

Learning how to learn: eight lessons for impact evaluations that make a difference

By Ben Ramalingam

This Background Note outlines key lessons on impact evaluations, utilisation-focused evaluations and evidence-based policy. Drawing on recent and ongoing work by the Overseas Development Institute (ODI), the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP), the International Initiative for Impact Evaluation (3ie) and others, it is aimed at researchers, policy-makers and practitioners working on development and humanitarian issues.

Introduction

The international aid system has faced the challenge of assessing its development and humanitarian impact in a systematic and credible fashion since at least the 1970s. Prior to this, according to one leading authority, there was ‘little interest in the impact and effectiveness of aid: the (assumed) need for aid was seen as a sufficient basis for providing it’ (Riddell, 2009).

Things changed, and rightly so. Efforts to understand, define and measure impact led to a wealth of books, studies, papers and research programmes, and a number of dedicated networks and organisations. Despite all this activity, there was a distinct lack of progress in understanding the impacts of development and humanitarian aid, whether at project, programme or country-wide level. As Roger Riddell noted in one of the most comprehensive reviews to date, focusing on whether aid has worked or failed: ‘there is no authoritative, overarching evidence to clinch the argument’ (Riddell, 2007). Attempts usually started with high ambitions but were all too often scaled back, narrowed in scope and made ‘more realistic’ in the face of the complex realities of development efforts. As a result, the arguments for and against the effectiveness of aid policies and practices remained patchy, partial and inconclusive.

Frustration at this lack of progress was articulated in 2006 with publication of the Center for Global Development’s Evaluation Gap Working Group paper, *When Will We Ever Learn?* (CGD, 2006). This influential study presented the lack of rigorous impact evaluations as the missing piece in learning about social development efforts. It advocated a renewed approach to aid evaluation that would bring greater accuracy and credibility to assessments of impact and, by extension, to development policy and practice. It recommended the creation of a new entity to help fill this ‘Evaluation Gap’. And it set out the challenge in clear terms: that 10 years on from the publication of the report, more and better rigorous impact evaluations would need to be in place if we wanted a stronger evidence-base for making decisions.

According to Esther Duflo, one of the leading lights of the new impact movement: ‘Creating a culture in which rigorous randomised evaluations are promoted, encouraged, and financed has the potential to revolutionise social policy during the 21st century, just as randomised trials revolutionised medicine during the 20th’ (cited in *The Lancet*, 2004).

While the overarching goals of improving development policy and practice are widely shared, the randomised approach has been the subject of much heated and protracted debate, some of it reviewed below. However, even the staunchest advocates of such ‘rigorous’ approaches are starting to acknowledge that changing policy is not ‘automatic or easy’ (Dhaliwal, 2011).

A fundamental shift is needed from the ‘build it and they will come’ approach that has characterised much of the recent impact evaluation debate and practice. As the Impact Evaluation for Development (IE4D) Group stated in their January 2011 Principles for Action: ‘it is ... timely to work towards the fundamental changes both necessary and possible to make impact evaluation more relevant, credible and useful for development’ (IE4D, 2011).

The position taken in this Background Note is that methodological pluralism is the key to effective impact evaluation in development. However, the emphasis

here is not methods per se. Instead, the focus is on the range of factors and issues that need to be considered for impact evaluations to be used in policy and practice – regardless of the method employed.

This Note synthesises research by ODI, ALNAP, 3ie and others to outline eight key lessons for consideration by all of those with an interest in impact evaluation and aid effectiveness.

The real evaluation gap?

We are now five years into the process outlined in *When Will We Ever Learn?* One of the most important outcomes of the study, the International Initiative on Impact Evaluation (3ie) estimates that since the study some 800 impact evaluations of social interventions have been completed or are in progress in low- and middle-income countries. This is clearly an impressive achievement. But, at the risk of sounding facetious, this is an output-based assessment (with no before-and-after comparison).

More seriously, progress towards greater use of evidence in development and humanitarian aid has been uneven at best. Particular methods for impact evaluations are not, it would appear, a magic fix to bring about improvements in evidence-based policy and practice. All methodologies face two sets of limitations: those regarding the knowledge they generate; and, related to this, limitations around the use and uptake of that knowledge.

First, it has been shown that the knowledge that results from any type of particular impact evaluation methodology is no more rigorous or widely applicable than the results from any other kind of methodology.

Angus Deaton has written convincingly on this in relation to randomised control trials (RCTs). He argues that, in ideal circumstances, randomised evaluations of projects are useful for obtaining a convincing estimate of the average effects of a programme or project. However, ‘the price for this success is a focus that is too narrow to tell us “what works” in development, to design policy, or to advance scientific knowledge about development processes’ (Deaton, 2009). Dani Rodrik takes a similar line, arguing that the typical randomised evaluation is undertaken in ‘a specific locale on a specific group and under specific experimental conditions [and] its generalizability to other settings is never assured’ (Rodrik, 2008).

A study from Public/Private Ventures concurs: ‘While RCTs are a vital mechanism for assessing program impacts, they are a means to a certain kind of knowledge under certain circumstances. They are not an end in themselves. RCTs and other rigorous impact studies can easily be imposed inappropriately (and at significant cost), leading to unfair and unhelpful

generalizations about program effectiveness’ (Public/Private Ventures, 2011). While the focus here has been on RCTs, the general point is that the only gold standard in impact evaluation is methodological appropriateness (Patton, 2008a).

Second, on take-up and influence: filling the perceived evaluation gap has not led to a wealth of new learning and better decisions. This is partly because of the above mentioned limitations of knowledge and the lack of generalisability of RCT findings. But it is also because evidence – regardless of its origins – is not absorbed into policy and practice in linear and straightforward ways (ODI, 1999; and various others). A session organised by the Abdul Latif Jameel Poverty Action Lab (J-PAL) at the March 2011 conference of the Network of Networks on Impact Evaluation (NONIE) notes that, ‘while the use and acceptance of randomized evaluations has increased in recent years, the translation of the results from such studies into policy impact is neither automatic nor easy’. A series of case studies has been commissioned by 3ie, looking at exactly this issue and will be published later in 2011 (Weyrauch and Langou, 2011).

Despite the volume of impact evaluations since 2006, much remains unchanged in the aid sector. Mistakes are still seldom acknowledged and frequently repeated. Policy-based evidence still outweighs, by some way, evidence-based policy. Standard operating practices still fail to give way to new ideas, or do so all too slowly. Research is still frequently sidelined, bad news is still buried, and the lack of results is still not publicised.

In the face of this lack of progress, some might conclude that the missing piece may not, in fact, have been ‘rigorous’ impact evaluation. Simply funding more such impact evaluations in considerable quantities has not been enough to produce better programme or policy decisions. As noted in a previous ODI opinion piece (Jones, 2009) questioning the assumed hierarchy of evidence, the gold standard has not proved to be a silver bullet.

To go one step further, in this area of work, silver bullets may well be red herrings. Issues of the partiality and contextual nature of knowledge and the complexity of uptake are faced by *all* forms of evaluative research, albeit in different ways. These limitations need to be understood, acknowledged and taken into account *for all methodologies* if the impact evaluation agenda is to reach its full potential in the development and humanitarian sectors.

This resonates with over ten years of findings from ODI’s Research and Policy in Development (RAPID) programme, which tells us that evidence – however derived – is a necessary but not sufficient component of improved performance and practice. As 2009 ODI research found, overstating the potential of RCTs – or indeed any specific

methodology – runs the risk of diminishing the whole of the impact agenda (Jones et al., 2009).

Howard White, director of 3ie, has also argued that the number of reports is only the first step. In a piece written in the lead-up to the NONIE conference in March 2011, he highlights key challenges for the impact-evaluation community (Box 1).

The rest of this Background Note focuses on lessons from ODI, ALNAP, 3ie and other research, which address, in broad terms, the first three of these challenges.

Key lessons for impact evaluations that work

Just as there have been repeated calls for aid evaluation to move beyond outputs to measure downstream impacts, it is also important for the same principles to be internalised in thinking about impact evaluations themselves. There are eight specific insights from previous research that are worth considering to maximise the possibilities of take-up. These fall into three broad categories, as follows, and are explored in turn below.

- **Institutional readiness**
 - 1) Understand the key stakeholders
 - 2) Adapt the incentives
 - 3) Invest in capacities and skills
- **Implementation**
 - 4) Define impact in relation to the specific context
 - 5) Develop the right blend of methodologies
 - 6) Involve those who matter in the decisions that matter
- **Communication and engagement**
 - 7) Communicate effectively
 - 8) Be persistent and flexible

Lessons on institutional readiness

Lesson 1: Understand the key stakeholders

An impact evaluation – like any utilised evaluation – must balance the priorities and interests of a range of different stakeholders if it is to contribute to improved effectiveness. A vital first step, therefore, is to determine who needs to know about impact – or the lack of it – and why. Is an impact evaluation for donors, the tax-paying public, the implementing agency, the wider academic and research community, national actors or the people most directly affected? Is the assessment for learning or accountability? As Robert Chambers has put it: ‘the starting point would be to ask about the political economy of the evaluation: who would gain? Who might lose? And how? And, especially, how was it intended and anticipated that the findings would make a difference?’ (Chambers, 2009).

The key issue is whether the questions being posed in the impact evaluation are relevant to these needs. If they are not, then there is a high likelihood the evaluation will not see substantial take-up (Patton, 2008b).

Box 1: Five key challenges facing the impact-evaluation community (White, 2011)

1. Identify and strengthen processes to ensure that evidence is used in policy.
2. Institutionalise impact evaluation.
3. Improve evaluation designs to answer policy-relevant questions.
4. Make progress with small n impact evaluations.
5. Expand knowledge and use of systematic reviews.

Like all evaluations, impact evaluations relate to two distinct institutional priorities: accountability and learning. Yet, there is an inherent tension between these two objectives. For example, the framing of impact evaluation in terms of *accountability for results* can encourage risk aversion among implementing agencies and, therefore, undermine the learning and innovation required for *improving performance* (Stein, 2008; OIOS, 2008).

Choices regarding the purpose and scope of impact evaluations are political and have important implications for the selection of appropriate methodologies, the kinds of knowledge and conclusions generated, and follow-up and use of these. It is crucial therefore, that adequate time is factored in for the meaningful participation of all stakeholders in defining the purpose and scope of impact evaluations (Patton, 2008b; Sandison, 2006; Proudlock et al., 2009).

Lesson 2: Adapt the incentives

Without the right incentives and capacities, evidence – however rigorous or technically proficient – will not be picked up and used (Proudlock et al., 2009). The incentives that create a demand for impact evaluations and their findings, in particular, need to be better understood and strengthened (Jones, 2009). Research published on evaluation use in the humanitarian sector found that the biggest influences on decision-making were donors and senior peers in other organisations (Sandison, 2006). Both sets of actors are crucial for better impact utilisation, and they need to put their money, and their actions, where their mouth is.

First, donors need to ask for different kinds of impact evaluations to justify funding decisions, but they also have to avoid over-extending these asks to the point where they stretch the available evidence. They also have to invest in strengthening impact-assessment capacities and incentives within international agencies and in developing countries. The UK Department for International Development (DFID) is perhaps the exemplary donor in this regard, and others would do well to follow suit.

Second, leaders across the development system also have a crucial role to play in asking for, and demonstrating, evidence-based decision-making. In

particular, there is a need to bring considerations of impact firmly into existing systems for monitoring, learning and evaluation. There are a few examples of leaders who are using the right kinds of rhetoric, but too few examples of behaviours that follow up on this. Rajiv Shah's work at USAID, and his endorsement of the new USAID evaluation strategy, is a prominent example at the time of writing. 3ie has also found that leaders in national governments play a crucial role in ensuring effective and credible evaluations (Briceño and Gaarder, 2009).

Of course, incentives are complex, and there are all kinds of institutional reasons why demand for policy-based evidence might outweigh demand for evidence-based policy. But sustained and coordinated efforts from these two sets of actors – donors and aid leaders – would be a good starting point for a transition from the current dynamics. This would also prevent impact being turned into an 'escape hatch' whereby donors and those at the policy level can shift responsibility for aid failures to implementers.

Lesson 3: Invest in capacities and skills

Just as important as the incentives that stimulate demand for impact evaluations are the capacities and skills that enable them to be carried out effectively. Addressing capacities is vital if impact evaluations are to be institutionalised as a key tool for aid accountability and learning.

This means not just training, but a whole suite of learning approaches: from secondments to research institutes and opportunities to work on impact evaluations within the organisation or elsewhere, to time spent by programme staff in evaluation departments and, equally, time spent by evaluators in the field (Foresti, 2007).

It is all too common to focus capacity-building efforts on developing countries but, as recent research has shown (Hallam, 2011), this is just as much an issue within international agencies. Agencies must make good earlier commitments to becoming learning and knowledge-based organisations, and see the strengthening of the evaluation function as a central part of such efforts.

Lessons on implementation

Lesson 4: Define impact in ways that relate to the specific context

Effective assessment of the impacts of a development or humanitarian intervention requires clarity on the underlying logic or 'theory of change', based on a solid understanding of the needs that particular programme is seeking to address. Such impacts can be positive or negative, primary or secondary, direct or indirect, and intended or unexpected.

There is wide agreement that impact can be defined

in various ways. Perhaps the most important consideration is to clarify what impact means in ways that are specific to particular sectors, interventions and contexts.

Broadly speaking, it is important for impact evaluation to take account of and reflect the simple, complicated and complex aspects of any given intervention. Simple aspects can be tightly specified and standardised; complicated aspects work as part of a causal package; complex aspects are appropriately dynamic and adaptive (Rogers, 2009). Work on bringing the ideas of complex adaptive systems thinking to bear on evaluation practices are especially relevant here (Patton, 2010; Ramalingam et al., 2008).

Such definitions need to enable practical implementation of an effective and timely assessment, and be sensitive to the needs of the stakeholders and their interests.

Lesson 5: Develop the right blend of methodologies

This element of the impact evaluation agenda has proved to be a very active, and even dominant debate, but for many a highly distracting one. Because of the prominence of this issue, debates on aid evaluation have become polarised around questions of methodology. The growth in RCTs is seen by some as the silver bullet to aid effectiveness issues, and by others as anathema to the kind of adaptive learning that is at the heart of successful development.

Research shows that it is vital to determine what methods are appropriate to the users' needs, the given context, and issues of data, baselines, indicators, timing and comparison groups. Three broad categories of methodology can be identified – experimental, inductive and participatory (Hulme, 2000). It is important to note that, when it comes to policy influence, the randomised, the inductive and the participatory are complementary approaches, generating different kinds of knowledge that will be of varying levels of importance and value for the key stakeholders identified above.

As 2009 ODI research found, impact evaluation's potential to shape donor investments and national-level policy decision-making has been constrained, in part, because of insufficient attention to diverse methodological approaches to evaluation. The authors 'suggest strongly that in all sectors there is a strong need for critical reflection on the suitability of methods to development questions and required knowledge' (Jones et al., 2009).

Lesson 6: Involve those who matter in the decisions that matter

The literature on evaluation use (Patton, 2008b) suggests that key stakeholders need to be involved throughout the evaluation process. This should, at

the very least, ensure timely feedback on issues of design and on emerging results. At the other end of the spectrum, stakeholders can be involved in the planning, design and interpretation of results.

The reality is that this places considerable demands on evaluators and the evaluation process, especially in regards to time and resources. As such, stakeholder engagement needs to be managed with care – too much stakeholder involvement could lead to undue influence on the evaluation, and too little could lead to evaluators dominating the process (Patton, 2008b).

The stakeholder group that is most often neglected in evaluations are the recipients and intended beneficiaries of aid. Work at Tufts University on impact evaluations (Catley et al., 2008), supported by the Gates Foundation and working in partnership with five leading NGOs, found that improved participation could help overcome some of the ‘inherent weaknesses’ in conventional impact-assessment approaches. Benefits included a shift away from a process focus to measuring actual impacts, an emphasis on community-based as opposed to externally-derived indicators of impact, and ways to address issues of weak or non-existent baselines. ALNAP case-study research (Proudlock et al., 2009) also found that that the whole process of impact evaluation, and particularly the analysis and interpretation of results, can be greatly improved by the participation of intended beneficiaries, who are after all the primary stakeholders in their own development and the best judges of their own situation.

Lessons on communication and advocacy

Lesson 7: Communicate effectively

Well timed communication efforts focusing on policy ‘windows of opportunity’ are vital for successful uptake of evidence (Young and Mendizabal, 2009). This needs to take into account both the hard side (reports, datasets, etc.) and the soft side (meetings, social networks and social media). The *Joint Evaluation of Emergency Assistance to Rwanda* is acknowledged as one of the most influential evaluations in the aid sector (Buchanan-Smith, 2003) but its influence was not immediate or automatic. The follow-up to the evaluation consisted of a small network with a budget and a part-time secretariat. In the 15 months after the evaluation was published, the members of the follow-up network contributed in different ways to over 70 events and meetings (Borton, 2004).

All too often communication strategies are seen as a one-way process, whereas the reality is that policy and practice uptake is a dynamic process that can take considerable amounts of time, persistence and ongoing dialogue. Even with quality evidence, some practices in the aid system have taken decades to be changed. As well as a supply of credible evidence,

effort needs to be made to understand the demand for evidence. 3ie and its partners have focused on impact evaluation capacity-building for policy-makers as a means and an entry point for political buy-in and to influence policy (3ie, pers. comm.), and have also explored the demand side of the impact evaluation (Weyrauch and Langou, 2011).

Lesson 8: Be persistent and flexible

ODI research (Young and Mendizabal, 2009; RAPID, numerous) has found that the biggest influence on research uptake is political context. No specific methodology can influence this, but finding ways to capture the political imagination and spot windows of opportunity are vital. This requires evaluators and other stakeholders (the evaluation champions) to scan the horizon for new ways to bring the evidence to bear on how things are done. When the evidence flies in the face of accepted wisdom or standard operating procedures, it is especially problematic, and the influence timeline for seeing changes on the ground might be years, if not decades. Short-term approaches to influence ongoing debates and dialogue need to be balanced with medium-term approaches to change policy and practice.

Conclusions

These eight lessons are a synthesis of numerous different studies representing a range of perspectives and approaches. However, one thing that is shared across all of these studies is a sense that there is a need to move beyond the overly technical focus of much of the recent debate on impact evaluations.

In particular, when considering use and influence, the overarching conclusion is that the key factors in impact utilisation are as much to do with human, organisational and political factors. To put it simply: it is becoming clear that no single methodology – experimental, inductive or participatory – has a monopoly on policy influence. Wider acceptance of this would help to enhance and strengthen the impact evaluation agenda in the future, not least by finding common ground between the opposing schools.

As the previously cited paper from Public/Private Ventures notes, regardless of methods: ‘evaluations should be designed in ways that invite practitioners to make use of the results and to adopt solid practices based on evidence ... organizations and their leaders need to own and trust information in order to use it. Partnership on the ground ... is indispensable if the goal is to deliver evaluations that actually improve program quality and effectiveness’ (Public/Private Ventures, 2011).

In collective efforts to take the impact evaluation agenda forward over the next five-plus years, all of us in the international development and humanitarian commu-

nity need to be willing to *learn how to learn*. The process will often be slow and uncomfortable, not least because of what it says about the institutional and political difficulties of improving the international aid system. But the challenge is one with which we all need to grapple.

Written by Ben Ramalingam, Senior Research Associate, Overseas Development Institute (b.ramalingam.ra@odi.org.uk) and Visiting Fellow, Institute of Development Studies. The author is grateful to the following for their comments and inputs: Annette Brown, Christelle Chapoy, Alison Evans, Simon Hearn, Andy Norton, Michael Quinn Patton, Patricia Rogers and Howard White.

References

- Borton, J. (2004) 'The Joint Evaluation of Emergency Assistance to Rwanda', *Humanitarian Exchange*, Issue 26, March.
- Briceño, B. and Gaarder, M. (2009) 'Institutionalizing Evaluation: Review of International Experience'. 3ie Research Paper. London: 3ie (www.3ieimpact.org/reports/India_Report_DFIG.pdf).
- Buchanan-Smith, M. (2003) *How the Sphere Project Came Into Being: A Case Study of Policy Making in the Humanitarian Aid Sector and the Relative Influence of Research*. Working Paper 215. London: ODI.
- Catley, A, et al. (2008) *Participatory Impact evaluation: A Guide for Practitioners*. Medford, MA: Feinstein International Center, Tufts University (<http://wikis.uit.tufts.edu/confluence/display/FIC/Participatory+Impact+Assessment>).
- Center for Global Development (CGD) (2006) 'When Will We Ever Learn? Improving Lives through Impact Evaluation'. Report of the Evaluation Gap Working Group. Washington, DC: CGD (www.cgdev.org/content/publications/detail/7973).
- Chambers, R. (2009) *So that the Poor Count More: Using Participatory Methods for Impact Evaluation in Designing impact evaluations: different perspectives*. 3ie Working Paper 4. London: 3ie (www.3ieimpact.org/admin/pdfs_papers/50.pdf).
- Deaton, A. (2009) *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development*. Princeton, NJ: Research Program in Development Studies, Center for Health and Wellbeing, Princeton University (www.princeton.edu/~deaton/downloads/Instruments_of_Development.pdf).
- Deaton, A. (2010) *What Can We Learn From Randomized Control Trials?* Chicago: Milton Friedman Institute (http://mfi.uchicago.edu/events/20100225_randomizedtrials/index.shtml).
- Dhaliwal, I. (2011) *Conditional cash transfers and health: unpacking the causal chain* (<http://nonie2011.org/?q=node/88>).
- Foresti, M. (2007) 'A Comparative Study of Evaluation Policies and Practices in Development Agencies'. Report for AFD Evaluation Department. London: ODI.
- Hallam, A. (2011) 'Harnessing the power of evaluation of humanitarian action: An initiative to improve understanding and use of evaluation'. Draft ALNAP Study. London: Overseas Development Institute.
- Hulme, D. (2000) 'Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice', *World Development* 28 (1): 79-98.
- IE4D (2011) *Impact Evaluation for Development: principles for Action* (www.keystoneaccountability.org/sites/default/files/IE4DFinal%20Jan%2017.pdf).
- Jones, H. (2009) 'The "gold standard" is not a silver bullet for evaluation', ODI Opinion 127. London: ODI (www.odi.org.uk/resources/download/2811.pdf).
- Jones, N. et al. (2009) 'Improving Impact Evaluation Coordination and Use'. A Scoping Study commissioned by the DFID Evaluation Department on behalf of NONIE (www.odi.org.uk/resources/download/3177.pdf).
- The Lancet* (2004) 'The World Bank is finally embracing science'. Editorial, 28 August.
- OIOS (2008) 'Review of Results-based Management at the United Nations'. Report of the Office for Internal Oversight Services (OIOS). Paris: UN OIOS.
- Patton, M. Q. (2008a) 'State of the Art in Measuring Development Assistance'. Address to the World Bank Independent Evaluation Group, 10 April, Washington, DC (www.worldbank.org/ieg/conference/results/patton.pdf).
- Patton, M. Q. (2008b) *Utilisation-Focused Evaluation*. 4th Edn. London: Sage.
- Patton, M. Q. (2010) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.
- Public/Private Ventures (2011) *Priorities for a New Decade: Making (More) Social Programs Work (Better)* (www.ppv.org/ppv/publications/assets/329_publication.pdf).
- Proudlock, K. and Ramalingam, B. with Sandison, P. (2009) 'Improving humanitarian impact evaluation: Bridging theory and practice' in ALNAP (ed.), *ALNAP 8th Review of Humanitarian Action: Performance, Impact and Innovation*. London: Overseas Development Institute.
- Ramalingam, B. and Jones, H. with Young, J. and Reba, T. (2008) *Exploring the Science of complexity: ideas and implications for development and humanitarian efforts*. ODI Working Paper 285. London: ODI.
- Riddell, R. (2007) *Does Foreign Aid Really Work?* Oxford: OUP.
- Riddell, R. (2009) 'Is Aid Working? Is this the right question to be asking?', Open Democracy, 20 November (www.opendemocracy.net/roger-c-riddell/is-aid-working-is-this-right-question-to-be-asking).
- Rodrik, D. (2008) 'The new development economics: we shall experiment, but how shall we learn?'. Revised version of paper presented at Brookings Development Conference, 29-30 May (www.hks.harvard.edu/fs/drodrik/Research%20papers/The%20New%20Development%20Economics.pdf).
- Rogers, P. (2009) *Matching Impact Evaluation Design to the Nature of the Intervention and the Purpose of the Evaluation in Designing impact evaluations: different perspectives*. 3ie Working Paper 4. London: 3ie (www.3ieimpact.org/admin/pdfs_papers/50.pdf).
- Sandison, P. (2006) 'The Utilisation of Evaluations' in ALNAP (ed.) *ALNAP Review of Humanitarian Action 2005*. London: ODI (www.alnap.org/publications/RHA2005/rao5_Ch3.pdf).
- Stein, J.G. (2008) 'Humanitarian Organisations: Accountable – Why, to Whom, for What, and How?' in M. Barnett and T.G. Weiss (eds), *Humanitarianism in Question: Politics, Power, Ethics*. Ithaca and London: Cornell University Press.
- Weyrauch, V. and Langou, G. D. (2011) *Sound Expectations: From Impact Evaluations to Policy Change*. 3ie Working Paper 12. London: 3ie (www.3ieimpact.org/3ie_working_papers.html).
- White, H. (2011) 'Five challenges facing impact evaluation on NONIE' (<http://nonie2011.org/?q=content/post-2>).
- Young, J. and Mendizabal, E. (2009) 'Helping researchers become policy entrepreneurs', ODI Briefing Papers 53. London: ODI (www.odi.org.uk/resources/download/1127.pdf).



Overseas Development Institute, 111 Westminster Bridge Road, London SE1 7JD, Tel: +44 (0)20 7922 0300, Email: publications@odi.org.uk. This and other ODI Background Notes are available from www.odi.org.uk.

Readers are encouraged to quote or reproduce material from ODI Background Notes for their own publications, as long as they are not being sold commercially. As copyright holder, ODI requests due acknowledgement and a copy of the publication. The views presented in this paper are those of the author(s) and do not necessarily represent the views of ODI.
© Overseas Development Institute 2011. ISSN 1756-7610.