is inaccurate and outdated.

Scientific impact evaluation draws on insights and techniques from natural, social, systems and complexity sciences. It emphasizes multiple kinds of valid comparisons, including the use of counterfactuals when appropriate, triangulation across methods and types of quantitative and qualitative data to improve measures and analysis. Sophisticated diagnostic tools for appropriate (and often mixed) methods selection and measurement are increasingly available and should be utilized. Appropriate norms, guidelines and criteria should also be used to provide assurance that any method or set of methods meet technical and ethical standards.

Scientific rigor should be judged in terms of appropriate data, method selection and implementation. This kind of rigor identifies and addresses threats to internal and external validity that are present in any measure, method or research design. Validity also has an inherently cultural dimension – and all data, methods and analysis must address the challenge and opportunity of multiple cultural understandings of reality.

4. AVOID A NARROW FOCUS

Development has many dimensions and many development initiatives in the 21st century have multiple goals. While it is not possible for every impact evaluation to cover all aspects of a development initiative, impact evaluation should produce an adequately comprehensive assessment of results.

To enact these cardinal principles, the practice of impact evaluation should be more scientific, avoid a narrow focus, recognize differences and close the loop.

A range of designs can be used to discover and describe different impacts. It is possible for experimental and quasi-experimental methods of impact evaluation to pay attention to differing impacts if relevant variables can be identified and measured in advance. However, it is frequently difficult, if not impossible, to anticipate (and account for) many relevant factors. Impact evaluation approaches that draw on a wider range of credible evidence, especially multiple kinds of comparison and triangulation, make it possible to include variables that may not have been identified or anticipated

There is an ethical imperative to identify and analyze unintended consequences, especially negative ones. Impact evaluation should deliver a balanced assessment that includes intended and unintended, positive and negative impacts. It should assess the distribution of benefits and costs of initiatives, address effectiveness, legitimacy, efficiency and sustainability.

5. EXPLAIN HOW AND WHY AND RECOGNIZE DIFFERENCES

Improvement in impact evaluation means building knowledge about what works, under what conditions, for whom, how and why. It means assembling knowledge about when and in what ways initiatives should be transferred to other settings and how to generate and use evidence to inform policy and practice. When impact evaluation does not recognize local knowledge, cultures or situational differences in its results, then it leads to 'one-size-fits-all' policies to the detriment of improving development.

Data about the average intended effects of a project are critical but often not sufficient to inform good policy or practice. Information is needed about the results for different groups (particularly for the most disadvantaged) to inform judgments about the value, transferability and scalability of a development initiative. Information is also needed about the contexts in which the project is successful to inform conclusions about whether the findings can be generalized. Few development initiatives are expected to work the same way in all cultures and settings at all times. The task is to help understand

how development may be affected by local and global contexts. Understanding how and why impacts differ is at the heart of good impact evaluation.

For some initiatives it will be possible to identify in advance how planned activities and outputs could produce clearly defined results. For many development initiatives, however, no single actor, factor or causal mechanism working independently will be sufficient to produce results. Rather, impact could be achieved through the combined efforts of multiple actors, with overlapping and interacting causal mechanisms, acting in particular contexts. Solutions to development problems will involve changing beliefs, attitudes, relationships, capabilities, conditions and behaviors and often need to be worked out over time with constituents as well as with partners. For these types of initiatives, an iterative and adaptive approach is essential and impact evaluation will need to document and support this.

6. CLOSE THE LOOP

Impact evaluation is more effective when it is part of a robust learning and improvement system that integrates cycles of planning, implementation, monitoring, assessing and reporting. Such learning and improvement-oriented systems translate credible evidence of an effective development process or initiative into actionable guidance, especially for policy makers and citizens. They do so in ways that are appropriate for effective and immediate utilization. Embedded impact evaluation creates synergies between immediate improvements and longer-term assessment of results.

The effectiveness of development initiatives will improve when those who fund and manage them are held accountable by those most affected - positively and negatively. This implies that public reporting of impacts should reflect the ways in which a development initiative responds to and balances the needs and perspectives of its various constituents. Impact evaluation design needs to explicitly consider: (a) how the initiative translates its understanding of what is happening into action; (b) how it adapts and improves; and (c) how different constituents of an initiative understand and act on evidence about the results reported.

Impact evaluation is only one type of evaluation that can contribute to improved development, and resources for it should not detract from other types of evaluation.

7. CHANGE THE SYSTEM

While we recognize the importance of power and politics in improving impact evaluation for effective development, we also argue that everyone involved in evaluation of development initiatives can take action and contribute to positive change. We have identified below an initial list of steps to reform the system.

What should evaluation practitioners do?

Evaluation practitioners are critical actors in this system. They are the implementers of impact evaluation policies and approaches on the ground. Practitioners, particularly those who work in their home countries, have power, knowledge and access to the projects, programs and policies where development action is implemented. While they do not normally have a mandate to make institutional change, independently there are many steps evaluators can take to move us all towards a vibrant field of IE4D.

Challenge yourself; challenge your clients.

- Be a model of good practice; maintain high quality using appropriate evaluation standards; ensure that you follow impact evaluation practices across disciplinary boundaries.

- Recognize that your clients may not be clear about what they need, what degree of certainty evaluation can provide and what can realistically be achieved within their time and resource constraints. Help them to identify their needs and make decisions that will make impact evaluation useful.

- Promote impact evaluation from a systems perspective. Understand your assignment in the wider development context not solely in the narrow parameters of the existing initiative.

Promote appropriately rigorous methods.

- Be a learner as well as an innovator: do not become complacent in your practice. Be aware of changes in development practice, and the broader external environment. Follow and

participate in debates, seek out and contribute to new tools and methods and contribute to standards for good impact evaluation practices.

- Recognize that your values and practices are shaped by your history, as are those of others. Doing impact evaluations in other cultural contexts is therefore extremely sensitive and requires active dialogue with, and participation of, key stakeholders.

Build vibrant evaluation networks.

- Seek out and work with others through associations and communities of practice. These provide not only professional support but are also key mechanisms for reforming the policies and agencies that shape impact evaluation practice. Do not settle for passive engagement in your field and passive acceptance of popular trends in practice without carefully thinking through the implications.

Fundamental reform requires each of us to ask, "How can we do better?" and "What do we need to do differently?" to implement these changes.

Contribute to improved reporting and knowledge sharing.

- Tailor reporting and communication of impact evaluation results, approaches, methods, and lessons to reach a broader range of audiences across disciplines, practice, policy and research.

What should those who commission and manage evaluation do?

Those who commission and manage impact evaluations are in an extraordinary position of privilege. Whether they recognize it or not, they have the power and authority to change the way impact evaluation for development is conceived, commissioned, managed, reported and disseminated. If they fail to recognize, or act on, the prevailing inequities and biases of much current development evaluation practice they perpetuate inappropriate practices. Those who commission and manage evaluations must be strong, independent thinkers, well grounded in development and evaluation theory and practice. They should be

transparent about their values and practices and brave enough to invite constructive criticism and to listen to evaluators who will 'speak truth to power'. Evaluation commissioners and managers must be committed to learning from success, but must also be prepared to face the failures of development and shatter ill-founded dogmas, question self-serving assumptions and challenge complacency among stakeholders. While this may be a tall order in many political environments and organizations where 'safe spaces' for evaluation and learning are limited, there are practical steps that all commissioners and managers can take to create a supportive environment and a set of incentives to work towards IE4D:

Be clear how impact evaluation is intended to contribute to improving development

- Conceptualize and frame evaluation around clearly defined theories and assumptions of what contributes to effective development and improvements in the lives of people and the systems upon which they depend. (Use explicit definitions of development and hypotheses about the role that rights, choices, freedom, economic growth, sustainability, accountability, etc. play in development processes).

- Focus on outcomes and impacts - not only outputs - to the extent feasible and appropriate.

Reflect standards and guidelines

- Set, promote, use and support the improvement of standards and guidelines for development evaluation that reflect international good practice and regionally adapted standards (such as the African Evaluation Guidelines).

Extend Boundaries

- Commission evaluations that push the boundaries of existing approaches to impact evaluation and that evaluate beyond the interests of those in charge of specific development initiatives.

Broaden the focus

- Join forces with other organizations and agencies to broaden the focus of what is evaluated in order to capture what matters in development.
- Evaluate the the drivers and root causes of unsustainable development and the effects of developed country policies on developing countries in areas such as foreign investment, trade, globalization, migration, pollution, and intellectual property. Include evaluating capacity development at individual, institutional and systems levels as a critical step towards achieving development outcomes and impact.

Embed evaluation within the management and leadership of organizations.

- Make sure evaluations consider the way an organization or initiative plans, monitors, assesses, and reports back to its constituents and the wider public.

Untie funds for impact evaluation for development

- Manage the procurement process to prioritize the use and development of high quality regional, national and local evaluation skills.
- Set up pooled national funds at arm's length from power-holders to enable less empowered and marginalized groups and organizations to choose their own evaluators to assess development initiatives.

Invest in evaluation infrastructure

- Invest in capacity and field building in impact evaluation at multiple levels – developing individual skills, organizational infrastructure, networks of practitioners and professional evaluation associations.

Improve reporting

- Commission and support improved reporting and communication of evaluation results, methods, and lessons, in order to reach a broader range of audiences across disciplines, practice, policy, and research. This could include authoritative, user-friendly briefs on progressive approaches to, and examples of, impact evaluation to extend and balance the existing body of information.

What should thought leaders, agenda setters, educators, and mentors do?

Those who create impact evaluation knowledge and teach, mentor or influence those who commission and practice impact evaluation can also contribute to positive change. As the field of evaluation - including impact evaluation - becomes more reflective, it calls for champions who can communicate its value, who can be influencers among decision-makers and who can motivate and engage the broader public. They need to translate theory into practice and improve theory based on good practice, as well as disseminate new ideas, develop new concepts and challenge conventional wisdom.

As development is increasingly recognized as trans-disciplinary, so too should training and capacity building in impact evaluation bring together the fields of social, natural and emerging sciences, management and cultural disciplines to provide much needed insights into complex questions of behavioral, institutional and structural change. Thought leaders, agenda setters, trainers and mentors can support IE4D through the following strategies:

Broaden the collaboration

- Provide 'safe spaces' for evaluative learning, analysis and reflection within academic and research environments, and in civil society and government.
- Bring together citizens, civil society, business and public sector to examine new ways to collaborate on impact evaluation.

Plan for different types of accountability and learning

- Incorporate accountability frameworks, transparency requirements, inclusive methodologies, and citizen engagement in new ways of practicing development evaluation.
- Promote improved reporting and knowledge sharing of evaluation results, methods, tools, advice, and examples to reach a broader range of audiences across disciplines, practice, policy, and research.
- Ensure training and professional development incorporates multiple approaches and methods in impact evaluation, drawing from different disciplines, perspectives and worldviews.

FINAL THOUGHTS

We are privileged to work in an expanding field where our findings and processes can change lives for the better. Evaluation approaches are currently in the spotlight among politicians, policy makers and philanthropists who seek to maximize the impact of their resources. It is therefore timely to work towards the fundamental changes both necessary and possible to make impact evaluation more relevant, credible and useful for development.

We extend an invitation to those who want to make a difference, to those who want to bring about change. There is much to do – let's get to work.