

Practical responses to current monitoring and evaluation debates

Case studies

INTRAC Conference, November 2013



Understanding the Impact of Amnesty's International Members

Gemma Cartwright, Hayley Baker and Leda Avgousti. Amnesty International

SUMMARY

Amnesty International's offices around the world work to mobilise members, supporters and activists around campaigns seeking to end human rights violations. In nations where Amnesty does not have offices, people wishing to become involved are known as international members, supporters and activists. These relationships are managed and driven by the Growth Programme. Amnesty's ambition is to have over half a million international members, supporters and activists taking several actions per year by 2015. Systems to support M&E are key in ensuring that realistic targets and an engagement strategy are established. There is a need to ensure that M&E shows and supports the contribution of international members, supporters and activists to human rights change, as well as growth in numbers and quality of engagement.

METHODS

- Amnesty are working with a third party to develop a customer relationship management system (CRM) that will allow communication with international members, supporters and activists.
- Customer relationship management (CRM) is a model for managing interactions with current and future members, supporters and activists using technology to communicate with them, monitor their actions for Amnesty and also understand their behaviours etc. more so we can monitor and frame future campaigns in a way that helps us realise our goals.
- The CRM will also monitor their activity and evaluate it to ensure we are reaching our target and to assess whether approaches are working.
- Social media is being used to increase membership and engagement and we are keen for the system to support M&E of this.
- Linking campaign indicators based on a theory of change for the contribution of international members in human rights impact and each specific campaign will be used to attempt to understand the contribution of international members. This will also drive the supporter journeys developed for them also.

OUTSTANDING QUESTIONS

- We are keen to understand how we can monitor more online activity cheaply and effectively, as well as offline activities by members who hold their own activism events.
- We are also keen to know how others understand the impact of their activities around social media and how we can monitor offline activities.
- We are keen to learn from others how they have successfully monitored and given strong evidence for the impact of the contribution of members or supporters of campaigns.
- What other thoughts do others have?

REFERENCES/FURTHER INFORMATION

www.amnesty.org/join



ATT Campaign event in Guinea

CHALLENGES AND SOLUTIONS

- Many of Amnesty's international members, supporters and activists are in nations where internet connectivity is minimal due to poor infrastructure, but SMS and mobile phone communication is popular.
- Development of a CRM and ensuring consistent data is a challenge, especially when then comparing it to previous data and in parts of the Amnesty movement where data has been collected in different ways.
- It is difficult to determine immediate impact of international members taking part in campaigns as the change desired is either often not measurable or long-term.
- To create strong campaigns based around clear and measurable theories of change.
- Improved M&E within international members themselves, helping them understand the importance of feedback to us, but what's in it for them?

KEY LESSONS

- To date, we have been using the insights and data which can be generated by social media but are keen to understand how others are assessing return on investment of their online activities against their campaign objectives.
- We are aware that to truly assess this work we need to also have other mechanisms running alongside the CRM to ensure qualitative data and impact generated is recorded, and processes of engagement assessed more deeply.
- Having evaluated using supporter journeys with campaigns involving international members we have seen more integrated campaign activity which in the longer term should contribute towards increased impact.



Project Effectiveness Reviews: Oxfam GB

Claire Hutchings, Oxfam GB

SUMMARY

As part of Oxfam Great Britain's Global Performance Framework (GPF), 26 randomly sampled projects are selected each year and their effectiveness assessed, 12 using quasi-experimental evaluation designs. One of the projects selected was Nepal's 'Climate Change Adaptation and Advocacy Project'. The project attempted to strengthen the resilience of households living in seven communities across the Alital VDC area in the far-west of the country. The intervention could have been evaluated using a clustered randomised control trial. However, as this was not planned for at the design stage of the project, a quasi-experimental impact design was pursued ex-post. Intervention communities were purposively chosen, and similar comparison groups identified, and PSM and MVR were used to statistically control for any measured differences between the two.

METHODS

- A quasi-experimental impact evaluation was implemented to assess effectiveness. Evaluations are conducted by Oxfam GB Impact Evaluation Advisors with national consultants managing local teams of enumerators.
- Communities in Alital VDC, both where the project was implemented and not implemented were mapped out in order to identify valid comparison groups.
- A total of 437 households were randomly selected for interview, including 173 from the seven communities which had participated in the project, and 264 from eight communities selected as appropriate for comparison.
- At the analysis stage, the statistical tools of propensity-score matching (PSM) and multivariable regression analysis (MVR) were used to control for differences between the intervention and comparison groups.
- The effectiveness of the project in affecting 30 "resilience characteristics" was assessed through the process.
- Composite indices were developed to aggregate the data associated with the 30 characteristics, following the Alkire-Foster methods used by the Oxford Poverty and Human Development Initiative (OPHI) to measure multi-dimensional poverty.

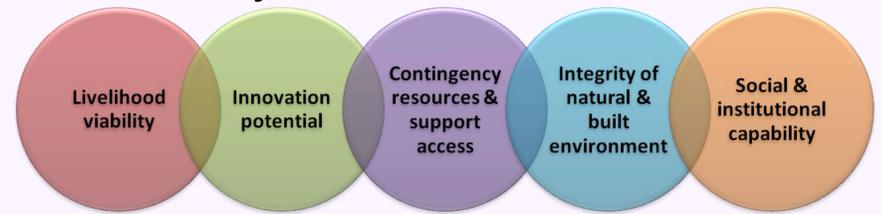
SOLUTIONS/KEY LESSONS AND WAYS FORWARD

- Explore pragmatic ways of better theoretically and empirically validating the particular set of "resilience characteristics" defined for a particular context.
- Investing more time in mapping the individual project's theory of change onto the multi-dimensional construct that has been developed to measure resilience. While we would still want to look at all those characteristics deemed relevant in that context, as there is important learning there for future programming, we want to be clearer about where the project could reasonably have expected to influence change.
- Invest for time in qualitative research ahead of the evaluation to ensure that the evaluation design is well tailored to the local context and individual projects. While this will not be statistically representative and generalisable across the population, it will allow us to better contextualise the empirical results and to acquire greater understanding of potentially complex issues. Wherever possible, undertake complementary evaluation of any governance or policy related interventions delivered by the project, to build a more comprehensive picture of the resilience of the communities we are working with. Wherever possible, undertake further research to help highlight how and why the project has positively impacted these particular results.
- Oxfam GB has only applied the approach to measuring resilience to several single different, ex-post impact evaluation studies. Its utility as a programme design and/or outcome monitoring tool has, consequently, never been "tested", and this is an area for further work.
- Encourage more investment in evaluation designs, ex-ante.

REFERENCES/FURTHER INFORMATION

- Bishop, David. (2013) Project Effectiveness Review: Climate Change Adaptation and Advocacy Project, Nepal. Oxfam GB. <http://policy-practice.oxfam.org.uk/publications/effectiveness-review-climate-change-adaptation-and-advocacy-project-nepal-303492>
- "A Multidimensional Approach to Measuring Resilience", Oxfam GB working paper—<http://policy-practice.oxfam.org.uk/publications/a-multidimensional-approach-to-measuring-resilience-302641>
- Oxfam GB. Project effectiveness reviews. <http://policy-practice.oxfam.org.uk/our-work/methods-approaches/project-effectiveness-reviews>

Key Dimensions of Resilience



Key Dimensions of Resilience assessed in the Effectiveness Reviews

© Oxfam GB

CHALLENGES

- Because 'resilience' is only ever evident after a shock has occurred, effectiveness is being measured against a complex theoretical construct that sets out characteristics of resilience. With the determinants of resilience varying by context, one of the advantages of the measurement framework is its flexibility. However, significant effort is needed to identify the specific characteristics that are relevant to the context in question. And many of the characteristics are inherently challenging to measure, several, for example, are only practically measurable with perception and/or self-reported data, which are known to be fraught with measurement error.
- Second, the project's own objectives were analogous, but not identical as the construct was developed after the project was designed.
- The effectiveness review focuses on the household as the unit of analysis, and therefore is only able to only consider the impact of the project at household or community level. It did not consider intra-household dynamics, or any impacts on the capacity, readiness etc. of duty-bearers, most notably the state.
- Where we do see impacts, or evidence of no impact, we need further information to understand why.

Digital Data Gathering in Concern Worldwide's Monitoring & Evaluation

Kai Matturi and Samuel Fox, Concern Worldwide

SUMMARY

Technology opens new data collection opportunities. In the case of M&E the direct inputting of survey responses into handheld devices can greatly reduce the time required for data processing and analysing. In 2011 Concern Worldwide decided to utilise ICT in its M&E practice. After a successful pilot project, a specific type of ICT handheld tool was chosen, Digital Data Gathering (DDG), to collect, analyse and store quantitative survey data. It is too early to make a definitive assessment as to the impact of DDG on the organisation's M&E practice. However, two years into the project it is possible to paint an emerging picture of the challenges, successes and lessons that DDG has so far unearthed.



Theresa Maurice Ojong – M&E Officer, Child Survival Programme & Albert Yanguba - Operations Research Advisor, interviewing Aminata Mansaray – programme participant in Baoma Village, Waterloo, Western Area, Sierra Leone for live data collection on the 3rd of August 2012. Photographer: Ciarán Walsh.

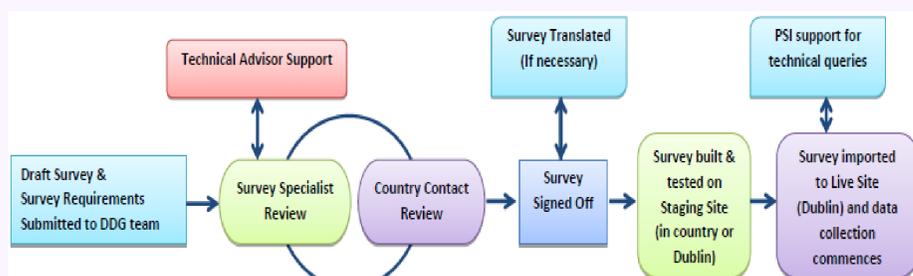
METHODS

- Software that allows users to capture diverse quantitative data sets.
- Hardware that hosts the data collection software and is able to quickly and easily download, facilitate, and upload completed surveys to a globally accessible "cloud-storage" via a web-based server.
- Data is transmitted from a mobile device to a server via a SIM card and connection to a mobile network or through an internet connection.
- GPS tracking in the mobile devices to track locations of data collection and data collectors.
- A DDG steering group comprising members of the policy, IT and overseas units provides overall guidance to the project. This group steers the work of the DDG project team. The project team is made up of a survey specialist and an IT specialist. It is tasked with rolling out the DDG implementation plan.

CHALLENGES

- There is a tendency for country management teams to perceive DDG as an IT solution to data collection rather than as a management tool to inform programme decisions.
- The roll-out of DDG is perceived by country programmes to be overly centralised and bureaucratic due to quality control systems ensuring necessary data is captured against programme frameworks and indicators.
- The current model of DDG is expensive.
- DDG does not fix a badly designed M&E system on its own.
- Removing the perception that DDG is a standalone intervention.
- Given the sensitive nature of information that is transmitted through DDG, it is critical that steps are taken to minimise security risks.

QUESTIONNAIRE DESIGN PROCESS



PROGRESS

- Since 2011, the DDG team has trained 220 staff across 11 countries.
- 18 countries of operation will be using the technology by the end of 2013, and all (25) countries by the end of 2014.

SOLUTIONS

- Include DDG costs as part of programme proposals.
- Gradually decentralise the design, building and implementation of surveys using DDG, whilst maintaining quality.
- Explore the possibility of using open software as a means to reduce costs.
- Design an IT platform that can be used as a project management tool.
- Limit data collection to what is absolutely necessary for the goals of the programme.
- Be transparent about what data are collected, how they are shared, and how they might be used in the future.
- Anonymise data collection.

SUCCESSES AND KEY LESSONS

- Over six thousand survey instances conducted since project inception. Increased quality in baseline/endline survey data for evaluative purposes.
- Reduced time spent on quantitative data inputting. Availability of real time data to manage and adapt surveys as they are conducted.
- The need to have a standardised menu of outcome indicators used across country programmes.
- The importance of having senior management at country level on board ensures that DDG is used as an information management tool.
- The need to integrate DDG into routine M&E practice from the start.
- Building the capacity of 'committed staff' so that skills are maintained within the organisation.
- As witnessed by the number of countries that have been trained on the use of DDG there is an appetite for this 'new' approach to M&E.

REFERENCES/FURTHER INFORMATION

For more information contact kai.matturi@concern.net or sam.fox@concern.net

Developing a Planning, Monitoring, Evaluation and Learning (PMEL) System for Global Witness

Jenny Ross (Consultant and INTRAC Associate) and Gavin Hayman (Global Witness)

SUMMARY

Global Witness (GW) has been developing a comprehensive PMEL system for international campaigning over the last four years. The initial approach drew on 'best practice' and focused on campaigners completing planning templates that were primarily used as a tool for accountability to senior managers. Whilst this approach did bring some 'order' and created a campaign impact 'dashboard' it was not popular with staff and of limited value to senior managers. Since 2011 GW has been working with an independent consultant to develop a more participatory PMEL system focusing on improving campaign impact that reflects better the organisational culture. The centrepiece of the approach is the annual 'Planning Week' where campaigners share their campaign logic/strategies and have the opportunity to receive feedback and learn from their peers (and senior managers). The system continues to develop and adapt to internal and external changes in the campaigning context.

METHODS

Planning

- Campaign strategy and tactics/activities templates have been developed that ensure a focus on outcomes and a clear link between activities and outcomes.
- Annual strategy/'Planning Week' involving campaign presentations and challenging of campaign logic outlined in templates.
- Campaigners have access to support from an independent consultant to complete the templates and prepare for the 'Planning Week'.
- Campaigners adapt their plans based on feedback during the 'Planning Week'.
- Participatory approach to developing planning processes and tools to promote ownership and enhance their value.
- The approach has a strong focus on outcomes with indicators and data collection linked to the fully articulated campaign strategy.

Monitoring

- Campaigns are monitored through the use of impact logs and monthly and quarterly campaign impact reports are produced.

Evaluation

- A questionnaire and follow up online survey are used to evaluate campaign 'Planning Week' and feed into the process for the following year.
- A separate project is documenting the key 'impact stories' from Global Witness's 20 years of campaigning.
- Campaigners are encouraged to conduct mini-evaluations after major activities (for example a G8 summit) in order to document learning for internal use.
- Any new initiatives/campaigns are passed through a series of questions/tests which evaluate the initial campaign logic and potential for impact. This provides an 'objective' bar for new campaigns.

Learning

- The annual 'Planning Week' provides good opportunity for sharing and learning. Sessions have focused on key targets (e.g. World Bank or European Union) or tactics (e.g. use of social media or undercover investigations).
- Campaigners are encouraged to share their experience through a series of lunchtime meetings (brown bags).

CHALLENGES

- Campaigns are dynamic and progress can be non-linear which makes planning for monitoring and evaluation challenging.
- It is difficult to be 'objective' about GW's contribution to a change or predict how change will happen.
- Ensuring that campaigners make clear links between the outcome that they want to achieve and activities.
- Often PMEL systems focus predominantly on 'upwards accountability' to managers and donors which is often at odds with the need to learn and share mistakes in order to improve impact of campaigning. Often there is an emphasis on quantitative data as opposed to qualitative information as this is seen as being more 'objective'. For example, campaign plans in some organisations will include targets for numbers of newspaper articles or numbers of parliamentarians that have been lobbied regardless of the 'value' of those activities.
- GW has staff based in the US and it was a challenge initially to involve them fully in the PMEL system, particularly the 'Planning Week'.
- Campaigners have taken ownership of the process but expectations have been raised that need to be met in terms of delivering value for them. For example, campaigners have called for greater feedback on their campaigns plans from managers and more accountability or information sharing from senior management about the organisation's overall campaigning vision.
- There is no blueprint for what makes a good campaign and often experienced campaigners do not know how to pass on their knowledge and skills. As GW has grown, individual campaigners have less personal contact with the founders and Campaign Director in which the institutional knowledge and history is embodied. Finding ways to share knowledge and build capacity in new campaigners is key.
- The key questions and tests that are used to test new initiatives are not necessarily applied to existing campaigns and agreeing to stop campaigning or engaging on an issue is still difficult (no 'exit strategy').

SOLUTIONS

- Initially, the GW system was focused on ensuring that all campaigns produced campaign plans which helped senior management to have an overview of campaigns across the organisation. But this has been balanced by a conscious effort to encourage campaigners to share what they have learnt (both successes and mistakes) and to encourage critical thinking through the PMEL process.
- The focus on learning and sharing has been supported by an investment in increased management capacity by GW which has allowed 'Planning Week' to take on more of a strategic focus as reviewing of campaign plans is now a line management responsibility.
- As an organisation run by campaigners GW understands that the 'value' of the information collected through impact logs has to be explained through an understanding of the campaign strategy or logic. GW focuses in general on 'quality' rather than 'quantity' in terms of activities. The organisation/senior managers understand that the value of one New York Times editorial is much greater than numerous articles in other media outlets. Campaigners are also encouraged to record information that is 'anecdotal' or 'gossipy' in their impact logs as in retrospect this is often the first indication of a change. This information is marked as confidential.
- Campaigners have been given the opportunity to feedback on key elements of the approach through (anonymous) online questionnaires. This feedback has been incorporated into the next year's system.
- 'Planning Week' is now a key part of the GW calendar – the majority of staff from the US office attended in 2012 and a party was held on the Wednesday night to support teambuilding and networking amongst staff.
- Some campaigners have chosen to supplement the planning approach with use of the 'critical path' method for critical analysis and are eager to see this rolled out across the organisation.
- The 'impact stories' project tries to document institutional history and knowledge and draw out what GW has learned in 20 years of campaigning.

KEY LESSONS

- A blueprint approach to developing an PMEL system will not be effective. It is important to develop a system that fits with the organisational culture.
- The system focuses on the importance of campaigners being able to explain how they think change could happen and how they will win the 'argument' (rather than focusing on quantitative indicators) – this means the system can capture what has value and meaning within the campaign logic.
- GW is an organisation that is run by campaigners (the three founder Directors are still actively involved in campaigns). This has supported development of a PMEL system that works for campaigns rather than having to adapt a system which has been developed for service delivery/programme management or broader communications strategies.
- It has been critical to bring planning, monitoring and evaluation together – good planning makes it possible to talk about impact and gather meaningful information/data to monitor and evaluate progress.
- GW has a relatively non-hierarchical culture that has supported an approach which relies on a peer-led critical appraisal of campaign strategies. Campaigners have responded well to the challenge of explaining how they think that change will happen, how they will monitor progress and how they will know if they have succeeded.
- 'Planning Week' has been a significant organisational investment but it has been at the heart of changing the PMEL system for GW. Without the focus that it provides it may have been more challenging to roll out the system as campaigners are often focused on responding and reacting to what is happening today.
- Embedding clear criteria for engagement in a campaigning issue (and disengagement) has provided an 'objective' bar for campaigners to try and reach and created rigor in the logic of a campaign from the beginning.
- The focus on monitoring and evaluating the system itself has been valuable as the approach has adapted as the organisation has changed and as new challenges and opportunities have emerged in the external context.
- The PMEL system has been very helpful for other organisational functions (fundraising, finance, communications and HR) and 'Planning Week' has provided an opportunity for support staff to get an overview of organisational plans and activities.

REFERENCES/FURTHER INFORMATION

Jenny Ross (jennyrossconsultant@gmail.com)

Measuring the impact of media programming: lessons learned from baseline data

Laura Smethurst, Zoe Fortune, Catherine Harbour, Natalie Maplethorpe, Emily Richter, David Jodrell, Emily Le Roux Routledge, Adrienne Testa, Kavita Abraham Dowsing. BBC Media Action

SUMMARY

BBC Media Action is developing TV and radio programming to help improve maternal health behaviours in Bangladesh, Ethiopia, India and South Sudan. In 2012-13 baseline surveys were conducted among key target populations in Ethiopia, India and South Sudan. Standardised survey items were used across the three countries to measure health seeking practices and the potential drivers of change: knowledge, attitude, self-efficacy, social norms and inter-personal discussion (Fishbein and Cappela, 2006). We found that some measures, such as social norms, are less amenable to standardised measurement, and we faced challenges collecting quantitative data in the post-conflict setting of South Sudan. Full findings on programming impact and on the relationship between behaviours and potential drivers will be available in 2016/17.

METHODS

- Programming impact will be assessed using cross-sectional surveys, which will be conducted at baseline, midline and endline in each country over a 2-3 year period.
- Approximately 8,000 women with an infant aged 0-9 months will be surveyed at each time point.
- Secondary target populations, husbands and older female influencers, will also be surveyed in Bangladesh and India.
- Shifts in outcomes over time will be analysed.
- Differences in outcomes between women exposed to the programming and those not will be examined.
- Dose-response analysis will be conducted, where possible, to assess the effect on outcomes of different levels of exposure to programming.
- Measures have been piloted and survey instruments have been subject to retrospective testing.
- Electronic data collection using mobile phones with Open Data Kit was used in South Sudan.
- Formative qualitative research was conducted at the start of the project to inform programming and additional qualitative research will be undertaken to understand the impact and provide contextual information at country-level. Methods include projective techniques and health facility-based research.

SOLUTIONS

- Tools and methods are being refined in response to the finding that some drivers are less amenable to standardised measurement.
- The measurement of attitude and social norms will be less standardised across the four countries. Response scales will remain standardised, where possible. Scales measuring these drivers will be built using country-specific measures, but findings will be aggregated to allow for some comparability across countries.
- Individuals' self-efficacy to perform particular health behaviours will be measured using qualitative methods to help ensure it is done effectively and to avoid significantly increasing questionnaire length.
- We are considering the value-for-money of conducting quantitative research in South Sudan and are exploring alternative qualitative evaluation methods.
- We are conducting further qualitative research across countries to help us contextualise our findings and are exploring methods to help address potential self-report bias.
- We are using cognitive testing to develop self-attribution measures to enable us to report outcome and impact to the global logframe. Given the varied nature of outputs across countries, we will not be using standardised measures, although measures will be comparable.



A woman with her children in Ethiopia. © Rachel Simpson, 2013

CHALLENGES

- The evaluation design presents some challenges around attribution of impact. It is often difficult to identify pre-defined control areas for evaluating mass media interventions. It can also be challenging to ensure that sample sizes are sufficiently powered to be able to detect realistic shifts in outcomes. We are also mostly relying upon self-reported measures of behaviour.
- We found that baseline measurements can present challenges including value for money and feeding insights into production in a timely manner.
- The majority of our measures performed well cross-culturally, although some measures, for example, social norms and attitudes, are less amenable to standardised measurement. There were also some differences between countries in respondents' understanding of key concepts being measured.
- We found it challenging to measure individuals' self-efficacy using a limited number of measures.
- Quantitative data collection in post-conflict settings such as South Sudan presents challenges for quality assuring valid measurement, even when introducing the use of mobile technology to improve data quality and efficiency.
- We face the challenge of reporting outcome and impact across four countries to a global logframe using a self-attribution method.

KEY LESSONS

- Baseline surveys provide a benchmark baseline against which shifts in outcomes and the impact of exposure on outcomes can be examined. They have been important to inform programming and create a benchmark baseline for future measurement of change.
- Some caution should be taken when directly comparing findings across countries.
- It is important to ensure that measures of drivers such as attitudes and social norms are relevant to individual country contexts. Moving away from standardised measures may not allow measures of these drivers to be directly comparable, but should still enable insights to be aggregated across countries.
- When attempting to measure several potential drivers of several different health behaviours, the effective measurement of individuals' levels of self-efficacy to perform these behaviours may result in very long questionnaires, so it may be better to seek alternative methods, for example, qualitative methods.
- The use of additional qualitative research is essential to help us better understand impact and validate quantitative measures.
- Whilst the use of mobile technology in South Sudan presents some advantages to quantitative data collection, collecting good-quality data in such settings remains challenging.

Scoring perceptions of services in the Maldives: the power of instant feedback

Valentina Barca and Nils Riemenschneider, Oxford Policy Management

SUMMARY

OPM conducted a longitudinal survey in the Maldives, to monitor and evaluate the contribution of a human development project to health and education outcomes on different islands. The research methodology included a perception scorecard, implemented as part of a survey module and also in a group-based setting. Importantly, the survey was implemented using **mobile mini laptops** and a **user-friendly software programme**, meaning that the **results were compiled and analysed while still in the field**. **Feed-back sessions** were then organised with key stakeholders and research participants before the research team left each island, enabling further discussion, validation of findings, additional insights and the option for local stakeholders to use accurate statistics for their own purposes. The effects of this approach were seen in substantially stronger local engagement, greater usefulness and deeper insights.

METHODS

- The research included a quantitative survey module (Citizen Report Cards) linked to qualitative Key Informant Interviews and Focus Group Discussions using Community Scorecards.
- Specifically, the group-based scorecard activity generated perception data that could be triangulated with the survey module while prompting a deeper evaluative discussion to justify and explain the satisfaction scores that the group has given.
- The quantitative data was collected using Ultra Mobile Portable Computers.
- A simple Survey Software (with incorporated skips and other useful design features) was used to deliver the questionnaires and enter the data directly.
- At the end of the research process in each location, the quantitative data was analysed, triangulated with the qualitative data, and transformed into simple presentations.
- The findings were then shared in separate sessions with key stakeholders in the community (including personalised recommendations) and in a community-wide feedback session.



Feeding back results and discussing findings

CHALLENGES

- The use of Ultra Mobile Portable Computers (UMPC) and tailored Survey Software is of course not free of hassle. These can be: a) costly (purchasing of hardware and software, upkeep); b) subject to crashing/viruses/bugs/power shortages (as with any electronic equipment and software, but with added problems in remote areas); c) not ideal for last minute tailoring of questionnaires (changing anything requires new contact with the developer); d) can pose a security risk (several computers carried around very poor areas attract attention); e) can intimidate respondents (interviewer perceived as 'rich' and 'different').
- Feeding back results while still in the community was problematic in one location where our research team was associated with the implementing team which had not delivered expected results: crude statistics showing how badly services were perceived created uproar among community members who had been campaigning to improve the delivery of those services and had received no attention from central government.
- The main challenge posed by the project was the fact that at the time of the third round of 'impact evaluation', very little had actually been done in terms of project activities making it very difficult to assess any 'impact'.

SOLUTIONS AND KEY LESSONS

- The fact that no project activity had happened at the time of the third round of 'evaluation' meant that the focus was shifted towards monitoring and feeding back possible reasons for failure and solutions.
- The potential problems posed by UMPCs were mitigated by a) sharing the cost of the hardware purchase across projects; b) identifying a go-to IT person in case of any problems and ensuring paper-versions of the questionnaires were also available; c) establishing a 'helpline' with the developer over the first week of survey implementation; d) applying safety protocols; e) carefully explaining to respondents why computers were being used.
- To avoid a backlash from communities at the time of the feedback it is extremely important to stress the separation between the research team and the project implementation team.
- By introducing the feedback mechanism and adding an extra day's research in each community we were suddenly in a dialogue with the community, rather than just extracting information—moving towards a more participatory process. The rapid feedback also allowed local stakeholders to interpret results and use them for their own purposes.
- Having quantitative data meant we could provide benchmarking to previous years and other islands, which was perceived as very useful, as well as providing standard analysis of perceptions by respondent type.
- An important side-product of this process was an increased robustness of the research as findings had been discussed, validated and adapted with the help of community members (usually passive respondents in the research process).
- We believe this type of participatory research can generate accurate and generalisable statistics in a timely, efficient and effective way while empowering local people in a sphere of research that has traditionally been highly extractive and externally controlled.

REFERENCES/FURTHER INFORMATION

Riemenschneider, Barca and Holland (2013). 'Scoring perceptions of services in the Maldives: Instant feedback and the power of increased local engagement', within book "Who Counts: the Power of Participatory Statistics", Practical Action Publishing, 2013

Supporting Marginalised Girls In Sierra Leone: Plan International UK's Quasi Experimental Approach to Measure Improvements in Education

Dominik Bulla, Plan UK

SUMMARY

In 2013, Plan International UK kicked off the programme 'Supporting Marginalised Girls with Improved Learning Outcomes' (GEC). The DFID-funded programme seeks to improve learning, retention and attendance for girls in 5 districts of Sierra Leone. To measure change, Plan International UK is pursuing a quasi-experimental approach that relies on large-scale surveying of both treatment and control schools during baseline, midline, and endline. In addition, a monitoring system has been put in place to collect high frequency data. The baseline was completed in August 2013. This case study reflects upon key lessons learned from the process to develop and implement the M&E framework up to the completion of the baseline. Key lessons are grouped around 4 main items: scope, targeting, measurement and validation.



Supporting Marginalised Girls with Improved Learning Outcomes

METHODS

- A quasi-experimental approach (i.e. use of control schools and communities) to determine impact of the GEC programme.
- Main data collection through a standardised household questionnaire and learning assessments.
- Use of a baseline survey to determine the pre-intervention status of programme beneficiaries as well as programme targets and milestones.
- Use of a midline and an endline to measure programme progress and overall impact against programme targets and milestones.
- Collection of monitoring data on a quarterly basis to track programme progress on certain key impact indicators that require higher frequencies (e.g. school attendance).
- Sample size calculations based on expected minimum detectable effect sizes as well as operational constraints.
- Use of qualitative data collection tools to validate findings on hard-to-access topics such as sexual reproduction and violence at school.
- Review of the programme upon completion of the baseline.

CHALLENGES

- The treatment: Do we measure the entire 'package' of treatments or do we try to disentangle individual treatment effects?
- Lack of control clusters: What to do if there are simply not enough control clusters against which one can compare programme progress?
- Choosing the control clusters: What to do if the programme design does not allow for random selection of treatment schools?
- The effect size: What to do if you cannot always expect dramatic changes?
- The tools: how much time and effort can you invest into getting rigorous data collection tools. And how much should you invest?
- Scale of data collection: How do we ensure high quality of the data collected?
- Development of the M&E framework: How much of the above can you address if requirements change at short notice?

KEY LESSONS AND WAYS FORWARD

- A rigorous impact evaluation must be programmed into a project design early on. Their requirements often guides the design of the project delivery itself. Hence, the quality of evidence may suffer significantly if the evaluation strategy is basically the product of what is feasible given a project design.
- Complex project interventions also require complex evaluation strategies. If a project consists of several interventions then it might be easier to measure the 'full package' of project treatments than the impact of each individual treatment arm. This may not allow disentangling the effects of each treatment but at least prevents the overburdening of a project evaluation.
- The decision on the evaluation strategy must also be informed by whether or not a project can randomise programme participants across treatment and control. It is not always possible to conduct natural or quasi experiments that allow impact to be gauged, without the ethical concerns that often come along with it.
- Depending on the effect size of an intervention(s), rigorous impact evaluations may require large scale data collection. This does not only have budgetary but also logistical implications.
- Depending on what needs to get measured, robust data collection tools often need to get developed. This is particularly true for intervention areas whose outcomes are less tangible such as violence at schools, or social awareness. Developing such strong data collection tools requires understanding of how best to operationalise indicators but also sufficient time for field teams to pilot the tools before roll out.
- Impact measurement also requires good enough quality of the data collected. This is particularly true for projects that collect data themselves. Spot and back checks of the administration of the robust data collection tools in the field as well as double entry of the incoming data are just some examples of quality assurance of the information gathered. The lack thereof often jeopardises the robustness of final conclusions.
- Rigorous impact evaluations are resource intense. They require careful budgeting to allow for data collection and verification to the extent necessary.
- Despite all the challenges, rigorous impact evaluations provide strong evidence not only for measuring progress but also for establishing sound project targets and milestones. They are also important sources of organisational learning on what works and what doesn't in the realm of poverty reduction.

FURTHER INFORMATION

Dominik Bulla
Monitoring, Evaluation, and Learning Officer, Plan UK
Dominik.Bulla@plan-uk.org

Establishing Baselines in Diverse Programmes: Practical Action, UK

Mansoor Ali, Practical Action, UK

SUMMARY

Practical Action has 108 contracts, operating in 9 countries and funded by more than 25 different donors for a value of more than £14 million per year. The contracts are grouped into 4 goals titled: Energy, Agriculture, Urban and Disaster Risk Reduction. The contracts' results and outcomes are agreed with donors. Over 60% of Practical Action's total income comes from these contracts. At the global level, Practical Action has a global strategy (2012-17), supported by 6 regional strategies. Establishing and using baselines at the level of each contract is possible, while doing this at the level of regional and global strategies is not as straightforward. Since early 2012 Practical Action has undertaken initiatives to establish goal level baselines and learning systems. This case study presents the practical challenges and opportunities on this.

METHODS

- Country offices were supported to develop programme documents and produce maps of stakeholders and their relationships.
- An organisational logframe was developed and a system of quarterly monitoring and annual impact reviews was established.
- Practical Action has developed Key Performance Indicators to monitor programmes against strategies and put in place global standards/principles to improve impact assessment and learning.
- An annual list of projects' impact evaluation and capacity were developed to answer organisational, programme level and donors' learning questions.
- Multiple sources were used to establish the baseline.
- Logical assumptions were separated from the facts and an acceptance was promoted that these are open to testing.
- A learning culture was promoted over accountability reporting.

SOLUTIONS

The main outcome of various initiatives was the ability of Practical Action to improve communication on their work to the external world. The organisation used the evidence to enhance communication on their impact, learning and received appreciation from their donors. The internal trust and communication was also improved and collaboration was started between different units and country offices.

A very clear monitoring, evaluation and learning framework emerged to work at all the levels. Currently, 26 impact evaluations, 2 looking back studies, 2 strategic evaluations, a PhD research, quarterly and annual self assessment are underway. All of this is also enabling the organisation to improve its management, especially on decisions to invest resources.

REFERENCES/FURTHER INFORMATION

Practical Action (2013), A Manual of Key Performance Indicators. Practical Action, UK

Practical Action (2012), Organisational Five Year Strategy (2012-17), Practical Action, UK.

An interactive exercise to understand complexity



M&E International meeting, Practical Action, 2013.

CHALLENGES

- Baselines are an important part of the intervention logic but sometimes it is not easy to accept that changes to the plans are normal.
- Frameworks and indicators are important, but they must not be too narrow to exclude the unintended results and learning.
- Words, language and concepts require careful scrutiny, as they could change a lot from one culture to another.
- Identifying, extracting and testing assumptions is important. Assumptions may be easily confused with the facts.
- Organisations are very aware about the accountabilities needed and this sometimes restricts learning.
- The developing sector deals with a diverse range of target audiences, where a vision of success may also vary.

KEY LESSONS

Practical Action is delivering a range of projects in a diverse context under a single strategy. Establishing baselines is an important part of its monitoring, evaluation and learning systems. There are 2 key lessons from the experience of Practical Action:

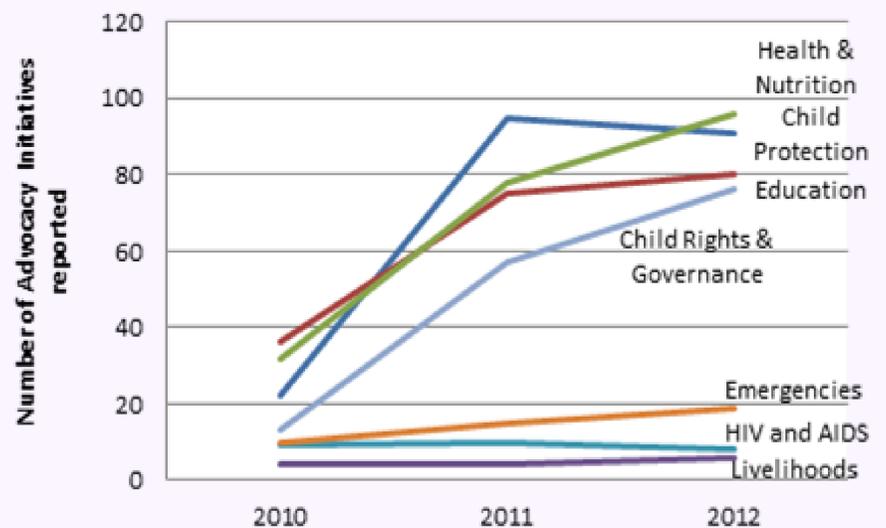
- 1) Organisational baselines must be able to capture the whole range of diverse work in its contracts. Therefore, data and information collection must develop multiple sources and methods to ensure that it covers everything it does through different units.
- 2) Organisational learning must not be put under risk due to the pressures of accountability from its diverse fundraising. Strategies need to be developed to inform all the stakeholders, especially institutional and individual donors.

Tracking advocacy efforts: Save the Children's Advocacy Monitoring Tool

Authors - Shani Winterstein, Eilidh Kennedy and Marta Arranz. Presented by Liisa Kytola

SUMMARY

In 2012, 81 Save the Children offices in 79 countries around the world reported on 380 advocacy initiatives undertaken in health, livelihoods and HIV/AIDS. Most of this was carried out at the national level, focusing on advocacy for policy change. Monitoring and evaluating advocacy work across so many countries and initiatives has proved particularly challenging. How to collate information and monitor it centrally was also a main challenge. In response, Save the Children has developed a global monitoring system that includes the Advocacy Monitoring Tool (AMT), which is a database used to collect and report information on advocacy undertaken. The AMT is used as a knowledge management and reporting tool which catalogues advocacy successes that can inform on advocacy planning at national, regional and global levels.



Number of advocacy initiatives reported in 2012, per thematic area. © Shani Winterstein

SOLUTIONS AND METHODS

The Advocacy Monitoring Tool was developed in response to the challenge of how to collate information and monitor it centrally.

- The AMT is an excel spread sheet which looks at various different pre-defined advocacy objectives.
- It is an annual measuring tool and country offices are required to report on their advocacy objectives.
- The AMT forms part of the global country annual reporting system, which includes a drafting and review process between country offices, regional offices, member organisations and thematic expert teams.
- In most cases the data is completed by the staff person leading this advocacy effort, this is often not the Monitoring and Evaluation (M&E) focal person.
- Save the Children does train the M&E staff to support the country office staff to complete the reporting template and have regional and global focal points in place that countries can reach out to if they have any questions.
- Advocacy staff are also trained on the importance of M&E in their work and how the AMT should be completed and utilised.
- Data is collected by theme and by region.
- If a country is reporting on continued advocacy efforts from the past year(s), this data is maintained in the spread sheet through a pre-population exercise, which makes reporting easier.

CHALLENGES

- Monitoring 79 country offices and collating information in a similar format across different country offices.
- Data is not gathered on all advocacy reported. Data reported is purely subjective in nature and therefore is vulnerable to dispute amongst actors and even in some cases within an office.
- Assessing Save the Children's specific contribution based on the information from the AMT can be challenging, and more work is needed on encouraging countries to document and present evidence and justification of their contribution to policy successes.
- Different levels of advocacy (campaigning, versus legislative reform etc.) are not clearly noted in the reporting which treats each advocacy effort with equal weight.
- The AMT does not provide a sense of 'good' advocacy or 'bad' advocacy.
- Specifically with the recent Save the Children transition to One Save the Children some offices have merged and this has resulted in some inconsistencies in reporting during the transition process.

KEY LESSONS

- At country office level the AMT is vital to advocacy initiatives and used as a reference for annual planning and strategy development. It is a useful barometer for measuring progress and holding those involved accountable.
- Use at Headquarter level in painting a global picture for analysis helps with institutional planning and technical assistance, the allocation of resources, direction of technical assistance and establishing links between similar initiatives across countries to scale up successes.
- Global campaigns are able to use the AMT to track advocacy messaging and activities globally to ensure maximum impact of resources and the implementation of our theory of change.
- The problem of reporting has been solved and gathering data in a uniform way has been successful. The biggest challenge is both processing and how to use this information in a systematic way .
- The tool has to be easy to use, with questions that leave little to interpretation. Training to advocates and M&E staff is important. The AMT continues to evolve, and it is important to make regular improvements based on learning from the quality of data, and the experiences of those filling it in.
- It is important to manage expectations about what tools such as the AMT can and cannot do by building understanding and the need for other methods.

REFERENCES/FURTHER INFORMATION

Shani Winterstein: shani.winterstein@savethechildren.org; Marta Arranz: marta.arranz@savethechildren.org; Liisa Kytola: l.kytola@savethechildren.org.uk

