



Praxis Paper 28

Getting to Grips with Evidence

How NGOs can tackle changing needs in the use of evidence and research

by Rachel Hayman with Jamie Bartlett

November 2013

Executive summary

Pressure from funders, beneficiaries, stakeholders and public authorities for the use of more advanced research skills and robust evidence has grown over recent years. As a result, many practitioners are grappling with issues surrounding the use of evidence.

Drawing on two webinars in May and June 2013, an online survey of practitioners, a literature review on evidence use, and discussions with practitioners and researchers throughout 2013, this paper explores issues surrounding the use of evidence by practitioners and offers ways in which practitioners can better engage with evidence in their own work.

It details some of the **challenges** that practitioners face, which include:

- Poor access to relevant research and evidence
- Weak prioritising of evidence internally
- Capacity and resource constraints, including lack of staff with the right skills, lack of time for research, and limited funding for research, particularly over the long term and for exploratory research
- Different understandings about the value of various types of evidence
- Understanding the best ways to categorise evidence and assess its quality.

The paper explores different **definitions of evidence**, including its use in policy making, evaluations and advocacy. It looks at different understandings of evidence, and reflects on the social contexts and political factors that determine what counts as evidence.

Quality of evidence is a major area of contention in current debates. We explore different ways in which evidence is assessed, including evidence hierarchies and matrices, Bond's Evidence Principles, and DFID's Strengths of Evidence criteria.

Key messages:

- **Using appropriate evidence.** Practitioners need to be clear about the purpose and take time to decide what forms of evidence are relevant and practical.
- **Improving quality.** Too often non-governmental organisation (NGO) research is disregarded as being of poor quality. If practitioners want their evidence and methodologies to be taken more seriously then there are simple steps that can be taken to improve the quality.
- **Challenge capacity constraints.** This includes understanding organisational culture around the use of evidence and priority given to it. Bringing gaps to the attention of research funders may help to orient capacity building resources in the use of evidence towards practitioners.
- **Open data up to scrutiny.** Major stores of knowledge on international development sit within NGOs. Practitioners need to find ways to make their data available and invite others to use it. This includes engaging with unexpected and negative findings.
- **Share more and learn from each other.** Pooling knowledge and resources is a way of addressing capacity constraints while building up strong communities of good practice.

Introduction

“As a practitioner I need to concentrate more on the use of evidence and try to rely more on that than my personal experiences.” (Comment from a webinar participant, 27 June 2013)

At INTRAC we have observed a growing concern amongst practitioners around how to gather, use and assess evidence better. Many NGOs are increasing in-house research capabilities, contracting consultants to carry out research work, or building partnerships with academic institutions for research projects. However, practitioners still seem to be struggling.

This paper is aimed at those practitioners who have to use evidence and research in different areas of their work, but who are looking for guidance and a better understanding of the issues surrounding evidence use. The paper highlights challenges facing NGOs, and puts these in the context of wider debates about the use of evidence in policy-making and practice. It proposes practical ways in which practitioners can engage better with evidence in their work, including how they can walk a line between evidence and experience. Finally it calls on practitioners to improve how they use evidence in order to respond to some of the criticisms levelled at NGOs. The paper draws on two webinars in May and June 2013,¹ an online survey of practitioners, a review of materials on evidence use, and numerous recent discussions with practitioners and researchers.

1. Why this paper, and why all the fuss?

The relationship between research, evidence, policy and practice is a topic of long-standing debate; and some influential and high-profile reports over the years have addressed the issue in considerable depth. Evidence, what we mean by it and how we use it has also become politicised in international development and has led to some polemical debates. A good example of this is the *Big Push Forward* in the UK.²

Yet, we continue to see fairly entrenched views that development practitioners and researchers inhabit quite different worlds and roles. Practitioners are seen primarily as *users* of evidence, while researchers *produce* it. The reality is more complicated. Practitioners *use* evidence all the time, in planning, programming, analysis, monitoring and evaluation, advocacy, learning, dissemination, marketing and lobbying. But they are also constantly *producing* it within these same processes.

There is no doubt that there is increased pressure on practitioners to use higher level research skills and more robust forms of evidence to demonstrate better the impact of their work in response to demands from beneficiaries, stakeholders, public authorities, and public and private funders. As a consequence, practitioners are experimenting with new approaches, looking to increase their capacity and skills, and looking for practical ways to address the challenges they face. In parallel, for several years a number of institutions, such as the World Health Organisation and the UK's Department for International Development (DFID), have been funding capacity building initiatives and research-to-use communication strategies. However, these have focused heavily on capacity

¹ 1 May 2013 'NGOs, Evidence Policy and Practice' Development Studies Association NGOs in Development Study Group webinar (www.intrac.org/pages/en/dsa-ngos-in-development-study-group.html); 27 June 2013 'Gathering and Using Evidence: Assessing the needs of practitioners', INTRAC webinar (www.intrac.org/pages/en/research-and-learning-within-civil-society.html).

² <http://bigpushforward.net/>; Eyben, R., I. Guijt, C. Roche, C. Shutt, and B. Whitty. 2013. 'The Politics of Evidence Conference Report' *Institute of Development Studies Conference Report*, The Big Push Forward; Green, D. 2013. 'So What do I take Away from The Great Evidence Debate? Final thoughts (for now)', on *From Poverty to Power*.

building amongst policymakers and government officials, with a view to pushing for more evidence-based policy-making. Much less attention has been placed on capacity building amongst NGOs.

In response, INTRAC began pulling together observations on the use of evidence and research by practitioners early in 2013. We were concerned that practitioners were not sufficiently involved or targeted, especially outside countries where these debates are raging (notably the UK and USA). This built on work we had been involved in throughout 2012 around research collaboration amongst practitioners and researchers.³

The following represent experiences of the challenges facing civil society organisations gathered from a group of practitioners and researchers in different parts of the world through an informal exchange of ideas. Many of these challenges are well-represented in the policy-oriented literature on evidence use. The difference here is that these emerge from NGOs and NGO service providers, looking specifically at practitioners.

Practitioners often do not have easy access to evidence and research

'Access' can be understood in two ways: first, in terms of physically being able to lay hands on data and the research publications and databases in which such evidence is housed; and second, in terms of being able to comprehend it when they do get it. On the first, most academics publish their research in academic publications, or log it in the databases of research institutions, public research bodies, or international institutions. Often these publications or datasets are only available to those who can pay or who are granted privileged access. For many practitioners, these sources become inaccessible because they do not have the resources to pay for subscriptions to journals, do not have easy physical access to libraries, or do not have the institutional support to access subscription-only databases. This can be a real problem for organisations based in countries with very limited research facilities; but even in a country like the UK where civil society groups should be able to get access to libraries and datasets, this often remains a major barrier to accessing research. The drive for more academic work in international development to be made publicly available through open access is making some inroads – for example many journals are now more openly available to those in developing countries – but there is a long way to go.⁴

A huge amount of evidence is also generated by national and local governments, and other public bodies. For the international development sector, this might be in the form of research conducted by ministries or public research institutions, including statistics and surveys; or in the form of evaluations or audits of particular programmes, projects or organisations. Not everywhere are these data publicly accessible even when they are not confidential, and we are only just beginning to see more regular publication of such materials in some countries.

Physical access is by no means the only issue. If you know what you are looking for and where to find it, you can usually get hold of it by one means or another. However, there needs to be an embedded culture – personal and organisational – that provides the inclination to look for information through such sources. Resources, capacity and technology (such as slow internet connections) are also issues, which we will come back to.

But there is another aspect to this side of access that practitioners rarely reflect on; that a huge amount of evidence is generated by practitioners themselves in their programme planning, and

³ See Cracking Collaboration project outputs: www.intrac.org/pages/en/cracking-collaboration-a-new-look-at-partnerships-in-international-development-research.html

⁴ For an overview of debates around open access see: www.theguardian.com/science/open-access-scientific-publishing; <http://blog.okfn.org/2013/10/21/the-revolution-will-not-be-in-open-data/>

particularly in monitoring and evaluation work. How many NGOs make available the data gathered for these purposes so that others can use and work with them? How many archives, shelves and computers hold valuable resources that no one can access? While there may be sensitivity around how such data may be used, and implications for the NGOs and their partners or beneficiaries, this is often not the main reason. It is very often about resourcing, having a culture of information use and sharing, and having good systems to manage the use of data, including by outsiders to the organisation.

Practitioners may not find existing evidence useful

The other side of 'access' relates to the accessibility of evidence from the perspective of comprehension and usefulness. A frequent complaint amongst practitioners is that research is presented in a way that is not **useful or useable**. Articles written with academic peer-reviewers (and promotion panels) in mind can be extremely dense, use a lot of jargon and technical language, and are incomprehensible to an audience with minimal expertise on the topic. Research is often only available years after the data collection was carried out. We therefore frequently see a disconnection between research conducted by research institutions and its use by practitioners.

Local organisations also sometimes struggle to see the value in comparative work from other settings. For example, an organisation working on women's rights in rural areas of a post-communist country may not consider work on women's rights from a completely different political or geographic location to be relevant. The value of comparative research may also be limited for monitoring and evaluation work, where the data needs to be specific to the project.

Sometimes it is more than just how research is presented and communicated. There are very different ways in which research can be conceptualised, designed, and carried out. Research may be considered irrelevant or not useful by practitioners because of how it was conceptualised. This reflects debates around whose knowledge matters and power relations in research and development processes, and it is important to recognise that different individuals involved in research may understand evidence very differently depending upon their educational or professional background. As one example, INTRAC has recently been supporting civil society groups to carry out action research projects. The design teams are largely made up of volunteers, and there has been a big debate about whether anything other than large-scale quantitative methods were valid. This reflected the educational background of the individuals involved.

The question of how to communicate research is being taken more seriously by some public funders, including pushing for better stakeholder engagement in the research process. However, we cannot overlook the importance of respecting different conceptualisations of evidence and knowledge.

Practitioners often lack capacity and motivation to use evidence

Capacity constraints are numerous, and financial resources are just one. Many organisations do not have dedicated research staff or staff with sufficient knowledge to design good research or to interpret research for use. Even in the largest international development organisations, researchers may be few in number and spread across monitoring and evaluation, programmes, campaigns and communications departments. Capacity is particularly problematic for small, voluntary organisations.

One way that practitioners can address the capacity challenge is through more research partnerships and collaborations. Such partnerships are fairly common, but there are a great many

barriers, including different attitudes, perceptions, strategic priorities, timeframes, and objectives.⁵ Collaboration often depends on existing relationships, and nurturing research relationships also requires time, space and resources.

Linked to this capacity question is the extent to which NGOs **prioritise** evidence and research. For action-oriented organisations, focused on their beneficiaries and missions, research and learning may not be the most crucial place to allocate scarce resources. This leads to a lack of prioritisation of research and learning, including learning from NGOs' work and experiences, and generating or drawing on research evidence. When budget cuts are required, research, learning and knowledge are often the first to feel the squeeze.

However, not using evidence well because of a lack of capacity can be very damaging. For example, when practitioners are invited into consultation meetings with government but come unprepared and without evidence to back up their positions, this damages both the consultation process and the credibility of the practitioners.

Practitioners face competing demands in evidence use

There are multiple uses of research and evidence by practitioners which can lead to competing demands. For example, research for evaluation purposes can be very different to that required for campaigns or advocacy, drawing on different methodologies and with different purposes in mind. The drivers for research are similarly disparate, and liable to change with strategic or funding needs and opportunities. This reflects the **political economy of research**. Use of research and evidence is affected by strategic, political and economic realities, which presents a challenge to practitioners and policymakers alike. Political realities may result in judicious use of evidence dependent upon the desired message. The desire for, or emphasis on, evidence-based programme interventions may consequently be undermined by how practitioners actually inform their interventions and use the evidence they have gathered in practice. Put more bluntly, does the message drive the research or does the research drive the message? Beyond the strategic side, however, the fact that practitioners are often deeply committed to their cause may colour their openness to viewing problems and situations from other perspectives, and therefore be more open to misinterpreting, misrepresenting or quietly ignoring evidence. This critique can be equally levelled at academics, researchers and policymakers of course.

Finally, practitioners are under pressure, particularly when producing evidence of impact. Too often evidence of impact data is gathered towards the end of an activity or intervention, long before any real or lasting impact is clear. Resources for studies to take place over a long period of time or after a significant time lapse are very scarce, which affects the quality of such evidence. Resources are also rarely available for exploratory research.

2. NGO needs, uses and challenges in using evidence

These preliminary observations reflect much of the current thinking around this topic, which can be found in numerous blogs, discussion forums and papers. However, we wanted to dig deeper to test these observations amongst a wider range of practitioners and consider how we might start to address some of the challenges.

In May and June 2013 we ran a short needs assessment using an online survey and held a webinar to discuss key issues. The survey was sent out using INTRAC's electronic bulletin and to

⁵ See www.intrac.org/pages/en/cracking-collaboration-a-new-look-at-partnerships-in-international-development-research-.html

targeted organisations. It was completed by 35 respondents, amongst whom 30 international NGOs and NGO support organisations were represented, based in many different countries. The survey asked open-ended questions around the use of evidence and research, changing needs in the past two years, challenges regarding the use of research and evidence, and priorities for addressing the challenges. We then organised a webinar to discuss the findings and open up the conversation. The International Rescue Committee (IRC) gave a presentation on their experiences of using evidence. The webinar was attended by 30 participants, including representatives from 11 different NGOs from several countries, as well as some academics and consultants. Some of the participants in the webinar and survey were the same, many were not.

What do practitioners want to know?

The main reason that practitioners took part in the survey and the webinar was a desire to learn from other organisations about how to use research and evidence effectively, to gain practical ideas and advice. This included enhancing the quality and relevance of evidence generated by the organisation itself and learning how to make full use of internal research data in programming, communications and dissemination. Participants also wanted to know more about how to work with *existing*, good quality and relevant evidence and use that in programming and other activities.

Furthermore, amongst participants there was a common interest in building a culture that values research and learning for better programming and evaluation, one that values research inherently for learning, improving practice, and sharing across the sector. Breaking down barriers internally between research and implementation was another area of interest.

A final concern that brought practitioners together was around qualitative evidence: to discuss what constitutes good quality in evidence that is gathered using qualitative methods; and to discuss how to promote the value of qualitative evidence.

Use of evidence and changing needs

Not surprisingly, the respondents to our survey indicated multiple uses of evidence, as figure 1 shows. Other uses included for fundraising purposes and to demonstrate impact. Given the observation above about learning often being under-valued and under-prioritised in NGOs, it was interesting that nearly all respondents ticked the 'for learning' option.

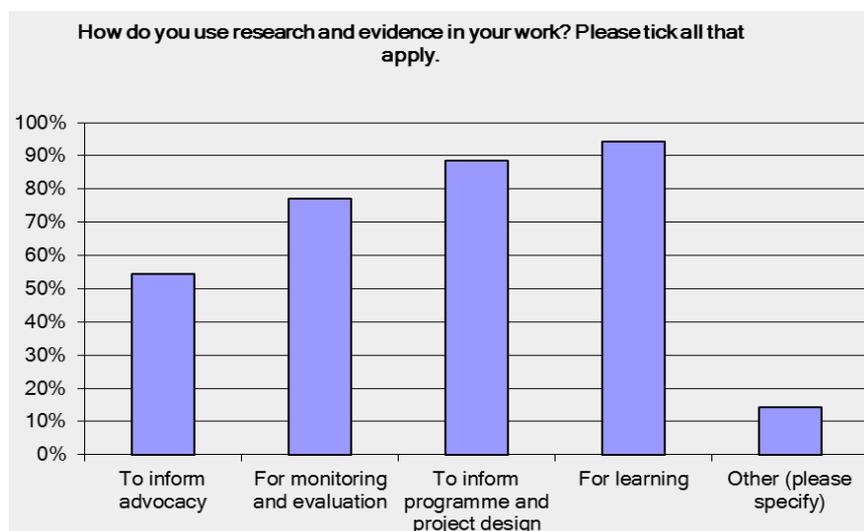


Figure 1: Use of research and evidence

To the open question of how needs around evidence have changed and why, 28% of respondents referred to changing **demands** for better data, evidence and information. When we analysed these responses further, we were able to categorise these demands into:

- downwards demands, i.e. requests or pressures coming from donors and clients
- upwards demands, i.e. coming from stakeholders or partners
- internal demands, i.e. for programming, management and internal accountability purposes.

Often the demands were multiple, covering more than one of these rough categories. The emphasis placed on upwards and internal demands is an important finding as very often there is an assumption in evidence debates that the changes are almost exclusively donor-driven. Although one feeds into the other, and the lines can be very blurred, it is important to note these different types of demands.

Survey question: please briefly explain the changes in your needs for research and evidence

“Increased demands for ‘robust’ quantitative and qualitative evidence for accountability and learning both as a result of an internal systems strengthening drive and external demands. At the same time there is interest in greater ‘beneficiary’ feedback to improve participation and programme performance.”

“The changes in our needs for research and evidence stem... from two key elements: a) the need to aggregate results organisationally and b) the [rigour] that we should aim for in doing so. These are needs closely linked to demands from institutional and strategic donors. It also responds to greater demands in terms of transparency and accountability in the sector as a whole.”

“Stakeholders have become more knowledgeable and therefore demanding, which is good. Now evidence and conclusions deriving from it need to be more clearly articulated, better sustained, and more sophisticated and diverse methods are needed.”

“Now geared more towards influencing policy audiences; previously more towards informing program managers.”

The other important change is around greater use of evidence for **influencing**:

- externally, e.g. for advocacy purposes, to influence other actors in the sector, and for accountability to donors and beneficiaries
- internally, e.g. for strategic planning and programming

Other responses we received included using research to test new ideas, approaches and methods.

Challenges faced by NGOs

Respondents in the survey were asked to list the main challenges that they faced around using evidence and research. These were also discussed during the webinar. For most participants the main challenges revolved around resources and capacity. This included **time** to search existing research, to carry out primary and secondary research, to read, analyse and reflect, and to go from research to dissemination; **staff** resources to be able to carry out sufficient gathering and analysis of evidence; and **funding**. Low **capacity** and **skills** to gather and use evidence at different levels were considered a problem, including amongst partners, in field offices and amongst programme and project staff. Indeed it is often at the country or field-office level that the greatest pressures are felt when primary data is being gathered. An understanding of – and support for – research is

therefore required throughout the system, not only in headquarters. Creating in-house research capacity and evidence 'literacy' is particularly problematic for small organisations.

A second set of challenges related to **knowledge** and **skills**, to:

- identify the evidence required in the first instance
- identify how and where to get that, including access to existing research
- identify what to use from a mass of data, some of which might be contradictory or irrelevant
- understand the evidence you are faced with.

Having the skills to make evidence accessible, and to disseminate and communicate it, were also mentioned.

"We often struggle with creating the space to reflect on evidence when people are so busy in day-to-day jobs."

"One of our biggest challenges is the varied degrees of skills across our organisation when it comes to implementing impact studies...so called 'aggregation' therefore becomes quite challenging."

"There is also the challenge of getting buy-in from all relevant staff of the value of dedicating time for evidence collection, review and identification alongside the busy work of getting projects implemented."

"Another challenge is to do with funding and program cycles. It can be really hard to integrate new learning into a program halfway through if the objectives are set for 5 years and we learn something that means we should make a huge shift."

(Written 'chat' amongst webinar participants, 27 June 2013)

Concerns about **quality** were raised, notably having confidence in the quality of existing research, and knowing how to ensure good quality in data collection and analysis, as well as quality in research design and management. The quality issue touched upon the nature of the evidence base. In many areas of international development it is hard to find easily accessible and useful syntheses of evidence to build on and work with. Practitioners also have to balance desires for quality against resource realities, including tight timescales in which to produce results or outputs, limited staff and budget.

In both the survey and the webinar there was considerable reflection on **prioritisation** and **expectations** of research and use of evidence by management on the one hand and by participants on the other. This included the need to get buy-in from many different actors to a more evidence-oriented approach. Amongst management and those involved in research there are very different views about evidence and its value, about what is feasible and the resources required for quality research. In some cases there is a 'fear factor', with people being intimidated by the research process. Differences also come out in mindset around values placed on evidence, research, knowledge and learning, and around the starting point for the process. Should we begin with the evidence or the decisions that management want to make based on the evidence? Do such differences lead to the potential misuse of evidence?

It was widely agreed in discussions that challenges around time, resources and space for better use of evidence are closely related to the value placed on evidence use, research and learning within an organisation, and the priority that is given to it. However, even where there is senior buy-

in for better use of evidence, integrating evidence into different aspects of the work of an organisation is never straightforward, especially during funding cycles or within externally-funded programmes. As the following quote demonstrates, practitioners are torn in many different directions in their use of evidence and research.

“Not having clear objectives prior to beginning a research process: are we researchers in an academic sense, or more goal oriented? Do we compromise quality of research, or our programme outcomes?” (Survey respondent on the question of challenges in the use of research and evidence)

A final area of debate was around the relevance and value of **different types of evidence**. Some participants were deeply concerned that qualitative evidence is increasingly considered as second-rate. Moreover, there was discussion about whether research should ever be taken out of context, and therefore the usefulness and relevance of generalisations based on evidence drawn from other contexts. This underscores one of our prior observations around how useful practitioners find research that is not considered relevant to the context in which they are working.

3. Stepping back: do we know what we mean by ‘evidence’?

The final point made above reflects a more contentious issue which underpins current evidence debates, namely what we mean by evidence. Very different understandings, definitions and interpretations exist. Development documentation is full of references to ‘evidence-based’ policy, ‘evidence-informed’ programming, ‘bodies of evidence’, ‘strength of evidence’, etc. What constitutes quality evidence or even ‘good enough’ evidence is highly contested and symptomatic of the politics surrounding evidence and results-based management. NGOs are consequently struggling with knowing when certain pieces of information can be counted as evidence and are juggling the different expectations of numerous stakeholders.

To help navigate this, we have drawn out two key themes from the literature on evidence: defining evidence and assessing evidence.

Defining evidence

The basic definition of ‘evidence’ is: *“the available body of facts or information indicating whether a belief or proposition is true or valid”*.⁶ However, there is no definitive guide or agreement on the meaning of evidence in international development. Evidence can come in **different forms**, depending on the origin, the purposes and the worldview behind it. We do not want to go into a long-winded debate on the philosophy (or philosophies) of knowledge, but it is important that practitioners recognise different ways of understanding ‘evidence’. Box 1 provides an example of different meanings and uses of evidence.

⁶ Oxford Dictionary of English (Online)

Box 1: Two different ways of looking at evidence

On 1 May 2013, the Study Group on NGOs in Development, organised under the Development Studies Association of the UK and Ireland, held a webinar on evidence. Two of the presentations engaged with very different forms of evidence:

Tiina Kontinen's presentation looked at **evidence in evaluations**, and explored how World Vision Finland has been looking for evidence that will act as "*a tool to let practitioners and actors know whether they are doing what they are supposed to be doing*", and simultaneously satisfy donor expectations of evidence. Using programme theories of change, they are working towards developing a quick and easy contextualised tool which can be used instead of, or alongside, traditional matrices and indicators to measure empowerment. The development and use of pathways has proved useful as they allow for the visualisation of empowerment. However challenges remain, such as how to combine the contextual paths with log frames and indicators. There are also questions about whether people's feelings of empowerment count as evidence.

Kate Gooding's presentation looked at **evidence in advocacy**, based on case studies of NGOs in Malawi. While many NGOs talk about 'evidence-based advocacy', what they mean by this varies considerably. First, the role of research in advocacy can be to identify advocacy issues, to understand them in more detail, or to prove the issues – three contrasting functions, with different implications for openness. Second, NGOs use different types of evidence in their advocacy, including formal research but also case studies of individuals, community testimony or reference to national policies. Some NGO staff feel these other kinds of evidence have more impact on decision makers than research. Third, NGOs disseminate research findings in different ways, and have different ideas about whether this dissemination constitutes advocacy. For example, some NGOs favour journal publications or conferences and view these as useful channels for influence, while others use workshops or meetings with decision makers. Several factors affect the approach to dissemination, including funding, staff capacity, and concerns about the legitimacy of INGOs to influence national policy.

Both presentations are at: www.intrac.org/pages/en/dsa-ngos-in-development-study-group.html

A key distinction in the literature is between evidence that is informed by research, and evidence that is based on more fluid forms of knowledge. This is captured in a fairly straightforward manner by both Nutley et al. (2013)⁷ and Newman et al. (2012)⁸ who differentiate between:

- evidence underpinned by research. In some literature you may also see this broken down into empirical evidence (informed by research) and theoretical evidence (which may be informed by research or by more informal, intuitive and reflective work)
- expert opinion and stakeholder views, also sometimes referred to as 'experiential' evidence, i.e. based on experience, practice or tacit knowledge.

Many practitioners will be very familiar with the latter forms of evidence as they are central to participatory work with partners and beneficiaries. Such experiential evidence often underpins campaigns-based or policy-oriented research, possibly emerging from case studies or research on specific target groups.

⁷ Nutley, S., A. Powell, and H. Davies. 2013. 'What counts as good evidence? Provocation Paper for the Alliance for Useful Evidence.' London: Alliance for Useful Evidence.

⁸ Newman, K., C. Fisher, and L. Shaxson. 2012. 'Stimulating Demand for Research Evidence: What Role for Capacity-building.' *IDS Bulletin* 43(5)

Box 2: Evidence-based policy-making

The 'evidence-based policy movement' is primarily associated with public managerialism in the UK and USA. In the literature we come across different definitions of it. Davies (2004), for example, makes a distinction between:

- evidence-based policy that puts available evidence from research at the heart of policy and implementation; and
- opinion-based policy which relies on selective use of evidence or untested views of individuals or groups.

Stern et al. (2012, 9) emphasise how the evidence-based policy movement focuses very strongly on the search for 'what works', the links between cause and effect in intervention.

A rather different perspective comes from Newman et al. (2012) who take a much broader view, describing evidence-informed policymaking as "*that which has considered a broad range of research evidence; evidence from citizens and other stakeholders; and evidence from practice and policy implementation as part of a process that considers other factors such as political realities and current public debates.*"

There is a good consensus that evidence-based policy-making is contentious and political, and that some of the underpinnings are problematic. Most notably there is a concern with the emphasis on specific forms of rigorous evidence, and the assumption that if policymakers are confronted with high quality evidence they will act upon it (see du Toit 2012; Stern 2012).

Sources: see footnote 10

More complex breakdowns are also available. Davies (2004)⁹, looking at types of evidence used in policy-making (see Box 2)¹⁰, for example, breaks 'evidence' into:

- impact evidence, for example reviews of the effectiveness of interventions, and in particular their impact on outcomes
- implementation evidence, i.e. aimed at explaining the 'how' an intervention works as well as the 'what' which is the primary focus of impact evidence
- descriptive analytical evidence, which may be based on descriptive surveys or administrative data
- public attitudes and understanding, for example evidence from perceptions surveys
- statistical modelling which uses regression methods to make assumptions about policy scenarios
- economic evidence, i.e. economic appraisal and evaluation, looking primarily at cost, cost-benefit and cost-effectiveness
- ethical evidence.

Ethical evidence requires a little more explanation. Davies highlights that in policy-making, trade-offs are standard practice and so how evidence is used is very often relative to cause, need and context. Policymakers will almost always consider much more than the evidence itself. He describes how explanations for policy decisions may be based on "*evidence of relative*

⁹ Davies, P. 2004. 'Is Evidence-Based Government Possible?' *Jerry Lee Lecture*, London: Cabinet Office.

¹⁰ Davies (2004); du Toit, A. 2012. 'Making sense of "Evidence": Notes on the Discursive Politics of Research and Pro-Poor Policy Making', *Working Paper 21*. Bellville: PLAAS, UWC; Newman, Fisher, and Shaxson (2012); Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies, and B. Befani. 2012. 'Broadening the Range of Designs and Methods for Impact Evaluations: Report of a study commissioned by the Department for International Development' *Working Paper 38*. London: DFID.

effectiveness, relative costs, people's perceptions and experiences, and of the social justice and ethics of [making trade-offs]."

A further distinction you may come across is between *Big Evidence* and *small evidence*, terms particularly associated with the work of *The Big Push Forward*,¹¹ and reflect debates around the politics of evidence:

- Big Evidence (Big E) is defined as evidence of what works under which conditions for whom
- Small evidence (small e) is defined as evidence about performance and results monitoring

A final way of thinking about evidence is in terms of a sliding scale of certainty. At the time of the webinar in June 2013 this was one of several approaches being explored by the IRC.¹² The 'best evidence' (at one extreme of the spectrum) attributes specific effects to specific programmes across contexts with a very high degree of certainty. Generally, the best evidence is considered to come from a robust body of high-quality research across several contexts. Additional research in the near-term would be expected to support existing conclusions, allowing practitioners to make programming decisions with confidence that they are providing positive impacts to the people they serve and avoiding negative impacts. Additional resources for learning or research should likely not be invested in programmes supported by the best evidence.

'Information' (at the other extreme of the spectrum) is defined as coming from experience or expert opinion and can communicate valid points, though with minimal certainty that these assertions reflect reality. Additional research will very likely change our understanding of which programme works best, on what outcomes, to what extent, for whom and in what contexts. Programmes supported by information only may be strong candidates for further investment in additional research and learning or a conceptual review of the logical framework.

For the IRC, the degree of certainty could be determined through judging the information against a checklist which asks:

- Does the study that it is from use the appropriate research design to answer the key question(s)?
- How strong is the study's internal and external validity?¹³
- How precise are the findings?
- How consistent are the findings in comparison to other studies?
- How do contextual factors or implementation factors influence the findings?

The best evidence likely would have an appropriate research design, a study with strong internal and external validity, precise findings, findings consistent with the wider body of knowledge on the subject, and a thorough understanding of contextual or implementation factors that influence the magnitude of outcomes.

Assessing evidence: strong versus good enough

Defining evidence is closely linked with categorising evidence and its quality, as the IRC example shows. There are various standards and tools that practitioners can engage with when considering how they rate evidence, whether in their own research or in evidence they draw on from elsewhere. What we provide here is a very brief overview of some of these tools and standards to

¹¹ Big Push Forward. 2013. 'The Politics of Evidence: Values, Choices and Practices for Evaluating Transformative Development' on *The Big Push Forward*.

¹² Presentation by Anjali Shivshanker, IRC, 27 June 2013. For more information contact: Anjali.Shivshanker@rescue.org

¹³ 'Validity' refers to whether the research findings really are about what they say they are about. Internal validity refers to the integrity of the findings, including whether the right variables were used to make causal claims. External validity refers to whether the results of the research can be generalised beyond the study.

raise awareness of the different approaches that exist. These differ across academic disciplines and development intervention sectors, and are all subject to critique from other perspectives. There is no single standard and we do not seek to orient NGOs in any particular directions. You may find that different standards are useful in different aspects of your work, for example standards for assessing quality of evidence in the health sector may not be useful when assessing the quality of evidence in empowerment projects. More importantly, you may find that your funders or stakeholders have specific expectations around what constitutes quality evidence, and you may have to be prepared to adapt.

Evidence **hierarchies** are often used in the ‘what works’ impact-oriented evidence literature. They have their origins in medical research.¹⁴ Although hierarchies vary, research based on randomised control trials and bodies of evidence that have been through a systematic review process (see Box 3) will sit near the top of the hierarchy. Observational, case study-based, empirical research will inevitably be found at the bottom (see Nutley et al, 2013).¹⁵¹⁶ Hierarchies have been criticised for reasons such as: emphasising study design over how the study was implemented; under-rating the value of good observational studies; excluding useful evidence by only focusing on the highest ranking studies; not taking sufficient account of theory and questions of what works and why; and not providing a basis for making recommendations. Primarily, critics are concerned that hierarchies focus too heavily on what works and not enough on how, why, etc. In such hierarchies, the methodology provides the first stage of categorising; quality is then based on an assessment of how well the research highlights study limitations, inconsistency of results, indirectness of evidence, imprecisions in the data and reporting bias. Similar to the IRC’s definition, a high quality study is one where the reader can be confident that further research would not change the results in any significant way.

Matrices for categorising evidence are more common in social policy, looking beyond ‘what works’ to ‘how’ the research/policy works and whether it matters to stakeholders. Matrices will pose a range of questions about the research, such as how does this work and is it right for these people? Quality assessment of the research will consider issues such as appropriateness in study method, appropriateness of samples, and context analysis. Annex 1 gives an example of a matrix.

¹⁴ See, for example, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system: www.gradeworkinggroup.org

¹⁵ Nutley, Powell, and Davies (2013)

¹⁶ Donaldson, S., C.A. Christie and M.A. Mark. 2009. *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* Los Angeles: Sage.

Box 3: Systematic reviews and bodies of evidence

A systematic review is a method that has traditionally been used in the medical sciences to assess 'bodies of evidence', i.e. multiple studies. Systematic reviews use a very strict process to identify and analyse research studies which address a research question. The review is not just a literature or narrative review, but aimed at extracting, testing and synthesising the evidence provided in research studies. Systematic reviews are increasingly being used in international development, notably by organisations such as 3ie, DFID and AusAID.

Example: in 2011 a team from the University of Edinburgh attempted a systematic review on the impact of different forms of aid on maternal health outcomes. The question was very complicated and not well suited to a systematic review. We nevertheless went through a painful process of trawling for studies that met particular inclusion criteria based on key words around aid, maternal health and developing countries. But more importantly we had to decide whether to include studies in the review based on the study design and methodology. The guidelines that we were working to for the review favoured experimental methods (randomised-control trials and quasi-experimental studies); observational studies based on qualitative methods were excluded. If no clear primary quantitative data was presented that demonstrated change in maternal health indicators as a result of an aid intervention, then the study was rejected. There was no real interest in processes of change, just the actual input and outcome data.

Studies that met these criteria were then subjected to a quality assessment which looked at study independence and potential bias, information on the types and flows of aid, study design, study methods, how data analysis was conducted, and alternative explanations for the results (confounding factors).

There is much debate around the value of systematic reviews in international development, because typically they focus on hierarchies of evidence and will therefore exclude information and reports that use less rigorous methods. Certainly in our review we were very concerned that we were missing lots of potentially valuable information because reports just did not quite fit our criteria. However, systematic reviews are a growing, and publicly available, resource summarising evidence on a range of international development topics. This is information that NGOs can work with when developing programmes or studies even if there are limitations in the approach. See: <http://r4d.dfid.gov.uk/SystematicReviews.aspx>

Source: Hayman, R., E. M. Taylor, F. Crawford, P. Jeffery, J. Smith, and I. Harper. 2011. 'The impact of aid on maternal and reproductive health. A systematic review to evaluate the effect of aid on the outcomes of Millennium Development Goal 5.' London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Tools and standards for assessing quality of evidence in international development research are becoming more widely available. Box 4 provides two very different examples. Approaches and priorities can vary widely, but ultimately most assessments of quality cover:

- Methodology used: was the design good in the first place? Design may cover the processes through which the research was carried out as well as the methods used
- Transparency: how much information is provided to the reader/user on how the evidence was generated, the limitations and potential biases in the data
- Credibility and trustworthiness of the data. This is sometimes referred to as the 'validity' of the evidence.

Box 4: Bond's Evidence Principles and DFID's Strength of Evidence

Two very different perspectives on quality in assessing evidence are provided by Bond's NGO Evidence Principles and DFID's Assessing the Strength of Evidence note.

In 2012-13 Bond (the British network of development NGOs) developed a guide and checklist for NGOs to assess the quality of evidence that they are collecting to measure and demonstrate their impact. The aim is to set common standards by NGOs for NGOs, building on NGO values and types of intervention. The standards aim to be practical and relevant, and assume that NGOs are using mostly qualitative or non-statistical methods. The NGO Evidence Principles toolkit was designed in consultation with many NGOs and is structured around five principles:

- Voice and inclusion, capturing beneficiary views on the intervention, who has been affected and how
- Appropriateness, whether methods are justifiable given the nature of the intervention and the purpose of the assessment
- Triangulation, i.e. drawing on a mix of methods, sources and perspectives
- Contribution, showing how change happens and how the intervention on contributed to that change
- Transparency, i.e. openness about data sources, methods, results and limitations.

A rating scale is used for each principle to determine whether the evidence can be assessed as weak, minimum standard, good practice, or gold standard.

See: www.bond.org.uk/effectiveness/principles; Bond. 2013. 'The NGO Evidence Principles pilot phase: October 2012-31 May 2013', Worksheets. London: Bond.

DFID produced a note for staff in February 2013 as a guide for assessing evidence, the purpose being to ensure that research "*is evaluated in a fair and balanced way*". It provides standards for assessing single studies and bodies of evidence (see Box 3). Single studies are ranked as high, moderate or low in quality against what it terms "*principles of credible research enquiry*" (p.10).

These principles are:

- Reflexivity and ethics of the researcher, i.e. how does the researcher situate themselves and what bias may enter the study because of the researcher's influence or position?
- Conceptual framing of the research, i.e. how does the research sit within existing knowledge, what are the assumptions, questions and hypotheses?
- Openness and transparency, i.e. how well are the design, methods and limitations explained?
- Appropriateness and rigour
- Validity, which includes measurement validity (are the right indicators being used), internal validity (is the technique used to explore causal claims satisfactory) and external validity (how generalisable are the findings beyond the study itself)
- Reliability, i.e. consistency and accuracy in the data
- Cogency in reporting, i.e. is there a clear, logical argumentative thread?

Assessments of bodies of evidence look at the quality of the single studies and the consistency in findings across a number of studies. Bodies of evidence are ranked as very strong, strong, medium, limited, or no evidence (p.20)

See: DFID. 2013. 'How to note: Assessing the Strength of Evidence', Practice Paper. London: DFID

Finally, it is also useful to reflect on what is viewed as poor evidence. For Davies (2004)¹⁷ poor quality evidence has:

- Unclear objectives
- Poor research design
- Methodological weaknesses
- Inadequate statistical reporting and analysis
- Selective use of data
- Conclusions that are not supported by data

This brief exploration of the literature and some examples demonstrates that there is no universally accepted and simple way of describing, categorising or assessing evidence. Inevitably there are political and power dynamics at work in the different ways in which evidence can be assessed and understood, and practitioners need to keep this in mind and constantly question how evidence is ranked.

"Scales for assessing the quality of evidence do exist... Such scales are not neutral or apolitical; they are built on particular assumptions about the ordering of evidence and, as such, it is important to question their foundations."

Source: Hagen-Zanker, J., and R. Mallett. 2013. 'How to do a rigorous, evidence-focused literature review in international development' Working Paper, London: Overseas Development Institute. p.15

4. Going forward: gathering and using appropriate evidence

Where do our preliminary observations, the findings emerging from our needs assessment activities, and our exploration of the literature take us? At the heart of the issue is what we want to use evidence for, how we want to or can gather it, and how to ensure that it is fit for purpose.

Appropriate evidence – ensuring it is fit for purpose

"Evidence quality depends on what we want to know, why we want to know it and how we envisage that evidence being used. In varying contexts, what counts as good evidence will also vary considerably."

Source: Nutley, Powell and Davies (2013, 6)

Practitioners are facing changing demands and pressures in how they engage with evidence. We observe a range of limitations that practitioners have to overcome, and we also recognise multiple purposes for which evidence is required.

This demands that practitioners reflect long and hard on ensuring that the type of evidence they are interested in or producing is fit for *purpose*, i.e. using appropriate evidence that is relevant for the specific need in hand. Evidence within programming or advocacy or evaluation or exploratory research can be very different. For an advocacy campaign, the most useful and relevant evidence may be some powerful, real-life and credible stories that will get the attention of those you are seeking to influence. For a high-level impact evaluation on a large programme in the health sector funded by a donor, you may require evidence that is extremely scientific and verifiable. On the

¹⁷ Davies (2004)

other hand, evidence for learning purposes may be much more exploratory, reflective and investigative.

Thinking about the purpose of evidence should also involve some reflection on the values that underpin the work of the organisation, a degree of reflexivity on the part of researchers and those commissioning research. The Bond Evidence Principles highlight a dimension that will resonate with many practitioners, namely including the voice of people who are affected by interventions.

The key factor is being confident about what sort of evidence you require, why you require it, and how you go about gathering it within the boundaries of practical limitations you may face and the principles and values you hold. At the same time, this has to sit within a clear understanding of the social and political context in which the evidence is being gathered or used. Despite the efforts of the evidence-based policy movement, the reality is that policymakers are driven by multiple incentives in their decision-making. Research-based evidence may come well below persuasive, politically-timely or actionable information. Power dynamics, political structures, institutional politics, and social contexts all play a role in how evidence is used. NGOs are affected by the politics and whims of the policymakers on the one hand, and will need to adapt their approach to evidence depending on the funder or accountability chain. On the other hand, however, the same issues also apply to NGOs themselves.

"If... we are to understand the role of evidence we need to look at how policy narratives work with it; how it is used; and how it is alternately marginalised or seized on, ignored or imbued with significance."

Source: du Toit (2012, 6)

Quality in evidence – being fit for purpose

There is a very big difference, in our opinion, between criticisms of evidence on the basis of its methodological approach (i.e. whether you use a randomised control trial or case studies) and criticisms of evidence on the basis of its quality (i.e. how well it is done).

If practitioners are to be confident in arguing for different forms of evidence for different purposes then they need to be doing this on the basis of that evidence being of a high standard. This is where the evidence push has arguably its greatest benefits for NGOs, in encouraging – if not pushing – for better quality in the production and use of evidence.

Based on our observations of NGO research and evidence use, we encourage practitioners to think through the following issues:

- How seriously do we review literature and evidence that already exists before gathering new primary data or commissioning another study? Are literature reviews tokenistic rather than in-depth because of constraints caused by time, in-house resources and access? If NGOs do not carry out good quality literature reviews at the outset, then they could easily be wasting time and money because they are reinventing the wheel rather than building on existing knowledge.
- How seriously do we analyse the data we gather, including looking at interesting information that emerges from primary data that may not be directly relevant to the work in question? Time and resource constraints seem to lead to a lot of unexplored and under-analysed data, particularly qualitative data that takes a lot of time to analyse well. However, there is no reason why those data should not be explored by others or revisited at another time.

- How well do we document and describe the methods we are using, the limitations of the study, and potential bias? Do we even have quality assurance standards in place and use them?
- How well do we report on unexpected or negative findings? Does the political context of NGO research lead us too quickly to gently sideline the data that does not quite fit the message? Yet such findings are vitally important for both organisational learning and for the wider sector.

In our needs assessment work the concern came out again and again about the value given to qualitative methods. Many of the biggest proponents of rigorous evidence recognise that qualitative and observational data have value, at least in their rhetoric. Practitioners should continue to fight this cause. However, if we are going to champion qualitative data, beneficiary voice and context-specific research – which some would describe as ‘weak’ forms of evidence – then we need to have the strongest possible quality in that work. This can be as simple as explaining fully what you did (processes), why it was appropriate (justifying the methods), what the limitations were in the design including contextual factors, and including this information in reports. Box 5 illustrates this.

Box 5: Aid to maternal health – why NGO studies rarely made the grade

In the systematic review introduced in Box 3 on the impact of different forms of aid on maternal health outcomes, very few studies by NGOs were included. The initial search of databases using a list of key terms around aid and maternal health brought up almost 1,700 sources. 211 of these were selected as potentially relevant after further exploration. Only 30 were included in the final pool for full analysis. Of these only three were studies of healthcare interventions by NGOs. So why did NGO material rarely show up in the searches and included studies when we know lots of NGOs channel aid to maternal health? There were big questions about the entire review methodology which partly explain this. For example, if the study did not actually mention aid and external financing in some way it would not have come through the initial search. And we were looking for evidence of impact which is extremely difficult to demonstrate if an evaluation of an intervention takes place before enough time has elapsed.

However, there are important lessons for NGOs to consider. Firstly, of the 211 studies included initially for full assessment, 154 were excluded because they failed to meet crucial ‘evidence of impact’ criteria. Reports had to provide sufficient data and information on the methodology. Many of our included studies – amongst them many from NGOs – just did not do this. They included policy briefings and evaluation reports which failed to provide sufficiently robust data. It was not good enough to present anecdotes from interviews or focus group discussions. There needed to be a strong analysis, based on statistics (which could be drawn from qualitative data) and ideally a good baseline that showed change in health outcomes over time.

Secondly, many reports scored weakly on quality assessment criteria. Too many studies just did not provide full information on the design of the study, methods used, alternative explanations for the results, and potential bias due to study design, context or lack of independence.

Alternatively, we think this might also be symptomatic of reports not being in the public domain, not subject to peer-review, not published nor made available in some way. An evaluation report that sits on a shelf does not figure in the available pool of knowledge for review.

The main message from this is that if practitioners want their evidence and their methodologies to be taken seriously, then they need to open themselves up to scrutiny to the extent possible and

make information available so that it contributes to global pools of knowledge. This will then provide a push factor to higher quality reporting.

Finally, we need to beware of dressing up information and data as evidence when it is not. This requires a good idea of what counts as 'good evidence' for stakeholders in the particular activity. However, it is also important not to deny the value of different types of information, and undermine its worth by making claims that do not stand up to scrutiny.

Delivering better quality evidence

Such posturing does not of course make it any easier for an over-stretched M&E department or research team in an NGO (if you are lucky enough to have one) to overcome the major challenges of limited skills, capacity, resources, differing worldviews and management or partner buy-in.

Some thoughts on what practitioners could do to start addressing some of the challenges:

Standards in evidence and research: do you have policies and procedures to categorise and understand evidence and its purposes in the work of your organisation? What policies and procedures do you have to assess the quality of evidence that you generate yourselves or that you commission from consultants? Are these fit for purpose? If you do not have any policies and procedures, what existing systems could you build on, for example those outlined in this paper?

Capacity gaps: can you do a stock-take of capacity and skills gaps around use of evidence and consider what you could and should address? Practitioners need to be honest about their skills and potential weaknesses. One thing that comes out in the reviews of capacity to use evidence amongst policymakers is that people tend to over-play their skills when they do self-assessments. We therefore need good ways of assessing capacity gaps. This is not necessarily about building up in-house capacity as we recognise that many organisations cannot do this. However, staff who are responsible for bringing in external consultants to help with reviews, research and programming also need to be able to engage with the work those consultants do. If the commissioning staff cannot comprehend the forms of evidence and the language being used, then they cannot adequately control for quality and ultimately make the best possible use over time of the evidence that is produced.

There are a growing number of programmes aimed at building the capacity of policymakers to use evidence. Practitioners need to engage with these programmes and arguably push for similar programmes for NGOs.

Organisational culture and buy-in: what obstacles are there to using research and evidence in your organisation? If we do not know what interventions work and how they work, how can NGOs achieve their objectives and fulfil their missions? Learning and evidence are not a luxury but a necessity to provide accountability to the people on behalf of whom organisations are acting or speaking, and to provide accountability to funders. Consensus is not required on meanings of evidence and its use, but we encourage NGOs to think again about how they engage with evidence.

Engage with the negative as well as the positive: we recognise that there is a lot of pressure on NGOs to demonstrate the positive impact of their work, and that there are concerns that highlighting failure may jeopardise funding. Research often throws up unexpected results. Quality and credibility of practitioner research will be greatly enhanced by engaging with unexpected and negative results. This is important for learning internally and more broadly in the sector.

Making better use of evidence

There are many ways in which better use of evidence will improve the working practice of NGOs. Here are some thoughts on how NGOs can make evidence work for them:

Thinking through uses of evidence: how do you currently use evidence? What other areas of your work could it improve, for example organisational improvement, strategic planning, accountability in its different forms, and learning for the next programme or project?

Take time when planning for research and evidence use: you need to take (and find) the time at the outset of a piece of work to explore the options, to get buy-in from different stakeholders, to clarify the purposes, and to pre-empt the questions about quality in methods and results.

Be open with the data you have: there are huge gaps in knowledge about many development issues and many parts of the world. We would strongly encourage practitioners to be more open about the evidence and information they have. This includes raw datasets as well as evaluation reports and programme documentation. NGOs often work in places that academic researchers cannot reach. This leads to big gaps in academic understanding about the issues facing people in fragile or remote contexts. NGOs should actively seek to share what they know and, just as importantly, what they do not know so that researchers can be encouraged to investigate the questions that practitioners want answered.

Be open to learning and sharing: we find there is a great appetite for sharing and learning amongst NGOs, and encourage practitioners to invest more time in this for the wider benefit of the sector. On relevant issues this should include not only practitioners, but also researchers, academics and policymakers.

Make use of what is available: there are more useful materials available than many practitioners realise. At the end of the paper we provide some useful resources that we have used in this paper and other work, which may help to address some of the challenges raised.

5. Concluding thoughts

At the start of this paper we quoted a webinar participant whose comment captured a key shift in the use of evidence by practitioners: the drive to move away from personal (experiential) knowledge towards 'real' evidence. This reflects major contradictions and contestations in evidence debates, as the following quote from du Toit emphasises.

"This is political terrain. Rather than a neat alignment between the interests of socially concerned researchers and rational, evidence-oriented planners, we find a complex process of struggle."

Source: du Toit (2012, 10)

The high-profile Politics of Evidence conferences and dialogues in the UK in 2013 attest to the political divides between proponents and opponents of 'evidence' in international development. But similar tensions exist between researchers, managers, communities and people when faced with the gathering and using of knowledge. Knowing what you are using, when, how and why are the building blocks for engaging better with evidence in development and are vital to managing these tensions.

We conclude with some final take-home messages that we hope may help practitioners to cope better with evidence debates.

There is no one-size-fits-all answer to the ‘what do we mean by evidence’ question

There is no definitive checklist to guide you in understanding or assessing evidence. But there are tools you can work with to devise your own standards that are appropriate to the particular piece of work in question.

Don’t discount what you know

Personal experiences and testimonies, well focused, well analysed, and tested from different angles, can form a critical mass of knowledge, which in many areas of development is crucial evidence and sometimes is the only evidence that is available. Such experiential knowledge has its value when used appropriately.

There is no excuse for poor quality

For us this is a critical message. Whatever your evidence is for, wherever it sits on ‘evidence’ ranking scales in terms of methodology, you need to be confident that what you have produced is appropriate to the purpose. Even small organisations, with limited resources and practical constraints can produce good quality research. That way practitioners can engage constructively in debates about appropriate evidence because we can be confident that what we are producing stands up to scrutiny.

Learn and share

Resources and capacity are big problems of everyone. Sharing ideas, lessons and good practice amongst each other is one way to build a strong community of practice in evidence use amongst practitioners. And by ensuring that donors, funders and managers are aware of the limitations may help with building up the case for addressing some of the constraints.

Acknowledgements

Thanks to all the organisations and individuals who took part in the survey and the webinar; and thanks go to the following people who provided comments on and input to this paper: Anjali Shivshanker, Kate Gooding, Tiina Kontinen, Sarah Mistry, Sandra Nutley, and Jonathan Breckon.

Finally, we are extremely grateful to INTRAC’s 2013-14 NGO Programme Members who support INTRAC’s research team to produce papers and convene learning activities: Broederlijk Delen, Cordaid, DanChurch Aid, ICCO, Norwegian Church Aid, Save the Children Denmark.

Useful open-access resources on evidence

- Understanding evidence - we recommend the following papers and blogs:
 - 'What counts as good evidence', by Nutley, Powell and Davies, 2013. This paper is UK-focused but is clearly written and provides very useful examples, annexes and links on standards in evidence. www.alliance4usefulevidence.org/assets/What-Counts-as-Good-Evidence-WEB.pdf
 - 'Broadening the Range of Designs and Methods for Impact Evaluations: Report of a study commissioned by the Department for International Development' by E. Stern et al., 2012. Chapter 6 is very useful on quality in evidence. http://r4d.dfid.gov.uk/pdf/outputs/misc_infocomm/DFIDWorkingPaper38.pdf
 - Follow debates on these blogs: The Big Push Forward (<http://bigpushforward.net>), Kirsty Evidence (<http://kirstyevidence.wordpress.com>)
- Bond Evidence Principles (www.bond.org.uk/effectiveness/principles): a handy toolkit for practitioners to assess their evidence against standards devised by UK-based international development NGOs.
- INASP (www.inasp.info/en): a lot of the information available on INASP's website is aimed at academics and librarians in developing countries, but they are also involved in capacity building programmes for policymakers. The website includes links to databases of open access materials.
- Research for Development (<http://r4d.dfid.gov.uk>): includes a link to systematic reviews funded by DFID.
- Evidence-informed policy making (www.who.int/evidence/en): supported by the World Health Organisation, the materials on this site are health-focused.
- Overseas Development Institute (ODI) (www.odi.org.uk/programmes/rapid): the research and development programme (RAPID) has produced some useful materials and toolkits over the years. These are very oriented at evidence for policy-making.
- The University of Leeds provides a range of papers on realist approaches to evidence (www.leeds.ac.uk/realistsynthesis/supreadings.htm)
- Critical Appraisal Skills Programme (CASP) has a range of resources available. These are at the more technical end of the evidence spectrum (www.casp-uk.net).
- The GRADE working group site (www.gradeworkinggroup.org/index.htm) provides a range of papers and resources. These are primarily health focused.
- 3ie provide a wide range of papers on evidence in international development, including systematic reviews (www.3ieimpact.org/en/evidence)

References

Big Push Forward. 2013. 'The Politics of Evidence: Values, Choices and Practices for Evaluating Transformative Development' on *The Big Push Forward*.

Bond. 2013. 'The NGO Evidence Principles pilot phase: October 2012-31 May 2013', *Worksheets*. London: Bond.

Davies, P. 2004. 'Is Evidence-Based Government Possible?' *Jerry Lee Lecture*. London: Cabinet Office.

DFID. 2013. 'How to note: Assessing the Strength of Evidence', *Practice Paper*. London: DFID. www.gov.uk/government/uploads/system/uploads/attachment_data/file/158000/HtN_-_Strength_of_Evidence.pdf

Donaldson, S. 2009. 'A Practitioner's Guide for Gathering Credible Evidence in the Evidence-Based Global Society.' In *What Counts as Credible Evidence in Applied Research and Evaluation Practice?*, S.I. Donaldson, C.A. Christie and M.A. Mark (eds.). Los Angeles: Sage.

Du Toit, A. 2012. 'Making sense of "Evidence": Notes on the Discursive Politics of Research and Pro-Poor Policy Making' *Working Paper 21*. Bellville: PLAAS, UWC.

Eyben, R., I. Guijt, C. Roche, C. Shutt, and B. Whitty. 2013. 'The Politics of Evidence Conference Report' *Institute of Development Studies Conference Report*. The Big Push Forward.

Green, D. 2013. 'So What do I take Away from The Great Evidence Debate? Final thoughts (for now)', on *From Poverty to Power*.

Hagen-Zanker, J., and R. Mallett. 2013. 'How to do a rigorous, evidence-focused literature review in international development' *Working Paper*. London: Overseas Development Institute.

Hayman, R., E. M. Taylor, F. Crawford, P. Jeffery, J. Smith, and I. Harper. 2011. 'The impact of aid on maternal and reproductive health. A systematic review to evaluate the effect of aid on the outcomes of Millennium Development Goal 5. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, London, UK.

Newman, K., C. Fisher, and L. Shaxson. 2012. 'Stimulating Demand for Research Evidence: What Role for Capacity-building?' *IDS Bulletin* 43 (5).

Nutley, S., A. Powell, and H. Davies. 2013. 'What counts as good evidence? Provocation Paper for the Alliance for Useful Evidence', London: Alliance for Useful Evidence.

Oxford Dictionary of English (Online)

Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies, and B. Befani. 2012. 'Broadening the Range of Designs and Methods for Impact Evaluations: Report of a study commissioned by the Department for International Development' *Working Paper* 38. London: Department for International Development.

Annex 1: Example of a matrix

Reproduced with permission from Nutley, Powell and Davies (2013: 16)

Box 4: A matrix of evidence to address various aspects of 'does this work?'

| Research question | Qualitative research | Survey | Case-control studies | Cohort studies | RCTs | Quasi-experimental studies | Non-experimental studies | Systematic reviews |
|--|----------------------|--------|----------------------|----------------|------|----------------------------|--------------------------|--------------------|
| Does doing this work better than doing that? | | | | + | ++ | + | | +++ |
| How does it work? | ++ | + | | | | | + | +++ |
| Does it matter? | ++ | ++ | | | | | | +++ |
| Will it do more good than harm? | + | | + | + | ++ | + | + | +++ |
| Will service users be willing to or want to take up the service offered? | ++ | + | | | + | + | + | +++ |
| Is it worth buying this service? | | | | | ++ | | | +++ |
| Is it the right service for these people? | ++ | ++ | | | | | | ++ |
| Are users, providers, and other stakeholders satisfied with the service? | ++ | ++ | + | + | | | | + |

Source: Adapted from Petticrew and Roberts 2003, Table 1, p.528.