CDI Centre for Development Impact
PRACTICE PAPER

Innovation and learning in impact evaluation

Assessing Impact in Dynamic and Complex Environments: Systemic Action Research and Participatory Systemic Inquiry

Abstract This CDI Practice Paper by Danny Burns explores the implications of complexity and systems thinking for understanding how change happens. This in turn has significant implications for impact assessment. The paper shows how a Systemic Action Research methodology can be employed for assessing impact in complex and highly dynamic environments.

1 Introduction

This paper is about the uses of Systemic Action Research (SAR) and Participatory Systemic Inquiry (PSI) for impact assessment (Burns 2006, 2007, 2011, 2012, 2013; Wadsworth 2001, 2010). SAR is an action research methodology which embeds design, planning, action and evaluation into a single process. It is an iterative learning process which supports real-time assessment across social systems. SAR can be embedded within programmes or layered into programmes later on in their development (Burns 2007). PSI is a shorter process, which allows a system to be mapped as a baseline against which changes in the dynamics of the system can be assessed (Burns 2012). PSI can underpin an action research process or it can be carried out as a process in its own right. SAR typically takes place over a period of 18 months to three years. A PSI mapping and analysis might take place over a 2–12-week period. SAR has been used in a number of national and global INGO programmes.[1] It has also been used in large-scale evaluations such as that of the Welsh Assembly Government's Communities First programme.

Generally, when we refer to systems we refer to complex webs of relationships between people, processes and the environment that they are situated within. The relationships include communication relationships, emotional relationships and power relationships. These are characterised by complex feedback loops, dynamics, thresholds and tipping points (see page 6). When we say that an intervention makes a contribution it is usually the result of interaction with other actions, interventions and contextual factors within a social and economic system. This means that it is critical to understand the relationships between different actions within a system and *how* they change each other. The focus of attention in a systemic inquiry is thus on the way in which the relationships between factors contribute to change.

Within any system, this complex web of relationships can form one or more equilibria. Over time these can become stable. As patterns of interaction are repeated, pathways are formed which are often made visible as norms. These repeated dynamic patterns can be described as system dynamics. They shape the extent to which individual agency is both possible and sustainable. In complexity terms we can see these patterns as representing a dominant 'attractor', which like a magnet draws people into established patterns of movement and behaviour. It is argued in this paper that changing system dynamics is critical to sustainable impact. This can be done by creating and mobilising around alternative attractors. This requires methods which make visible system dynamics; which can assess the extent to which they have changed; and which explains why they have changed.

Impact assessment requires an understanding of causality which in turn requires an understanding of how change

happens and why one (or many) thing(s) leads to another. Befani (2012) stresses that there are many ways to look at causality and that in situations characterised by the combination of multiple interacting causes we need to move beyond the limitations of counterfactual-based approaches. SAR approaches which subject detailed change narratives to critical and collective scrutiny provide a means for tracking the processes by which change in these contexts happens.

## 2 Action research and Systemic Action Research

Action research is a process where participants work through repeated cycles of evaluation, assessment, action planning and action (Reason and Bradbury 2008). This is typically articulated as a four-stage cycle. The Kolb cycle is most commonly quoted (Kolb 1984) but there are many versions. When the cycle is completed it starts again – the new situation is analysed; hypotheses are generated, refined or reformulated; new action is planned and so on. I have found it helpful to explicitly articulate a fifth stage in the cycle specifying the generation of theories of change (see Figure 1). Rogers (2008: 34) highlights the problem that 'simple logic models are… more likely to present a single theory of change, rather than representing different stakeholders' views about what are desirable outcomes and how these might be achieved.' This thinking is at the heart of the action research process where it is important to generate dynamic theories of change with each iteration of the process, and to reflect the many different theories of change articulated by people with different positionalities in any complex social system.

Cycles of this type are now in fairly common use. What makes action research different is the frequency of the cycles – typically no more than two months apart. This allows for a rapid iterative and real-time learning process where action is informed by analysis and analysis is informed by action. Action research is constructed around regular meetings (inquiries) of action researchers/key stakeholders interspersed by actions, activities and interventions. Action research processes are dialogic processes. This does not mean that all of the data comes from dialogue (although some of it does), rather that the process of analysis is dialogic. Impacts are assessed by the stakeholders (sometimes alongside outsiders) based on the data that they gather. It is thus an iterative process within which there is a continuous re-assessment of inputs, outputs, outcomes and impacts.

While elements of systems thinking are built into the many different approaches to action research, SAR differs in a number of ways.

Firstly, it explicitly focuses on system change, and brings in specific tools to surface and engage with system dynamics.

Figure 1 **Building theories of change into the action research cycle**

Evaluate impact → Analyse situation → Generate new Theories of Change → Plan action → Take action
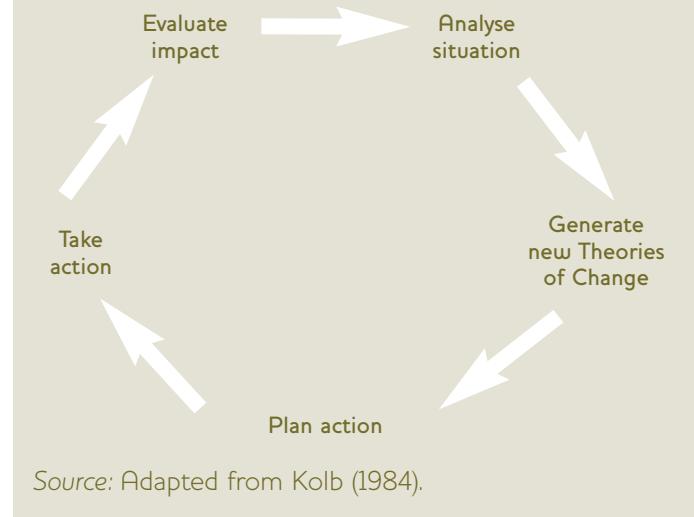
Source: Adapted from Kolb (1984).

One such tool is Participatory Systemic Inquiry, which has already been discussed (above). Here multiple stakeholders co-create big 'messy' maps of relationships. Sections of these are then distilled into smaller more systematic maps which show causal relationships, feedback loops, etc.

Secondly, it builds a learning architecture, which mirrors the complexity that it is engaging with. So instead of one core group and one core strand of inquiry, a SAR process may be comprised of many different inquiries operating in parallel involving different stakeholders working within different parts of a system. The learning architecture links the different learning strands to each other, and to formal programme structures; it also creates spaces for deliberation and sense-making up the hierarchy and into policymaking arenas. Sometimes these 'strands' are developed separately because conflicts of interest or even open conflict may make it difficult for people to be in dialogue directly with each other.

In this way, SAR can be seen as a meta methodology, which can contain many methods. In SAR processes that I have been involved with, methods include generation and processing of systems maps; rich pictures; monitoring of workload patterns; surveys; immersions, observation and transect walks; participatory photography and participatory video; open discussions and dialogues; story collection and analysis.

3 How change happens

SAR is conceptually built on a combination of systems thinking (Checkland and Scholes 2004) and complexity theory. The understanding of change, which is inherent in these concepts, has only recently started to be taken seriously within development (Ramalingam 2013) and evaluation discourses (Williams and Imam 2006; Patton 2011;

Rogers 2008; Stern *et al.* 2012; Woolcock 2013). Complexity theory represents a challenge to the linear assumptions about how change happens that underpins almost all funded development programmes. These tend to be based on a predictive logic, i.e. if a particular set of actions are taken these will result in a particular set of outcomes. This works very well where there are known solutions to problems, and stable and replicable contexts, but is often problematic in real-world social contexts which are constantly changing and in which multiple actions are combining with each other time and again to produce a highly unpredictable environment. This paper argues that in these contexts interventions that are able to adapt through real-time learning will be most successful, and the answers we are looking for should read something like this: 'when we did this, this is what happened, and this is how and why it happened', because it is only when we know *how* and *why* that we are in a position to adapt our knowledge to different environments.

The underpinning assumption of this paper is that much change in contemporary development situations is not linear (Rogers 2008), and change which is more predictable is often governed by complex feedback loops, which has significant implications for impact evaluation. While it is not practical here to look at all aspects of complexity and systems, this paper looks at four which have particular implications for impact assessment:

- Change is emergent.
- Interventions often produce unpredictable and unintended consequences.
- Change is constrained by system dynamics *and* sustainable change requires changes in system dynamics.
- Latent change often leads to tipping points characterised by sudden major transformation.

Change is emergent

Change is constant and dynamic and unpredictable. Any action within a system can lead to system change because all parts of a system are connected. Any move within a system both changes the system and changes what participants see of the system. What they see of the system determines where they perceive that they can go and what they can do within the system. This means that after each step in the change process a different set of options or pathways are available than were there (or perceived) before the action was taken. Similarly, new constraints are constantly emerging. Navigating this change successfully requires a higher degree of programming flexibility than is typical in current programme designs. In dynamic contexts, programmes need to be reassessed and re-oriented at regular intervals.

An example of this is provided by Figure 2. The programme objective may be to get from A to 1. The

Figure 2 An iterative approach to decision-making
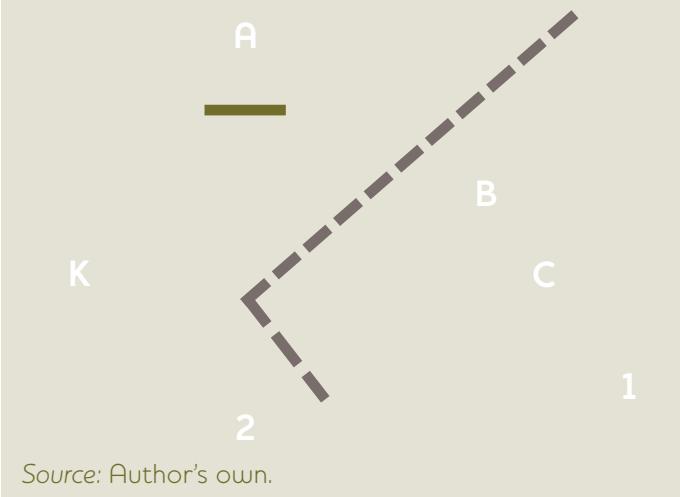
Source: Author's own.

theory of change is that A will lead to B will lead to C will lead to 1. Unfortunately the programme planners did not know that there would be a flood and a huge un-crossable river (grey broken line) is now flowing between A and B.

So while many programme managers would feel tied to B (and keep looking for new ways to get over – or under – the river) because they have to follow the theory of change that has been pre-constructed, a flexible programme manager might decide to go to K. She does not know the route from K to 1 but it looks more promising, so she heads toward K. When she gets to K she is able to see 2 which had previously been obscured. Now she has more options. If she still wants to go to 1 then K turns out to be a better route, but having moved past the wall (green line) which obscured 2 she may discover that 2 is a better place to go.

This is a very simple example. Here it is only the environment that has changed. But it is also possible that A, B and C are moving, and that I (the subject) am changing and many other changes are happening at the same time (which is why we can often only see the best opportunities while we are acting rather than before). What I am illustrating here is that changing the way in which we think about planning fundamentally changes the way in which we need to think about impact.

As part of some work that I did with the non-governmental organisation (NGO) SNV Kenya on the Lake Victoria Water and Sanitation Programme, we facilitated inquiries in four small towns in Uganda. One of these took place in Nyendo where we were exploring the problem of toilet provision in the market. There were toilets in each of the four corners of the market, all managed by the local council. Problems had arisen: men were not paying for

using the toilets, making them uneconomic to run; a number of toilet attendants had been beaten up; and where there were problems with the water supply, it always seemed to be the women's toilets that got shut down. It emerged during our inquiry process that the lorry drivers and the fishermen wanted to run their own toilets. Interestingly, this opened up the possibility of a solution to the problem for women, which did not appear to be there before. This new solution challenged two core assumptions: (a) that toilets could not be run by anyone other than the council, and (b) that each toilet block had to be multi-sex. So if we take this as an example of adaptive programming, we could no longer assess the effectiveness of the toilets in the market on the basis of the old programming logic. A new baseline has to be constructed, the theory of change has to change, and the data has to change. So in short, adaptive programming requires adaptive impact assessment.

This resonates closely with contemporary thinking on impact assessment. As Rogers (2008: 39) points out: 'Emergent outcomes may well require an emergent logic model – or in fact one that is expected to continue to evolve.' Patton has called this 'developmental evaluation', arguing:

> Developmental programming calls for developmental evaluation in which the evaluator becomes part of a design team, helping to monitor what's happening, both processes and outcomes, in an evolving, rapidly changing environment of constant feedback and change (Patton 1994: 313).

Implications for impact assessment In this context impact can only be assessed against the decisions made during the course of the programme, not against the programming logic at the start of the programme. This means that process and content (including any changes in direction, and the rationale for them) must be meticulously documented as we go along.

If we are using baselines to assess impact then we need a process for generating rapid and real-time baselines, as well as the outcome data which relates to them. If, as in the example above, we decide to go to 2 we need to establish a new baseline at K because we have seen the possibility of achieving something different and we are assessing a different impact.

We can also conclude from this that impact cannot sensibly be assessed in relation to a macro theory of change, because in a highly complex and dynamic environment there is so much change within the system it is impossible to predict the outcomes of most actions. Theories of change need to be generated iteratively at every deliberative stage of the process, and this is what the learning architecture of SAR enables.

Interventions often produce unpredictable and unintended consequences

Changes in one part of the system have an impact on other parts of the system. These can be minor changes or can be much more significant where, for example, you get domino effects running through a system. As a result of the complex and dynamic environments that interventions are implemented in, programmes can have unpredictable consequences. These can occur within the boundaries of what is considered to be the domain of the intervention, or beyond those boundaries. For example, some interventions might be successful, but only because they have sucked away all of the skilled workers into that project, at the expense of a whole set of other projects which now start to decline. Charities Evaluation Services cite an example of negative unintended consequences on their website:

> Installing CCTV cameras reduces the level of drug crime and street prostitution in the Borough of Lowdown. Unfortunately, levels of drug crime and street prostitution increase in neighbouring boroughs that do not have CCTV.[2]

This pattern of displacing negative activity is common. Once again, the intervention can be evaluated as successful in its own terms but only if the displacement of negative impacts is not taken into account. Another common unintended consequence of development activity is dependency. There are many examples of refugee camps that people do not want to move from because conditions back home are much worse than in the camp, or of development impacts which only last while there is ongoing intervention with ongoing inputs of resources. There are also numerous examples of loans which have sent people spiralling further into poverty because they could not manage the debt, and so on.

In another piece of work I did with SNV Kenya, we were exploring the impact of a clinic supplying antiretroviral drugs to treat HIV/AIDS. While availability of the drugs provided a direct response to HIV/AIDS, the clinic started to put intolerable strain on housing and facilities in the surrounding neighbourhood. This was due to the substantial increase of the local population as a result of people needing to live close to where they received their daily dose of the drugs. For most people living in poverty it was not viable in time or monetary terms to travel to the clinic (Burns 2007). The beneficial impact of the intervention had a negative impact for others in the surrounding population (and those with HIV/AIDS) who were living there.

Unintended consequences may be a result of systemic patterns but they are not always non-linear. A classic example of unintended consequences of improved service delivery can be an increase in the number of complaints as a result of people knowing about a service. Although

the response at first glance might be regarded as perverse, it is on closer examination highly predictable. On a macro scale we might look at Western slum clearance programmes. People who were rehoused into tower blocks benefited from radically improved facilities and frequently their material conditions were significantly improved. But the systemic effect of this was that the community was decimated. This led to a range of negative impacts, including a decline in mutual aid, depression and a loss of hope for generations of people. The programme logic of those interventions would denote their impact as positive. In hindsight many commentators saw them as very negative. The systemic effects of these changes are quite linear and predictable, but they are very significant.

Effective impact assessment has to take into account these complex change processes. This is recognised in a number of contemporary impact assessment documents, including Development Assistance Committee (DAC) guidelines on assessing development assistance which stress the importance of assessing:

> The positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended. This involves the main impacts and effects resulting from the activity on the local social, economic, environmental and other development indicators. The examination should be concerned with both intended and unintended results (OECD n.d.).

Implications for impact assessment The major implication of unintended consequences is that any assessment process has to look at the wider system upon which the intervention impacts. Once again, this requires some process of system mapping. It will also be helpful to ask explicit questions within impact evaluations such as those made explicit in this argument. For example:

■ Has this positive impact simply displaced negative outcomes to another place?
■ Has this positive impact been at the expense of other localities?
■ Has this intervention created dependency?

A second issue is that what appears obvious after the event does not always seem so at the time. Large-scale changes like slum clearance should be carried out on an iterative basis to see what the impacts will be. Stakeholders need to be engaged in the solutions to problems as they evolve.

Change is constrained by system dynamics *and* sustainable change requires changes in system dynamics
If our aim is to create sustainable change, we have to show more than that an intervention created a change. We have to show that the system dynamic changed. What is meant by a system dynamic?

A situation in the Kalangala islands that I have written about provides a good example (Burns 2014, forthcoming). Here there was a complex set of system dynamics that were fuelling the spread of HIV/AIDS. The core change dynamic highlighted by local people was the economic system. These local people are part of fishing communities where the movement of fish prompts movements in people – when the fish move, the men move with them, but the women (often with children) do not move. So the men take new women and the women are forced to take new men because there are no independent livelihoods for women on the islands. Formal 'interventions' mostly relate to sex education, but even the most knowledgeable woman who is educated not to engage in multiple relationships, had no choice because of her economic circumstances. The outcome of the intervention might be that she had a much better education, but for a change in her life, which impacted on her susceptibility to HIV/AIDS, we need to look for changes in the system dynamic, which in this case centres around the relationship between economic processes and sexual behaviour. This might tell us that the sorts of interventions we were taking were not appropriate to the task we need to accomplish. Better education was not enough to affect the underlying structure of incentives and constraints that resulted in the spread of HIV/AIDS.

Evaluation systems themselves are very good illustrations of system dynamics. If there is a culture of programmes being evaluated against log frames, this will put in place a set of incentives for organisations seeking funding to propose programmes that are based on interventions – the effects of which are broadly known – rather than attempts to find solutions to intractable problems.

So it is not only important to shift our core question from a focus on 'did the intervention work?' to 'did the intervention make a difference?' (Stern *et al.* 2012: ii), but also to ask whether the intervention made a difference to the system dynamic. Given this, it is crucial that the evaluation process is able to see the system, see the system dynamics, and to assess how these are changing over time.

Implications for impact assessment The critical factor here is that it is necessary to focus our impact assessment on how the system dynamic has changed, not on what changes have taken place *within* the system dynamic. This requires detailed mapping from multiple stakeholders. It is then necessary to look explicitly at what changes might shift the system dynamic. These can then be tested through action research, and assessed through later mapping of the system dynamics.

Latent change often leads to tipping points characterised by sudden major transformation
Multiple pressures within the system eventually lead to tipping points. This can look like extremely sudden large-scale change, but under the surface changes may have

been taking place for a long time – for example, changes in discourse. In complexity terms, alternatives to the dominant norms have the potential to create what are called 'attractors'. These gradually draw supporters into their orbit until they reach a critical mass when they are able to replace the dominant attractor which represented the status quo.

Some classic examples of tipping points can be seen in personal relationships. Lots of tiny changes happen which each change the relationship in subtle but significant ways. Then one partner does something small which triggers what appears to be a reaction that is disproportionate. Perhaps he forgot to book the night out that he had promised. For her it was 'the last straw' (the straw that broke the camel's back) and in that moment it became clear that the relationship was over because she reached a tipping point.

More broadly, these tipping points can be seen in relation to wider changes in society. For example, in a number of Asian countries there has been a sudden shift in behaviour towards older people, who are now regarded as a burden rather than deserving care. But also, interventions may be contributing to a major change in the underlying pattern of attractors but the tipping point may not yet have been reached, so it might appear as if there has been no impact at all. A good example of this is smoking in the UK. For decades research has shown that smoking causes cancer, but this 'evidence' alone has not been enough to create change in the system. Attempts to reduce the level of smoking have ranged from public health messaging, to alternative role modelling, to attacks on the tobacco industry, to local peer pressure, etc. Shifts in public attitudes have evolved. No longer do people find it acceptable to be in smoke-filled rooms. Perhaps this has to do with a rising tide of new narratives about health more generally, or about child protection. Perhaps some of these shifts came about through committed individuals in key roles, or shifts in corporate power. All of these things and many others were happening at the same time, shifting the underlying system dynamics. At a key point it became possible to introduce legislation that previously would have been politically impossible. The system had realigned, and once that happened the flood waters broke, resulting in a smoking ban in public places and in cars carrying children, visceral health warnings on cigarette packets and plain cigarette packets. The system realigns to make one action possible, then everything tips.

Figure 3 illustrates very clearly the different sorts of impact trajectories of various kinds of interventions. Woolcock (2013) argues that the point at which impact is assessed is crucial because of the dynamic nature of change.

In Figure 3, if impact assessment is done at t=1 then all of the initiatives appear to be doing well. If, on the other

Figure 3 **Understanding impact trajectories**
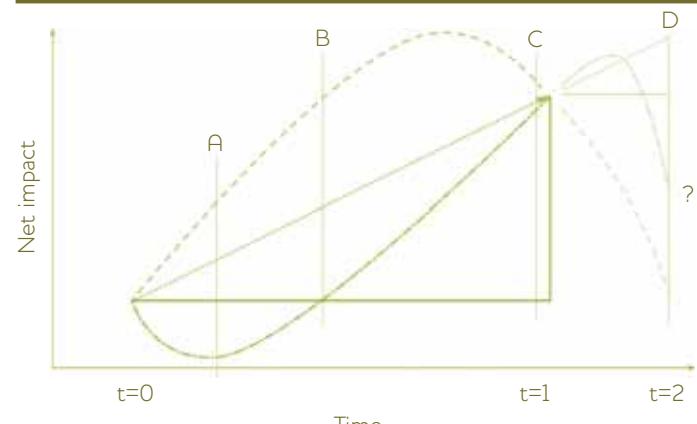
Source: Adapted from Woolcock (2013).

hand, it is t=2 then some appear to have had no effect at all and so on. The critical issue is to understand that change can happen in different timescales and in highly non-linear patterns.

Implications for impact assessment To assess impact we need to assess the underlying change as well as the surface level change. This might, for example, involve looking at how discourse is changing; or seeing where alternatives to the dominant system are emerging – and the extent to which they have traction and are building support. It also involves assessing how the environment – including power relationships – is changing. Might these shifts open the door to more radical change? Might new interventions nudge the system closer to tipping? The critical issue here is that you can assess interventions as having no impact when the changes that have been made may have contributed significantly (alongside others) to creating the conditions for a tipping point.

4 How Systemic Action Research can help
So how is Systemic Action Research (SAR) able to respond in assessing impact within complex systems?

At one level, the impact assessment element of SAR is straightforward. It is based on multiple stakeholders from across the terrain in which an intervention has been made (and beyond) making judgements about what is happening, what action to take, and later about whether it has created change or not. They determine what constitutes success and what data needs to be collected to assess that. Because it is a multi-stakeholder process in which people are engaged from across the system, the interpretation of the data is subject to strong critical scrutiny. It is thus the design of the learning system which enables it to respond effectively to complexity, not so much the methods themselves or the data that is assessed. The following are critical design features:

An iterative process

The iterative nature of the process enables participants to:

- assess impacts in real time;
- adjust action to maximise impacts;
- adjust programme goals to maximise impacts;
- set new baselines and indicators of success against those baselines.

Small steps are taken, the situation reassessed and adaptive action is then initiated. An iterative learning process allows you to see both unintended and intended consequences as they emerge in real time. This is a crucial part of a real-time impact assessment. In an action research process, every time you meet you ask the following questions:

- Are we still confident that we are going in the right direction?
- Are our success criteria the same as they were (last time we met)?
- Did the theories of change that informed our last actions still hold?
- Do we need to change our intervention?
- Do we still have the right people in our inquiry?
- Do we have all of the information that we need?
- What methods do we need to collect that information?
- What has changed in the system, and what specifically has changed as a result of the action we have taken?

The quality of data is also radically enhanced by using multiple methods and generating different types of data from different sources. Because at each stage of the process questions are asked about the methods that are needed to achieve each new purpose, the methods are always fit for purpose.

Tested in action

A critical dimension of action research is that theories of change that are generated are tested in action. Take the example of the Re-thinking Vulnerability project that I co-facilitated with the British Red Cross: one of the emerging theories of change generated by the early inquiry groups was that people who had been helped by Red Cross volunteers, who lived in communities with higher levels of vulnerability, and who subsequently became volunteers themselves, were the volunteers most likely to be able to create effective change within those communities. This theory of change became the basis for a pilot programme across the Red Cross that intentionally recruited people with this profile, put in place different resources and processes to support them, and then tested that assumption. Once the theory of change was verified as a result of this action, the learning was fed into the development of a new UK service strategy. The outcomes of the action provide evidence of the theories of change that have been advanced in the analysis. Embedding the

testing of action into the learning process adds substantive rigour to any analysis. Here we can see that the collective analysis is not enough to base new programming on. The collective analysis identifies a pattern, which needs to be further substantiated through action. The success of that action is assessed by the multiple stakeholders in the inquiry process.

Recorded in real time

As participants engage with these questions they are recorded. Every meeting, every decision and every action is documented at each stage of iteration. Dialogues, informal conversations, video material, reflective analysis, outcome and impact data, collective analysis, etc., are all closely documented and this provides an embedded and highly transparent audit trail of knowledge, decisions and impact. With the data collected and analysed as you go, gaps in the data are identified while you work. This means that as well as having real-time information upon which to base decisions, you are not trying to assess impact with poor data that is almost impossible to retrieve later. This process enables participants in SAR processes to track changing baselines in the context of highly dynamic situations and emerging choices.

Built on a systemic analysis of how change happens

SAR processes look at issues systemically so it is possible to see how things that would normally be outside of the gaze of an impact assessment are impacting on what is happening. The systemic mapping generates strong baselines about the dynamics of change against which to assess any interventions; the process by which change happens can be tracked; unintended consequences into other domains can be seen and assessed. Furthermore, by focusing on system dynamics the evaluation and impact focus is centred on the things that matter – not every possible change that might happen. This means that the data that needs to be collected can be more focused on those dynamics.

A participatory process

Because it is a participatory process involving multiple stakeholders, SAR groups can identify what they think success looks like, and what indicators might denote that success. Impact can be assessed against these indicators. As SAR is based on a large-scale learning architecture, the detailed changes that are identified by action research groups can be tested across the wider system. Resonance of narratives, theories of change and the meaning of indicators can be tested in multiple arenas. Because systemic inquiry builds evidence in multiple arenas and at different levels (e.g. community level, public service providers, policymakers) it can test the resonance of evidence emerging in one domain in the other domains.

This is a powerful process of triangulation. Triangulation is not unique to action research processes. What is significant here is the way in which the SAR architecture enables knowledge to be assessed in multiple domains. This means that what starts as rich localised knowledge can be triangulated so that meaning is generated from participatory data at scale.

In an action research process the assumption is that an analysis of impact can best be done by the stakeholders. This enables ownership of the results, but also provides a higher quality of analysis. For example, a group of local women who do not use a multi-million pound investment in public toilets, can interactively explore their stories of why. These might range from the fact that they fear rape, that they cannot afford the charges, that the toilets are being vandalised, that they do not like to wash in a public place, etc. They can explore the interrelationships between these things, and their conclusions will be the ones that are relevant to what any future intervention might look like. Collective analysis (which is a strong feature of SAR) is less vulnerable to researcher bias because it involves multiple perspectives that are close to the ground (as well as external voices), subjecting data and analysis to real-time critical scrutiny. It is more ethical from a normative perspective because it adheres to the well-articulated principle drawn from the disability movement – 'nothing about us without us'.

Quality data rooted in relationship building

A good analysis of impact depends on the quality of the data about what is happening, and about people's views and perceptions. Interventions that have been completed to high technical standards and which appear to be the optimum technical solution, fail because they have not taken into account culture and norms. The quality of information available to evaluators can be low because:

■ there has not been sufficient time for deliberation or critical scrutiny;
■ there are inadequate records of what has happened and what has changed;
■ researchers cannot get access to honest information – for example, people tell researchers what they think they want to hear; and
■ researchers cannot get access, full stop, because they are not trusted by communities.

Action research is rooted in relationships. These can take time to build, but the rewards are that action researchers have access to a depth of knowledge that academic researchers often cannot get close to. Action research groups directly involve stakeholders who can work independently with their peers. A young person who is not attending school can engage with others who are not attending school because he or she is trusted by them. This is likely to

produce a higher quality of knowledge than interviews with a researcher, which reveal little because there is no trust. This approach aligns with perspectives in evaluation which stress the 'knowledgeability of the social actor'. Social actors know the reality better than the researchers because they live in it (Pawson and Tilley 1997: 162–3). Even if you believed that an 'objective' analysis was possible, there is no better 'objective' analysis to be had, because the researcher will never get access to what is really going on.

5 Challenges in the use of Systemic Action Research

SAR is of course not a silver bullet. There are many challenges, and there are many situations in which it is simply not the best approach to use. In this section I elaborate on a few of these.

An approach like SAR requires donors and programmes to be comfortable with not knowing everything before they start: SAR is based on 'learning as you go'.

The SAR process takes time, and is therefore best built into the programme from the start and certainly while it still has considerable time to run. Participatory Systemic Inquiry can be done more quickly but even if there is no ongoing inquiry, time has to elapse after any baseline system mapping in order to assess shifting system dynamics. So SAR is not really suited to quick impact assessments.

The most important skills in SAR processes are facilitation skills (followed closely by good documentation skills). Facilitators need to be trained to allow open conversations, to encourage new lines of thinking, to guide rather than to lead – skills that are not ubiquitous. This means that most of the action research work I do involves a capacity development phase.

In much of my recent writing I have highlighted the trade-off between participatory and systemic approaches (Burns 2012). A traditional action research process is firmly locked within a group of stakeholders (group, community) who have a collective problem that they want to solve. They completely own the process and all of the decisions. A systemic approach works with multiple people across the system. Because of this, it is much better at surfacing and understanding the complex power relationships and system dynamics, but the process is driven from multiple points across the system and therefore the facilitators that work across the system hold considerable power. While I have identified some ways of mitigating this and maintaining a strong participatory ethos within Systemic Action Research (Burns 2012), there is still room for the development around this tension.

SAR is a research and evaluation methodology which is most appropriate for use in complex and highly dynamic

environments. As an iterative process of assessment, it is especially suited to working with environments where change is assumed to be systemic and transformative.

Bearing these issues in mind I have found that this approach works most effectively:

■ when assessing the impact of a whole programme or organisation – for example, the impact of volunteering on poverty;
■ when assessing the extent to which deeper system dynamics have been changed – thereby supporting sustainable change;
■ when enabling ongoing assessment in dynamic environments and emergent change processes;

■ when there is time for learning. Typical timeframes for an SAR approach will be one to two years. They are designed to be embedded within programmes, not for rapid end of programme evaluations. Some of the tools associated with SAR, such as systemic mapping, can be used for summative evaluation, but SAR is best used as a formative assessment process where the data that is collected along the way can be used;
■ where there are diverging perspectives on what has happened and what counts as success.

For those who are interested in methods that can engage effectively with change in complex and dynamic systems, SAR may be an appropriate option.

Notes

1   British Red Cross 'Rethinking Vulnerability' programme; VSO 5 country study of the impact of volunteering on poverty currently being directed by IDS.

2   www.ces-vol.org.uk/about-performance-improvement/about-monitoring-evaluation/planning-for-monitoring-evaluation/outcomes/outcomes-impact.

References

Befani, B. (2012) 'Models of Causality and Causal Inference', in E. Stern, N. Stame, J. Mayne, K. Forss, R. Davies and B. Befani (eds) *Broadening the Range of Designs and Methods for Impact Evaluations*, DFID Working Paper 38, London: Department for International Development

Burns, D. (2014, forthcoming) 'Systemic Action Research', *Encyclopedia of Action Research*, London: Sage

Burns, D. (2013) 'Systemic Action Research: Changing System Dynamics to Support Sustainable Change', *Action Research* 12.1: 3–18

Burns, D. (2012) 'Participatory Systemic Inquiry', *IDS Bulletin* 43.3: 88–100

Burns, D. (2011) 'Facilitating Systemic Conflict Transformation Through Systemic Action Research', in D. Körppen, N. Ropers and H.J. Giessmann (eds), *The Non-Linearity of Peace Processes – Theory and Practice of Systemic Conflict Transformation*, Opladen/Farmington Hills: Barbara Budrich Publishers

Burns, D. (2007) *Systemic Action Research: A Strategy for Whole System Change*, Bristol: Policy Press

Burns, D. (2006) 'Evaluation in Complex Governance Arenas: The Potential of Large System Action Research', in B. Williams and I. Imam (eds), *Systems Concepts in Evaluation: An Expert Anthology*, Washington DC: American Evaluation Society

Checkland, P. and Scholes, J. (2004) *Soft Systems Methodology in Action*, Chichester: John Wiley

Kolb, D.A. (1984) *Experiential Learning: Experience as the Source of Learning and Development*, Englewood Cliffs, NJ: Prentice Hall

OECD (n.d.) *DAC Criteria for Evaluating Development Assistance*, www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm (accessed 11 July 2014)

Patton, M. (2011) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, New York: The Guilford Press

Patton, M. (1994) 'Developmental Evaluation', *Evaluation Practice* 15.3: 311–20

Pawson, R. and Tilley, N. (1997) *Realistic Evalution*, London: Sage

Ramalingam, B. (2013) *Aid on the Edge of Chaos*, Oxford: Oxford University Press

Reason, P. and Bradbury, H. (2008) *Handbook of Action Research*, London: Sage

Rogers, P. (2008) 'Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions', *Evaluation* 14.1: 29–48

Stern, E.; Stame, N.; Mayne, J.; Forss, K.; Davies, R. and Befani, B. (2012) *Broadening the Range of Designs and Methods for Impact Evaluations*, DFID Working Paper 38, London: Department for International Development, http://r4d.dfid.gov.uk/pdf/outputs/misc_infocomm/DFIDWorkingPaper38.pdf

Wadsworth, Y. (2010) *Building in Research and Evaluation: Human Inquiry for Living Systems*, Melbourne: Action Research Press and Allen & Unwin

Wadsworth, Y. in association with Merinda Epstein (2001) *The Essential U and I: A One Volume Presentation of the Findings of a Lengthy Grounded Study of Whole Systems Change Towards Staff–Consumer Collaboration for Enhancing Mental Health Services*, Melbourne: Victoria Health Promotion Foundation

Williams, B. and Imam, I. (eds) (2006) *Systems Concepts in Evaluation: An Expert Anthology*, Monograph 1.1, Washington DC: American Evaluation Society

Woolcock, M. (2013) 'Using Case Studies to Explore the External Validity of "Complex" Development Interventions', *Evaluation* 19.3: 229–48

> " If our aim is to create sustainable change, we have to show more than that an intervention created a change. We have to show that the system dynamic changed. "

Centre for Development Impact (CDI)

The Centre is a collaboration between IDS (www.ids.ac.uk) and ITAD (www.itad.com).

The Centre aims to contribute to innovation and excellence in the areas of impact assessment, evaluation and learning in development. The Centre's work is presently focused on:

(1) Exploring a broader range of evaluation designs and methods, and approaches to causal inference.
(2) Designing appropriate ways to assess the impact of complex interventions in challenging contexts.
(3) Better understanding the political dynamics and other factors in the evaluation process, including the use of evaluation evidence.

IDS Institute of Development Studies

itad

UKaid from the British people