‘Measuring and Evaluating Outcomes in Practice’
NPC Conference 23.5.2013

These notes were made by Catherine Sutton who attended this conference on behalf of the Impact Assessment Project for Arts Inform. They are a personal response and not an official record of the event.

Overall notes

1. Very striking that speakers are those who have started out as ‘field workers’ in their specialist area (e.g. a health visitor, a youth worker) and have been drawn to evaluation though the need to understand the context and impact of their work better and have then become expert in evaluation. However, nobody was the same. There was a lot of difference in terminology. A sense that this was quite an emerging area. A lot of head scratching still going on. A sense of contextual challenge (financial, political) but also a sense of excitement and optimism that we are getting better at doing this and that it will = quality in our work.

2. The words ‘evaluation’ and ‘impact assessment’ were used fairly interchangeable. Many people referred to M&E (Monitoring and Evaluation) as their area (not impact). Rather part of the job of evaluation was to show impact

David Pritchard (Head of Measurement and Evaluation, NPC)

- Medium sized charities rarely asked for evidence of impact – so what is the need for what we do

- There is increasing appreciation of outcomes and accurate measurement. Payment by Results is increasing – though it’s problematic

- There are new evidence-based initiatives – e.g. Project Oracle http://www.project-oracle.com/ from the GLA

- What should we be asking of Government in terms of access to evidence and data (in order to inform our understanding of need and also to provide comparison for control purposes

- What should we be asking of academic researchers in terms of making their evidence accessible (physically) and accessible (readable!) and publishers to make it affordable

Charlotte Weinburg (Exec Director, Safe Ground)

Safe Ground works with prisoners and their families through arts based programmes

- They have built an evidence-base over the last 9 years, comprising 1 pieces of work. Its purpose is to show the impact of their methodology

- This is external evaluation and research

- Safe Ground has a staff of 7 and is small - £ 350K turnover. They are always asked for evidence. But they find that the more money people invest, the less they ask

- They have a Business Manager, Adam Mole, who used to look after evaluation and impact. The value of evaluation to the charity was using the evidence to support strategic planning and this is why Adam Mole was promoted in this way. INTERESTING

- The Truth About Evaluation:
  - It's complicated
  - It's expensive (but they build evaluation into their full cost recovery model)
  - It's frightening - because you may find out things you'd rather not. It may not be possible to show your impact – or the evaluation may show negative impact.

- The Ministry of Justice (MoJ) have recently opened up their data. This is a major step as they can now look evidence for e.g. reconvictions. Previously it was not possible to track people so couldn't get beyond intermediate measures.

- Evaluation can throw up surprises. Bear this in mind when asking your questions. Example given of working on a TV soap opera in Nicaragua. In order to build new character of a teenage girl a questionnaire was constructed by the writers to be completed by a sample of teenage girls. One question was 'Who do you talk to about your problems' Suggested answers to tick were 'My best friend', 'My Mum' – everyone expecting answer was best friend. Charlotte W pushed to have 'No-one' as a response. This came out top and supporting info said 'I want to talk to my mum but I can't'. So character rewritten, new character of mum added into script...

- Look for hidden indicators as to what your investment / success looks like. Example given of a prison project – v tough inmates – outcome was them giving a presentation to fellow inmates. On morning of presentation Charlotte came in and they were voluntarily mopping the floor. This demonstration of pride in their work etc was a great an indicator as anything else.

- What impact does impact have on policy? Difficult – often policy decided outside of evidence. But evidence can point up links between areas. Organisations should collaborate and share – e.g. Prisoners Learning Alliance. By pooling collective knowledge and evidence the case to Govt is stronger.

- Always use negative feedback. It will tell you something. Really dig into it. It may contribute to your future quality assurance.

- Ask the right questions. Mix different kinds of evidence.

Dr Elizabeth Young, Director of Research and Policy, Home-Start UK (a charity working with families)

This was a complex, quite technical talk. I need to look at the online presentation again but here are a few nuggets

- Do theories of change and logic models lead to measuring the right outcomes? In truth they are only as good as:
    - The metrics available
    - Interpretation of results
    - Attribution of value in a complex system (there may be other inputs as well as yours that are responsible for the impact)

- Beware of oversimplification – you can lose treasures in the data. E.g. when studying how families have changed – you have to look at the individual families as well as the group results in order to see the full impact

- Think about having a Single Outcome Measure – an overriding outcome regardless of what the needs are on an individual basis – e.g. Families are coping better

- Use the counter-factual. Think about what the world would be like without your organisation in it. This will help to pinpoint what you do and what the different outcomes are.

- A scale of credibility – where are you on this

Credibility ⇨					
Anecdotal Evidence / Quotes	Case Studies	Self-reported change	Before and After Survey	Control Groups	Randomised Control Trial
Basic				Advanced	

- Planning for research: need to organise / understand Budget, Stakeholders, Timing, Participants, Tools

- A note on sampling
    - Quantitative data. Need numbers from subgroups that are proportionate to the total %. E.g. if 50% of participants from Surrey, 30% from Middlesex and 20% from Bucks then the numbers sampled should reflect this.
    - Qualitative data. You just need to represent the full range of views. % representation not necessary.

- Attribution
  You can ask people you are working with to what extent the change they have experienced is attributable to your activity. This is OK but will always be subjective.
  A comparison control group is the only sure test.

  Ideas for control groups:
    - Use waiting lists – these people will be comparable in terms of need but will not have had the activity
    - Compare results of impact of the major activity with 'small dose'. E.g. what is the difference in impact between a one off telephone counselling session and the large intervention of 1:1 long-term counselling

- Disincentives for honest feedback

e.g. a family may not want to give positive feedback about impact because they think that they will be discharged from the service. A family may not want to give negative feedback about impact because they are afraid their case will be escalated

- Homestart has a wide monitoring programme and data set to then customise for funders. But they would never create new monitoring for just one funder. 'If you talk to funders about what you can provide then this generally works and indeed helps the funder and you to better understand each other'

Kieron Kirkland, Development and Research Manager, Nominet Trust

Kieron was programmed to talk about Hard and Soft Outcomes, but he didn't. The resulting talk was complex, exciting.

- Financial value is inherently subjective.
- Social value is even more so.
- So in any model of change different stakeholders to whom you are accountable – e.g. Board, funders, beneficiaries – are bringing different value systems.
- You have to think about zooming in an out of your change model – different people will have different values, informing their theory of change.
- The evaluation duck:

This represents two things in evaluation.

1. Above the line we see the duck moving in the water – we can see / measure the distance travelled. This is an external articulation of success
2. But we also need an understanding of the operating context. Look below the water. How is the duck moving? What is going on – are there obstructions?

You need to show the relationship /interaction between what you do / achieve and the operating context.

- How well do you know your operating environment?
- If evaluation doesn't include wider hypothesis testing then you are only testing what you already know.
- Kieron is very keen on 'open data capture' with participants. Through storytelling, video, image. Asking them about the most significant change on a project. Ask completely open questions in order to get rich data. Then interrogate / categorise this – look for the patterns.

He also had a system for getting people to tell their 'story' then self-code.

- Biased data is still useful. But understand the bias and therefore use this data differently.

- Open data can be interrogated continuously. E.g. used for different funders, or indeed different types of funders, or across outcomes recognised by different sectors.

- There is more honesty in open feedback – there can be disincentives to answer accurately.

- Data capture – analysis / aggregation – translation – communication.

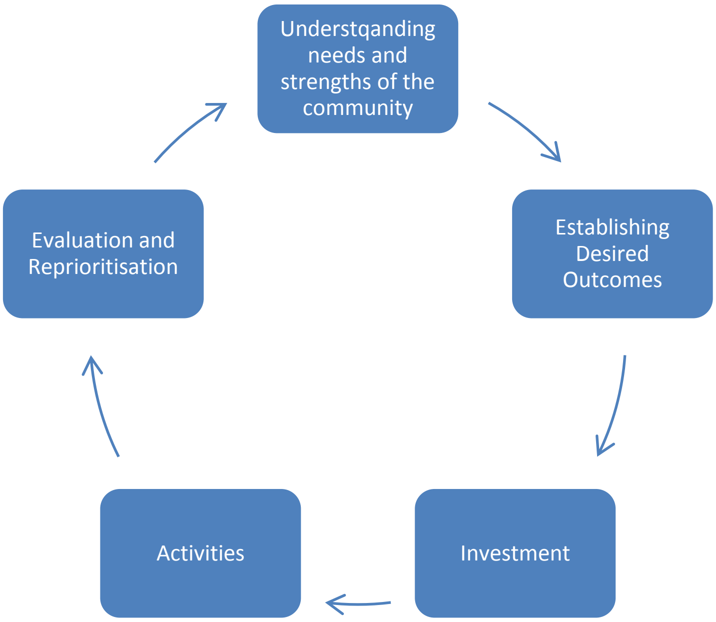
- Example of this system in use – Global Giving Project – multi-org piece of work, supported by Sensemaker Technology. This is on a huge scale so use of IT to do categorisation is necessary. But I can't help feeling this is a very similar approach to our qualitative one and our 2 methodologies. See what you think: http://www.globalgiving.org/story-tools/

Sophia Looney – Director Policy, Equalities and Performance, London Borough of Lambeth

This was interesting as it was from a local authority perspective.

- Talked about 'Cooperative Commissioning'. This means involving residents as shapers of commissioned services, not simply as recipients. 'Doing with' not 'doing to'

- Model of process below. Citizens lie at the centre



- NESTA Standards of Evidence
  http://www.nesta.org.uk/investment/impact_investments/assets/features/standards_of_evidence_for_impact_investing
  Lambeth is only achieving evidence of standard grade 1, occasionally 2.

- In its role as commissioner Lambeth is shifting its emphasis from outputs to outcomes.

- They are keen to look at outcome relationships – e.g. crime reduces / people have better homes.

- The best way to be successful with local authority relationships is through research and evaluation showing outcomes.

- They practice risk awareness, not risk avoidance. But with Payment by Results the relationship with risk will change. Better outcomes knowledge and research helps to manage risk and make funders feel more confident.

Social Investment

There were two speakers focusing on this. Their view is that social investors are still very finance-based and not comfortable with language around social outcomes. However, there is also a need to take the 'social impact people into a financial space.'

Question to funder panel

Will there be more need in the future for evaluation services?

- Local Auth – yes, quality is needed and attribution
- Trust – yes – we will be more outcomes focused
- Social Investment – Yes

Intermediaries (evaluation services) needed to support small organisations in finding a pathway

Breakout Group – One-off services

I didn't go to the Ongoing Activity, group partly because it was an arts organisation and I was interested in something different, and also because what was meant by ongoing and one-off was not particularly clear.

The organisation leading the session was SolarAid – one I would not have come across so it was interesting. Put simply, they are a social enterprise selling solar lights to communities, largely in Africa, where there is either no lighting or kerosene is burnt. The lights are cheap and as they use sustainable energy they are in a short time much cheaper than kerosene

- SolarAid previously asked leading questions, rather than open ones. The felt that even asking a question such as 'what was the outcome of this for you' was leading as it presupposed an outcome.

- Previously evaluation was for funders. Now it fulfils an internal function – to learn

- As a result the quality of messaging has improved as has internal quality. Everyone is clear about what the outcomes of the work are.

- Better to have fewer indicators and do them well. Before they had been tempted to over-claim.

- Open questioning. The Company believed that the impact of their lights was: economic, environmental (less kerosene burnt), wellbeing / health (kerosene not good for health) and education (children able to study longer with lights). However by asking open questions and also doing other monitoring such as giving people cameras to take pictures of the lights and asking children to draw pictures they began to understand that the recipients of the lights had completely different perspective on impact. It was all about wellbeing – pride at owning a piece of technology, family relationships – e.g. cooking and eating together, socialising (the lights meant that people stayed up)

See separate doc for write up of 'Evidencing Impact through Economic Analysis'.

Diana Barron, CEO, Coordinated Action Against Domestic Abuse – spoke about sharing measurement within a partnership

- How do we do measurement in such a way that we are embedding a learning culture in our organisation?

- Partnerships – agreeing metrics up front will strengthen a partnership – or else point up problems straight away.

- It's brave to say that something is not working – so that we can learn.

- How can we use the same information all the way along the process – i.e. for partners, funders, advocacy and policy makers?