International Conference on Evaluation, Ivàlua, Barcelona, 3 November 2011

Methodologies and approaches to evaluation in third sector organisations
By Dr Jean Ellis, Lead Consultant, Charities Evaluation Services

Charities Evaluation Services (CES) is a UK non-profit organisation which provides training, technical support and consultancy services to the voluntary and community sector on evaluation and quality. We train over 1,000 voluntary sector groups a year, support the development of evaluation practices through publications and technical support, and carry out external evaluations.

As a historian by training, I believe that a review of where we have come from can give us important learning. So in this presentation I am going to be talking about some of the trends that have emerged in third sector monitoring and evaluation over the last 20 years, how monitoring and evaluation practice has responded to policy and funding environments and priorities, and the opportunities and challenges that this has given rise to.

I shall speak above all from the experience of the UK; however the developments in evaluation in the UK third sector during this period have seen strong parallels in developments in monitoring and evaluation of social programmes more globally.

- **Early 1990s** – CES was set up in 1990 on the back of a government drive for greater efficiency and the adoption of more business like methods in the third sector.
- **Mid-1990s** – The second wave came in the mid-1990s, bringing with it more of a focus on results or the benefits of what organisations were doing, asking them not only how **efficient** they were but how **effective** they were.
- **From around 2004** there was an increased emphasis on the quality of **performance** in the sector and how third sector services were delivering **value for money.**
- This brings us to the **current situation** – In the last few years, the emphasis on **results** has strengthened with a focus on how the third sector can help to meet government social and economic targets while achieving savings in public funding, both in the present and the future; this has brought with it the language of **social value**.

Before we start we need to make sure that we have a common understanding of some of the language. I'll start by defining what I mean by monitoring and evaluation.

All organisations keep some form of records and notes and discuss what they are doing. This simple checking becomes **monitoring** when information is collected routinely and systematically against a plan.

When we **evaluate** we proactively gather data – including monitoring data if it exists. Evaluation also involves analysing and interpreting that data to arrive at judgements.

Value is a concept that is central to definitions of evaluation, although evaluation theorists differ in their emphasis. At CES we have suggested:

> *Evaluation aims to answer agreed questions and to make a judgement against specific criteria. Like other research, for a good evaluation, data must be collected and analysed systematically, and its interpretation considered carefully. Assessing 'value' – or the worth of something – and then taking action makes evaluation distinctive.*
> (CES (2009) *Practical Monitoring and Evaluation*)

First wave – early 1990s – more business-like sector and efficiency focus

The end of the 1980s had brought a sea change for the third sector in the UK. Before that period, as long as voluntary organisations did not stray too far into politics, they were not exposed to much public scrutiny. There was little demand on organisations to monitor their performance, and there was very little monitoring and evaluation support available to third sector organisations.

When evaluations were done, they were largely commissioned by funders and carried out by evaluation experts, external to the organisation. They were often viewed with suspicion, as evaluations were often used to justify funding decisions. Findings were only too often contained in lengthy reports, many of which were consigned, largely unread to filing cabinets.

However, the late1980s had seen the emergence of a new 'contract economy'. At that time many third sector organisations were losing their core funding. In order to get funding for specific projects and programmes, they needed to give a better account of themselves. As a result there was the beginning of a new positive attitude towards monitoring and evaluation from some in the sector.

By 1990, government policy laid out a framework of reporting mechanisms for greater third sector **accountability** and compliance and indicated that charities should in future behave like businesses.

Accountability, literally rendering an account of proper and beneficial expenditure of funds for the purpose intended, is one of the main purposes of evaluation. Accountability is of course important, but CES, which was newly formed and promoting monitoring and evaluation to the sector, at the same time wanted organisations to be aware of how they and their users could benefit by using monitoring and evaluation to understand their own work better and to improve it.

When we worked with organisations, we were telling them that evaluation could be good for them and their users. It could not only help them to meet external demands, but it could help them to improve and develop better services.

We supported both large and tiny organisations to carry out internal evaluations, using their own staff teams, and developing skills to self-evaluate. We saw the potential of evaluation as a reflective process, delivering learning, and being an important part of organisational management, and we trained around a simple five-step approach to monitoring and evaluation.

The simple tool that made a real difference and which we still use in our training today is the CES Planning Triangle. This simple triangle has been picked up and disseminated by support agencies across the UK and used in resources with a more global audience. What it does is to help organisations to make a distinction between what they planned to do, or are doing, and what they planned to change – or the difference they are making. This meant being clear about aims and objectives.

This is how organisations filled out the triangle.

As external evaluators, we struggled to find internal monitoring data either in place at all or of a quality that could be used. So we stressed the importance of setting up monitoring and evaluation systems when activities first start, so that baseline information – that is, the situation when the programme or project starts - could be gathered, and progress could be tracked against the baseline over time as an integrated part of work practice.

We introduced basic questionnaire, observation and interview techniques, and other simple tools and methods that could be used by project managers and frontline workers.

During this time, third sector organisations increasingly learnt to view external evaluation as more than just a necessary evil, and a new evaluation partnership emerged between external evaluators and project personnel responsible for internal monitoring. With limited budgets, one option was for consultants to work with those having internal responsibility for monitoring at the start point of their programmes, helping to put in place a sound framework and systems which would underpin and guide monitoring and evaluation activities.

Second wave – mid-1990s – focus on benefits to users and effectiveness

By the mid-1990s, monitoring and evaluation practice was moving higher up the agenda and the focus was to shift from what organisations were **doing** – their outputs, to the **results** of their programmes. This tendency increased enormously as a policy emerged of more explicit partnership between the state and the third sector. A welter of different funding regimes brought with them tighter monitoring requirements, while some funders were introducing a new or greater focus on benefits for users. This was the new language of outcomes.

From the 1990s there were increasing moves in the public sector in the US and other western economies to address outcomes – the changes resulting from work carried out – as a key part of measuring performance. These governments became focused on the real improvements intended from public services and social programmes – better health, reduced crime and anti-social behaviour, an improved environment and increased educational attainment.

In the US, some philanthropic foundations promoted outcomes approaches, and in the UK some funders followed suit.

In 2004 Judy Kusek and Ray Rist of the World Bank published *Ten Steps to a Results-Based Monitoring and Evaluation System*, which was to be translated into five languages and be used in government offices and NGOs across the world. The publication illustrated how traditional monitoring and evaluation had focused on monitoring the implementation of programmes rather than their results or benefits. They argued that attention needed to move to results – assessing outcomes by translating these into clear **indicators – signs of change** that could be measured**.**

In the UK pioneering work on developing outcomes approaches had been carried out in alcohol agencies and later on in the homelessness sector. Yet, it was hard for third sector organisations, as for public services, to shift their focus from **what they do** and establish a new focus on **what difference they make** or to make a distinction between the two.

CES' basic planning triangle remained an important first step to introduce to teams when we started work with them. With aims clear, outcomes could then be identified, and indicators could be selected. Indicators are the well-defined and measurable information that will show change is happening or has happened.

Some organisations, particularly those working in the field, say, of health or children and young people, have found it useful to use assessment tools that are used more widely in public services and have been clinically tested and validated. These are often downloadable from the internet, but real care is needed that a tool isn't chosen that is not a good fit with the work. Other organisations are developing their own assessment tools, or adapting off-the shelf tools.

For many organisations there was a problem. Government and funders are interested in community-wide measures of change, such as obesity levels, literacy rates, teenage pregnancy rates, levels of recycling and so on. The work of a third sector agency might not be of sufficient scope to show how they affected obesity level targets, for example, but so what they needed to demonstrate was how increased levels of exercise or increased numbers of play spaces could contribute to those higher level changes.

Some of the tools attempted to find a way to effectively measure what was called **distance travelled** – the progress made towards a final outcome, such as employment or maintaining stable housing. These were also called intermediate outcomes or steps along the way.

The dissemination of the Outcomes Star tool has been particularly influential. This is an outcomes tool developed by Triangle Consulting, together with various third sector networks in different service areas. This tool encourages key workers and clients to jointly assess individual progress over time. People who use this tool finds that it also develops a more client-centred way of working. There is now a family of these outcomes stars – about 11, for example for use in work with young people, with people with drug and alcohol dependency, homelessness etc.

Many community based organisations are looking at more interactive /creative methods, using video diaries, for example, or other visual evaluation methods. Organisations working in international development have had a history of using

participatory, community-based approaches and narrative to evaluate their programmes, and have struggled to fit these together with more formal monitoring approaches and performance frameworks. But, responding to the new managerial techniques, some of them created large, indicator-based monitoring and evaluation systems to apply across their world-wide programmes in an attempt to demonstrate impact in hard numbers and through aggregating results globally. In many cases these attempts have given way to a new understanding of the importance of an assessment of local contexts and circumstances, which is needed to interpret and give meaning to evaluation findings.

The third wave - from around 2004 – emphasis on performance and value for money

By the new millennium, the third sector in the UK was more and more accepted as a partner in delivering public services and the drive to **improve performance** was high on the agenda. Third sector organisations felt an increasing need to develop their monitoring and evaluation in order to compete in the market place, and to demonstrate value for money and improved benefits for beneficiaries.

However, monitoring and reporting was beginning to become a burden. Organisations were sometimes running a number of complex programmes, and found themselves accountable to a range of different regulatory bodies and funders.

In many cases organisations were not even using the data they were generating internally, in part because so much time was spent in collecting and reporting it. When we carried out a major research study in 2007, organisations told us that they were 'monitoring themselves to death.'

The UK third sector was not alone. In the US, in 2008 Joanne Carman wrote about non-profits 'drowning in data'.[1] Across Europe there were similar concerns. A 2007 publication with contributions from 12 authors from 7 different countries drew similar conclusions in relation to public services.[2] Several of the contributions talked about systems being awash with performance information, which hampered rather than fostered effectiveness and improvement.

Our 2007 research found that organisations were often struggling to bring their data together and make sense of it. They were collecting it, but couldn't always analyse it, or use it.

Evaluation was no longer just carried out by evaluation experts. There was a real challenge to train project managers and front line workers in the necessary skills. The good news was that there were new opportunities in the development of monitoring and evaluation IT software, and what we found in our research was that organisations that were using it were making considerable time savings and increasing the quality of their reporting. Yet the third sector has been slow to take up the opportunities offered by this new technology.

---

[1] Carmen, J (2008) 'Non Profits, funders and evaluation: accountability in action'.
[2] Marie-Louise Bemelmans-Videc, Jeremy Lonsdale and Burt Perrin (2007) *Making Accountability Work: Dilemmas for evaluation and for audit.*

The current wave – emphasis on savings and social value

As we reached the end of the decade, the 2000s, the percentage of public expenditure channelled through the third sector in the UK looked set to continue to grow in a major way.

Public and third sector services were now increasingly subject to the dominance of a market-based ethos, pressure of tighter budgets and the need to do more for less. This brought with it a new emphasis on demonstrating economic measures of value, and increasingly the sector looked at the private sector for inspiration.

In the past, evaluators have questioned how appropriate it was to assess social programmes using economic indicators and analysis. However, the UK government was now enthusiastic about a methodology that might demonstrate current and future benefits in terms of savings to the public purse while bringing measurable social benefits.

The Social Return on Investment (SROI) methodology, developed first in the US, puts a monetary value on the social and environmental benefits of an organisation relative to a given amount of investment. This leads to the calculation of a ratio of total benefits to total investments. For example, an organisation might have a ratio of £4 of social value created for every £1 spent on its activities.

The UK government invested in a three-year project to turn a complex methodology into one that was more accessible to third sector organisations. The method is now welcomed with enthusiasm by some funders and organisations themselves – who have seen an opportunity for using results expressed in monetary terms to advantage when competing for contracts.

This new method seems here to stay for now – responding as it does to the political and economic climate. But a number of pieces of research have expressed caution – particularly about using the SROI findings to make comparisons. Not only are certain assumptions and judgements involved in the method, but it's difficult to apply to some sorts of work, such as campaigning, and some kinds of outcomes, particularly personal benefits.

As well as this current focus on demonstrating social and economic value, there is also is also some pull towards asking the third sector to use more 'scientific' methods – more commonly understood in clinical experiments. These approaches – for example using control groups – have ong been regarded as the gold standard in public policy evaluation, but they are costly for third sector evaluation and fraught with difficulty.

As well as being more cosly, these approaches need greater technical skills than are available to internal evaluators. Interestingly, they run counter to the trend for monitoring and evaluation to be an essential element of everyday work practice. Of wider concern is whether the prime focus on demonstrating value through numerical formulae and scientific proof might shift emphasis away from learning lessons through monitoring and evaluation.

The current focus on outcomes has shifted attention away from evaluation of how things are done – the processes. We mustn't forget the importance of understanding the context in which our programmes are implemented and how these affect how we can benefit users.

Often simply asking 'What works' may be misleading. We need to make sure that we ask questions that require another level of response, such as 'What parts worked?' 'Who did they work for?' ' In what contexts did they work?' 'What else happened?' 'What were the broader and longer-term consequences.'

However, the demands of learning and of accountability tend to pull organisations in different directions.

In demonstrating value, organisations are encouraged to demonstrate their success, and that they are the ones that have achieved that change. By contrast, evaluation for learning requires a much more open enquiry about what change has or has not occurred. It means looking for information about failure as well as success and trying to understand what has caused or prevented change – also for which groups of people or in what situation change has occurred.

One method which is gaining currency is the theory of change, described by Connell and Kubisch back in 1998. A theory of change model helps to identify the theory which underlies any social programme about what will make it work. It shows the links between activities and short-term change, steps along the way and longer-term outcomes. It also makes clear all the assumptions that are made, for example about what will cause change.

A theory of change has a narrative that can help organisations show how the results they achieve – sometimes from a limited scope or remit – contribute to the bigger picture of social change. The model can also help to describe how a specific project may be part of a wider and more complex pattern of intervention – and how the project **contributes to change** rather than operates as its sole cause.

A theory of change is a flexible tool that can be used in conjunction with a diverse range of different monitoring and evaluation approaches, including those putting a monetary value on benefits.

A diverse monitoring and evaluation landscape

Third sector organisations that are starting in 2011 to engage for the first time with monitoring and evaluation face a dynamic and exciting, if somewhat confusing, landscape. The internet now provides a wealth of support in terms of literature, methodologies and tools. In the UK at least, there are also numerous support bodies, departments and consultancies that can add value in different ways to internal monitoring and evaluation efforts.

Ray Rist, one of the World Bank authors that had promoted results-based reporting in 2004, has reflected on the change from 20 years ago. At that time, monitoring remained undeveloped and organisations had only modest databases. Evaluations

were driven by funders and usually carried out by external evaluators. They were often characterised by long, written reports which few inside the organisation and even fewer outside, would read.

Evaluation is now being more and more carried out by non-evaluators, and the old distinction between monitoring and evaluation is breaking down. We can send out online surveys to large samples and retrieve instant response analysis. The medium for evaluation is changing, as video diaries, website statistics, blogs and social media add to the sources of evaluation data. A local group, such as a sports club funded as part of a programme, now has the potential to input data on its activities on a day-to-day basis to a web-based monitoring and evaluation system which can provide real-time data to programme managers and funders. Our monitoring and evaluation findings can now be stored on the internet and accessed by organisations working in our field across the world. Individual frontline workers no longer need to rely on experts to turn their data into smart charts and visual representations. Our users can write their own personal stories and put them on our websites.

However, the challenge remains of finding meaning in our work. Organisations still need to plan clearly, identifying their aims and objectives. All these different methodologies rely on clarity about outcomes, the changes that our work brings – the difference that we make. And it is this change from 20 years ago, being not only accountable for what we do and how we do it, but on our results – the difference we make – that continues to be a challenge, but an important one.

We also need to find effective ways of managing the data we collect. And we need to look behind the data to interpret it and understand what it means. Above all it means fostering a culture that allows us to take lessons from our evaluation and to increase the benefits of what we are doing. This will lead to a more creative third sector, where evaluation can foster improvement and innovation. This year at CES we held the first Learning and Innovation prize for third sector organisations, to encourage good practice in monitoring and evaluation.

The Leeds Survivor Led Crisis Service, the winner in the small charity category, said that a 'radical change to services as a result of visitor feedback was the beginning of our monitoring and evaluation journey.'

I will leave you with another quote from them, which I believe illustrates how far monitoring and evaluation in the third sector has travelled:

> *We are passionate about monitoring and evaluation as we believe it is the absolute right of our visitors and callers to know that we scrutinise our work in order to ensure that we are providing an excellent service which meets their needs.*

Dr Jean Ellis
Charities Evaluation Services
1 November 2011