



Evidence in Development – A Mokoro Seminar, 30 May 2018

By Rebecca Rhodes



There is a growing demand for the production of evidence in the international development sector to ensure that programmes and initiatives are generating their intended results. For those working in the sector there is now a familiar cycle of reviews, assessments and evaluations designed to serve the mutually reinforcing purposes of accountability and learning. These tools for monitoring and evaluation may come at a considerable cost, both to stakeholders in terms of time invested and to donors, governments and NGOs in terms of financial expense. However, they are now a routine part of programme delivery and a key feature of the current development landscape. Their inclusion as routine may be attributed to a political climate which, it can be argued, is increasingly sceptical of the value of development spending.

In Mokoro's most recent seminar, which was hosted in partnership with the Oxford Network of International Consultants (ONIC), we were joined by experienced development professionals, who explored and reflected on this highly relevant and evocative subject. The seminar was chaired by Mokoro Director and consultant Adam Leach with speakers Abby Riddell (Mokoro, independent consultant), Jon Bennett (ONIC, independent consultant), John Rowley (ONIC, independent consultant) and Claire Hutchings (Oxfam GB).

Abby Riddell

Abby Riddell opened the presentations with a retrospective approach, reflecting on previous trends and asking what the key questions were, looking ahead. As a senior education expert who has undertaken evaluations for 40 years, Abby centred her presentation on the education sector and the ways in which quality measures and learning achievement measures have changed within it – unfortunately not always to the benefit of programme 'beneficiaries'. She noted with disappointment that what she perceived to be an increasing complexity in evaluation methodologies and a greater frequency of evaluations had not led to better policy outcomes. She also pointed to the development of a sharper focus on value for money, noting that the desire to achieve the most 'bang for your buck' in terms of educational achievement had pitted finance ministries against education specialists.

Abby felt it was necessary to consider whether we have the capacity to and whether we do in practice meet different stakeholders' objectives in pursuit of evidence generation; do the evaluations we conduct meet the needs of programme beneficiaries in terms of content and methodology? Abby questioned whether, as evaluators, we are capable of challenging poor practices and actually do so, whether by rejecting unrealistic timeframes, challenging over-emphasis on value for money or opposing the use of 'fudged' data, and she posed the challenge to the audience of being more 'honest' in evidence generation. Abby was critical of the use of 'big data' meta-analysis and randomised controlled trials in the generation of evidence in the education sector, which in her experience leads to answers based on individual variables and specific questions which fail to address issues of systemic educational reform, the 'big questions' a minister of education requires solutions to in her portfolio of education policies.

Abby pointed to the growth in the complexity of evaluations and in statistical modelling as part of evaluations as an indication that what is valued in terms of the sort of data generated by evaluations may have changed. This also led her to question whether there has been a shift in who evaluations are undertaken for, with an increasing emphasis on generation of evidence for accountability purposes rather than the production of meaningful data and opportunities for learning for key stakeholders. There is also a question of who shapes the drive for and use of evidence and who may be

left out of the process. Demands for evidence are often imposed by donors, with other key stakeholders' requirements for the shape of the evaluation and/or the kind of data generated often left as an afterthought, rather than being a central consideration. The nature of the data generated may also exclude some stakeholders. Randomised controlled trials, for example, may be highly valuable, appropriate tools in certain settings, but capacity development, which should be included in the evaluation process in order to maximise the utility of the study, is often absent from them. Abby questioned whether and how we might make the process of generating evidence more inclusive, to improve the quality of the data generated and its usefulness.

Finishing on an upbeat note, Abby commented that in this atmosphere of 'fake news', she was delighted that evidence-based programming is what we hold as important.

Jon Bennett

The next to present was Jon Bennett who shared his thoughts and experience on the involvement of intermediaries and how we might navigate the plethora of methodologies available to those generating evidence. Reflecting on the growth of the role of intermediaries, Jon considered the 1990s as a decade in which a fundamental change was observed in the role of development practitioners, as intermediaries gradually came to occupy the space between donors and recipients. A specialisation took place, in which individuals and organisations in the sector became its specialists and regulators. Jon said that this change was particularly notable in the field of evaluation, with evaluators becoming desk-bound. Jon regretted that as a result, the new generation of development practitioners tend not to have the field experience required and look to create new, often complex, systems of evidence acquisition as an alternative means of 'making their mark'.

In the drive for the generation of evidence in development, Jon raised the question of who the evidence is produced for. He observed that the new methodologies employed in evidence generation create systems of knowledge which are top heavy and owned by experts and specialists rather than by the key stakeholders themselves. Evaluation products are often vast and impenetrable, thick documents not much read and least of all by beneficiaries. The data produced is oriented towards the needs of the donor client, with little regard for the need or capacity for learning outcomes. For Jon, this amounts to an inevitable control of the means of production and transmission of knowledge by 'specialists'. He highlighted the example of the 2004 Boxing Day tsunami response, in which the organisations on the ground were the first to respond and held key information on the status of those affected, but the aid community then descended, generated reams of documents and vast sums in aid revenue and in this way presented themselves as the 'first responders' on the scene. Jon provided a further example from an evaluation he supported in Syria which centred on local governance and accountability. He questioned the validity and timing of an evaluation of impact of measures to stimulate local governance in a nation ravaged by years of war. For him, this example reflected the misalignment between the choices of donors and of stakeholder communities.



Jon Bennett presenting

Jon argued that the increasingly sophisticated and complex methodologies which are produced in the pursuit of evidence generation put increasing pressure on those charged with data gathering and lead to increasing anxiety. With each new methodology it is necessary to carry out another round of data gathering, but there is a failure to ensure that the ground-level data collected is of sufficient quality, and Jon reminded us that 'if it's rubbish in, it's rubbish out'. The several layers involved in processing of data and translation can lead to data becoming corrupted.

Jon also noted that the language used to discuss issues related to aid has changed and is itself part of the means by which power is held by donors and intermediaries. The use of technical language may exclude stakeholders such as NGOs and beneficiaries, unless they also adopt the accepted vocabulary.

John Rowley

The next speaker was John Rowley, who used examples from his many years of work in monitoring and evaluation to highlight some of the potential pitfalls in gathering evidence on development but started with some positive examples. John described an assignment he supported in Northern Tanzania which had rehabilitated the system for water distribution, allowing women to collect water from taps rather than from a borehole at some distance from their community. As part of the M&E of the programme, women were asked to complete a daily time-use exercise involving a diagram, the results of which showed the additional time gained by the women as a result of the project. John quoted another example in which girls in Uganda were asked to take part in an exercise on how safe they felt going to school. They were invited to contribute their thoughts using a simple diagram. The resulting image provided a wealth of information. John used these examples to illustrate what he saw as effective M&E – the voices of those directly impacted by the project heard loud and clear. As John put it, 'I want these pictures to be the evidence' – a far cry from the quantitative indicators imposed by donors.

John also presented the results of an exercise with the staff of five organisations working in education in Tanzania. He asked the staff to score the indicators they used on two criteria: first, how easy the indicators are to observe and,

second, how useful and valid the results are. John presented a graph showing that the more valid observations were generally harder to collect and those that were easier to collect were usually less reliable. John questioned whether it is always true that better evidence costs more to collect.

In looking at who shapes the demand for evidence and the form that it takes, John spoke of the use of fund managers by donors to oversee M&E, a process which he saw as prioritising the interests and methods of the fund managers. He described a situation in which the management and analysis of data, including complex statistical analyses, and the content of reports were based on structures strictly imposed by the fund manager. John also considered the reliance on the services of a fund manager as the sole source of information to donors to be a cause for concern, because a failure to achieve success in the fund manager's statistical models might result in cuts to funding for projects and programmes that were otherwise working well for the people they were intended to benefit. John questioned how, as evaluators, we might shape the role of the fund manager – to what extent might we push back on unsuitable practices?

Posing a moral dilemma to the seminar audience, John asked what, as an evaluator, one should do if one was aware that a project was achieving desired results, though these results were not corroborated by the data demanded by the fund manager. This prompted much heated discussion about the 'fudging' of data by those who were willing to confess to having done it, why it might be required and how it was allowed to persist.

Claire Hutchings

Finally, Claire Hutchings drew on her experiences as Head of Programme Quality at Oxfam GB to discuss issues relating to evidence. Claire acknowledged the pitfalls and challenges that had been presented by other speakers, but noted that too little had been made of the importance and necessity of generating evidence to help us understand what we're doing right and to inform our commitments, and of the significant positive impact that efforts to strengthen our approaches to evidence generation have had.

Claire noted trends away from simple metrics towards what are often referred to as 'Hard to Measure' benefits, and how such measurement challenges can be drivers for productive change. She noted that it wasn't until her team was tasked with finding ways to measure women's empowerment that they came to realise that even within the organisation there were multiple, diverse and sometimes conflicting ideas on the meaning of empowerment and what a successful programme would look like. Working through approaches to measuring women's empowerment helped to bring out these differences and unpack hidden assumptions not only within Oxfam, but also, and importantly, with partners and with communities in ways that have helped to strengthen programming approaches.

But Claire acknowledged that the politics around evaluations is a cause for concern, with the potential to cause divisions. One particular area of conflict remains the differing weight given to different types of evidence generation, with continuing discord between those who prioritise randomised controlled trials and quantitative data, who may be concerned about the potential for bias of qualitative data or question its generalisability, and those who question how well quantitative data sets represent and relate to the people they are drawn from. Mixed methods methodologies are now an accepted part of evidence generation and we benefit vastly from employing different techniques, but we're still struggling to find the balance and must look for ways to come together at the core. Importantly, we must get better at valuing the data generated and owned by communities themselves, and recognise them as the most important agents of change in their own development.

Claire noted that studies were often driven by the wrong questions, centring on whether projects were 'delivered' rather than reflecting on any wider impact and learning objectives. Claire saw a trend shifting away from this, with donors becoming more engaged with learning outcomes and moving towards fit-for-purpose evaluation designs, but noted that there was still a temptation to overload evaluations with long lists of evaluation questions and evaluation criteria, which served to obscure the evaluation's core purpose. And this takes us back to the question of 'useful for whom?' On this point, Claire advocated the engagement of project participants and communities as evaluation users, which may require more time and more cost, but she saw it as a critical feature of well-run evidence generation. Bringing participants in to the decision on the scope and focus of an evaluation and engaging them in the analysis and in supporting our efforts to understand and validate data would allow us to look at results differently and not just through an imposed framework which may serve to skew rather than strengthen our interpretation. Claire further advocated returning to communities during the data collection process to give high-level results – a low-cost activity that helps to close the loop and gives community members the opportunity to interact with the data, a minimum courtesy that recognises the significant time people who are often very time-poor give to participating in studies.

In conclusion, Claire said that we should herald the improvements achieved in past decades whilst also recognising what has been lost. She suggested a pragmatic approach, adopting the best aspects of the large, costly surveys and studies and applying them to methodologies at a smaller scale, or 'taking a Ferrari and turning it into a well-functioning Fiat' as she put it, with a focus on fit-for-purpose study designs that put positive outcomes for beneficiaries, and their own agency, at the centre.

Finally, in striving for positive change in the aid sector, she requested greater sharing of data as the same questions are asked of the same communities too frequently. She also felt there should be more thought and care around data communication, with the use of visual presentations to make the information provided by data easier to understand.

What next?

A lively discussion followed the presentations, especially on the issue of prioritisation of accountability over learning which was noted from all four speakers, with questions as to whether this was in fact the case, how the situation had come about, and what could be done about it.

The conversation continued long after the official end of the seminar, which had clearly stirred up a lot of interest and debate amongst the audience members and speakers alike. We would like to thank everyone who joined us for this event, especially the speakers and Adam Leach, who chaired it.