Representing Theories of Change: Technical Challenges with Evaluation Consequences

Rick Davies[1]

1 Evaluation Consultant

Suggested Citation: Davies R, 2018, Representing Theories of Change: A Technical Challenge with Evaluation Consequences, CEDIL Inception Paper 15: London

About CEDIL: The Centre of Excellence for Development Impact and Learning (CEDIL) is an academic consortium supported by UKAID through DFID. The mission of the centre is to develop and promote new impact evaluation methods in international development.

Corresponding Author: Dr Rick Davies, email: rick.davies@gmail.com

Table of Contents

Abstract

This paper looks at the technical issues associated with the representation of Theories of Change and the implications of design choices for the evaluability of those theories. The focus is on the description of connections between events rather than the events themselves, because this is seen as a widespread design weakness. Using examples and evidence from Internet sources six structural problems are described along with their consequences for evaluation.

The paper then outlines a range of different ways of addressing these problems which could be used by programme designers, implementers and evaluators. The paper concludes with some caution speculating on why the design problems are so endemic but also pointing a way forward. Four strands of work are identified that CEDIL and DFID could invest in to develop solutions identified in the paper.

"Essentially, all models are wrong, but some are useful" George Box, 1987

"At the heart of all major discoveries in the physical sciences is the discovery of novel methods of representation …" Stephen Toulmin (1953: 103)

"Ninety per cent of problems have already been solved in some other field. You just have to find them." Tony McCaffrey (Marks, 2015)

Section 1

What is a theory of change?

Although the idea of using a Theory of Change as an aid to evaluation has been around for a long time (e.g. Weiss, 1995) interest in its use has been especially notable in recent years amongst international development aid organisations. Guides and literature reviews on Theories of Change have been published by Comic Relief (James, 2011), DFID (Vogel, 2012), ESPA (Vogel, 2012b), The Asia Foundation (Stein & Valters, 2012), UNICEF (Rogers, 2014), ODI (Valters, 2015), HIVOS (van Es et al, 2015) and UNDAF (UNDG, 2017). There have also been many blog postings on the subject (e.g. Green, 2011; Davies, 2016a).

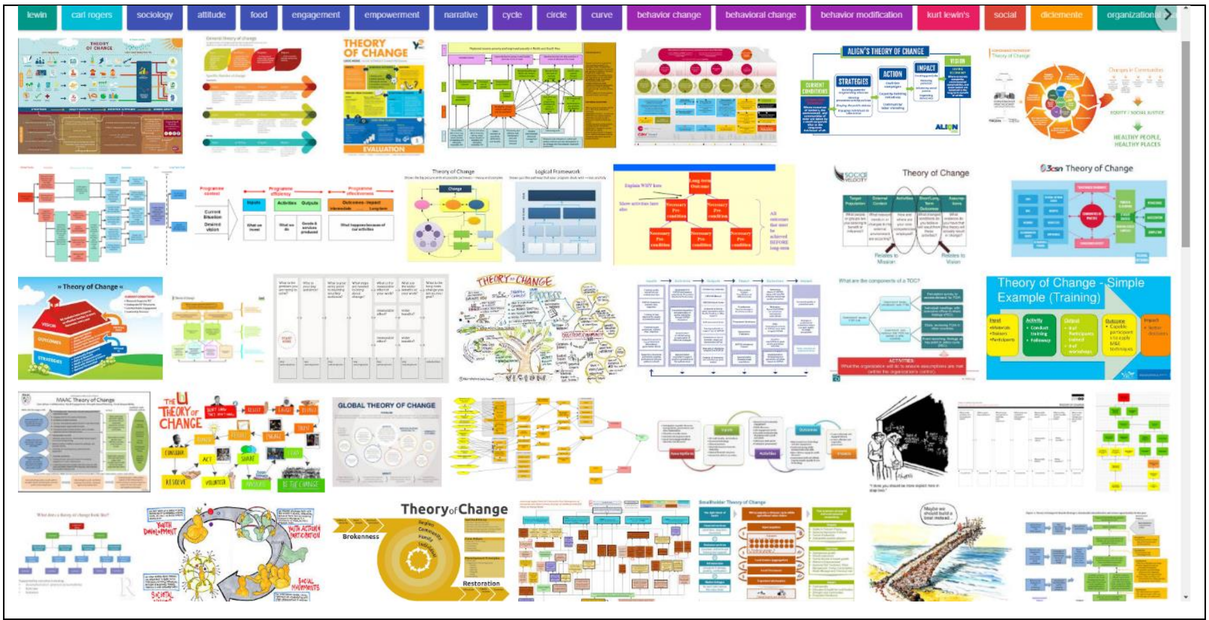
Carol Weiss, one of the earliest popularisers of the idea, described a Theory of Change as "a theory of how and why an initiative works." More recently in their review of the use of Theory of Change in international development Stein and Valters (2012) have explored various interpretations and concluded that despite the variety of views "Theory of Change is most often defined in terms of the connection between activities and outcomes, with the *articulation of this connection the key component* of the Theory of Change process" (emphasis added).

A Theory of Change typically involves some form of diagrammatic representation, usually supported by a text commentary. It is the diagrammatic representations that are the focus of attention in this paper. Diagrams are capable of succinctly summarily representing multiple parallel and intersecting causal pathways in ways that a textual narrative cannot. However, it is recognised that the narrative component of a Theory of Change can provide much needed detail on particular elements within a diagrammatic representation of a Theory of Change.

That said, there are some differences of opinion as to whether a Logical Framework matrix also qualifies for inclusion as a Theory of Change of the kind that will be discussed here. A Logical Framework matrix is a tabular structure for representing program logic in development projects, widely used by development agencies since the 1990s. The contents of its rows describe a sequence of "if…and…then" statements, connecting project activities, at the base, via linking assumptions and intermediary activities, to desired outcomes at the top. Given these features, it does meet the minimal requirements of a Theory of Change, as described by Valters (2015) above.

Representations of Theories of Change exist in many and varied forms, as can be seen in the results of a Google Images search for "Theories of Change", shown in Figure 1.

Figure 1: Varieties of Theories of Change



Google Images search result: "Theories of Change"

One reason for this diversity is the wide variety of contexts in which they have been developed. This is especially the case with Theory of Change of development aid projects being implemented across a range of countries and sectors, and which are the focus of this paper. The other is that a Theory of Change can be developed for different purposes. Since the 1990s there has been something akin to adaptive radiation[1] in the uses made of Theories of Change. They can be used at all stages of a project cycle: to articulate a programme design, to identify and build agreement among stakeholders, to inform the design of monitoring and evaluation systems, to focus individual evaluations, and to structure reporting to donors and other stakeholders (Stein and Valters, 2012; Mayne, 2015). There has also been some concept speciation, with distinctions now being made between Theory of Action and Theory of Change[2], the latter referring to how a social, political, economic and/or cultural change happens, and the former referring to how a particular program contributes to the change process. Distinctions have also been drawn between a Theory of Change and a Logic Model (Mayne, 2015; Dhilon and Vaca, 2018). Here the

---

[1] https://en.wikipedia.org/wiki/Adaptive_radiation

[2] http://www.kstoolkit.org/Theory+of+Change+%26+Theory+of+Action

phrase Theory of Change is used in the more inclusive and generic sense, as used by Weiss.

This paper is one of a series commissioned by the Centre of Excellence for Development Impact and Learning (CEDIL), which has been funded by DFID to develop and test innovative methods for evaluation and evidence synthesis. Hence, the focus on the evaluation functions of a Theory of Change, and in particular the exposition of *how* change is expected to happen. The focus is on the use of Theories of Change as *products* for evaluation purposes. While saying this it acknowledges that the *process* of developing a Theory of Change, especially participatory design processes, can make a significant difference to the ownership of a Theory of Change and this also has consequences for evaluation. But that process dimension is not discussed in this paper.

Section 2

What is the problem?

The development of a good representation of a Theory of Change involves managing at least two competing and valid requirements. One is for simplicity, to ensure readability and thus usability. Warnings are often given about the need to avoid undue complexity that will make Theory of Change unreadable by their intended users (Funnell and Rogers, 2011; Green, 2012). The second requirement is for sufficient detail, to ensure some match with the complexity of the real world. This is essential if the Theory of Change is to be evaluable. Evaluability has been defined by the OECD-DAC (2010) as ""The extent to which an activity or project can be evaluated in a reliable and credible fashion". An adequate Theory of Change can be considered as a necessary but insufficient basis for project evaluability[3]. Available data on what subsequently happened and appropriate stakeholder engagement are also necessary (Davies, 2013). If a Theory of Change is evaluable it should be possible to pursue Weiss's (1995:67) aim "...to examine the extent to which program theories hold".

The focus of this paper is on the technical challenges involved in developing an evaluable Theory of Change, and how these challenges might be resolved. Crudely summarised, diagrammatic representations of Theory of Change are typically made up of boxes and arrows. Boxes are filled with text descriptions of events, and arrows connect them, representing expected causal connections between these events.[4] The central proposition is that it is the connections between events in a Theory of Change diagram that is most

[3] The one exception to the need for a Theory of Change might be Goal-Free Evaluation (Scriven, 1991).

[4] As can be seen in a Google Image search, Theories of Change are sometimes represented in more metaphoric forms, using landscapes, houses, trees etc. They suffer from essentially the same problems as seen with more diagrammatic representations discussed here.

problematic, not the descriptions of those events[5]. Problems exist both in the content and structure of these connections.

These problematic features affect the plausibility and testability of Theories of Change. That said, a testable Theory of Change does not imply a commitment to blueprint planning and the impossibility of an adaptive approach. Theories of Change can and do get adapted in the course of programme implementation. Nor does it imply a requirement of complete and certain knowledge about the future. Theories of Change are hypotheses, which should be updated in the light of experience[6].

Both the analysis and the proposed solutions have been informed by different perspectives: the literature on social network analysis (summarised in Borgatti, et al, 2018), set theory views on causal models and inference (Goertz and Mahoney, 2012, Rihoux and Ragin, 2009), the concept of impact trajectories (Woolcock, 2009) and a recent CEDIL paper on causal chain analysis (Gough et al, 2018). In addition, the management of complexity in Theories of Change has been a long-standing interest of my own (Davies, 2004, 2005).

Sources: The main source of examples for the arguments being made in this paper is the Google Images search result for "Theories of Change" and "Logic Models" with a focus on images that meet the Stein and Valters' criteria of showing connections between entities. This sample includes examples from domestic as well as international programs, across a wide range of sector. A second source was a collection of 11 postings on Theories of Change to the AdaptDev email list in January 2018 by members of that list. These all related to international development aid programmes.

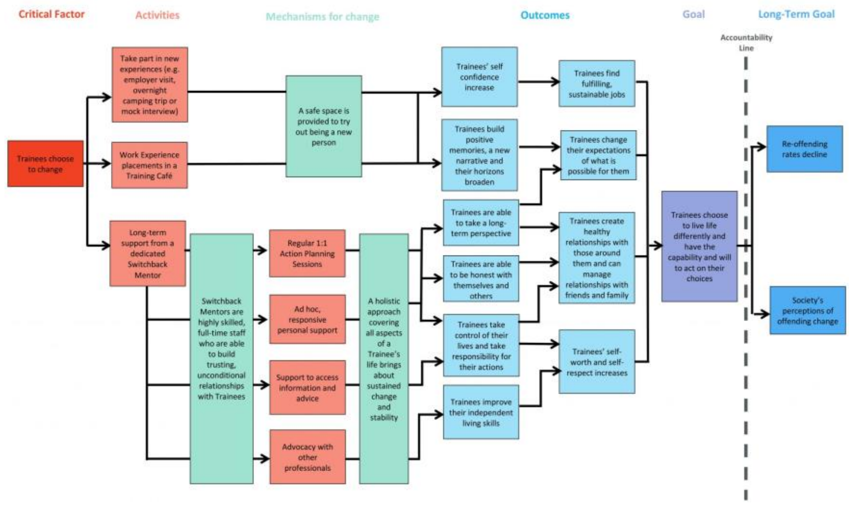
Problem 1: Unlabelled connections

It is almost universally the case that the arrows connecting events in a Theory of Change diagram are without text annotation, or any form of colour or shape coding, which tells the reader more about the nature of those connections. None of the 30+ examples found via the Google Image search shown in Figure 1 above provides any information about the nature of the linkages between events. The same applies to the "logic models" search result. We typically know nothing about timing, duration or scale of the causal connection, or anything about the actual mechanisms at work. At best, this information might be inferred from the text of the boxes they connect, as in Figure 2.

[5] For my criticism of the text content of "boxes", especially in Logical Framework models, see http://mande.co.uk/category/lists/the-logical-framework/#editor

[6] Along with dates and who participated in the revisions. This will create a trail of evidence on how change was perceived and managed over the course of a given intervention (Shaw, 2018)

Figure 2: A Theory of Change for reduction in reoffending by ex-prisoners



Switchback (2018)

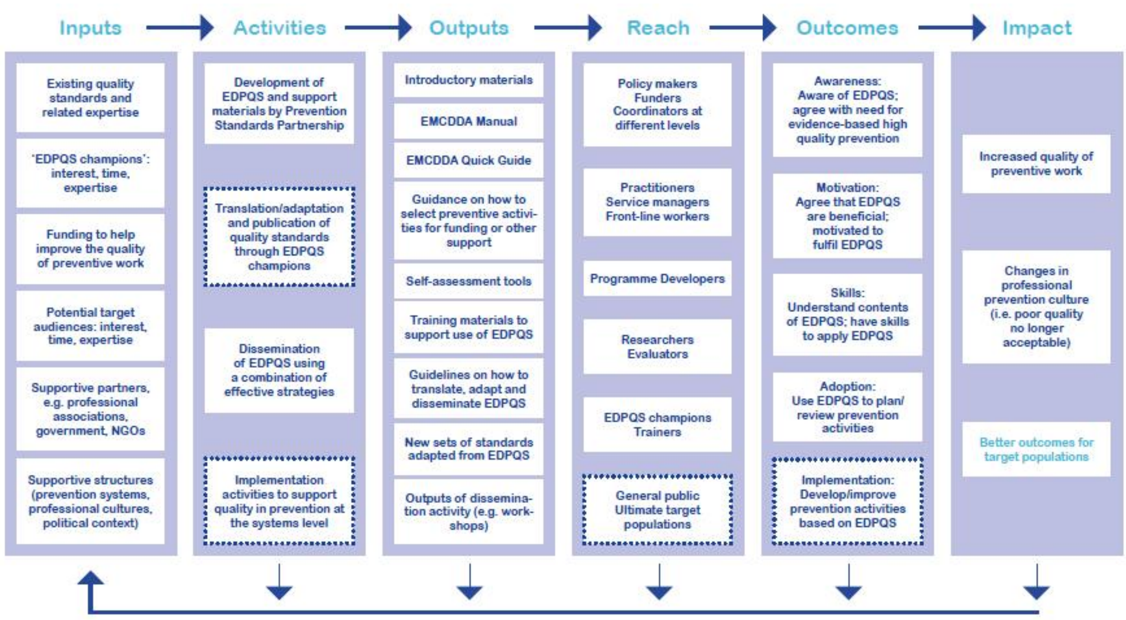
Figure 2 is also a good illustration of why diagrammatic representations of Theories of Change are useful, relative to text-only descriptions. Even though it uses time as the main axis, it would be difficult to create and read a narrative description of a Theory of Change that described and communicated the multiple parallel and interacting causal chains as shown in this diagram.

Problem 2: Missing connections

In some Theories of Change, the main problem is missing connections. These are typically, but not exclusively, seen in chain models that show how a list of one type of events is connected to a list of another type of events. For example, as in Figure 3. This problem can also be found in narrative descriptions, particularly in strategy documents, where the lists are described as vision, mission, strategies, themes, focal areas, outcomes, impact areas and sundry other abstractions, described in a sequence.

The traditional structure of the Logical Framework was problematic in this respect. It never had any means of explicitly connecting individual events in one layer with individual events in an adjacent layer (Coleman, 1987). However, more recent uses of the Logical Framework (DFID, 2017), following revisions to its structure (DFID, 2011) now nest sub-groups outputs under specific outcomes, going some way to address the problem of missing connections. It introduced and requires an impact weighting describing each output's expected contribution to the associated outcome – addressing an aspect of Problem 1: Unlabelled connections above.

Figure 3: Theory of Change for European Drug Prevention Quality Standards



EDPQS (2018)

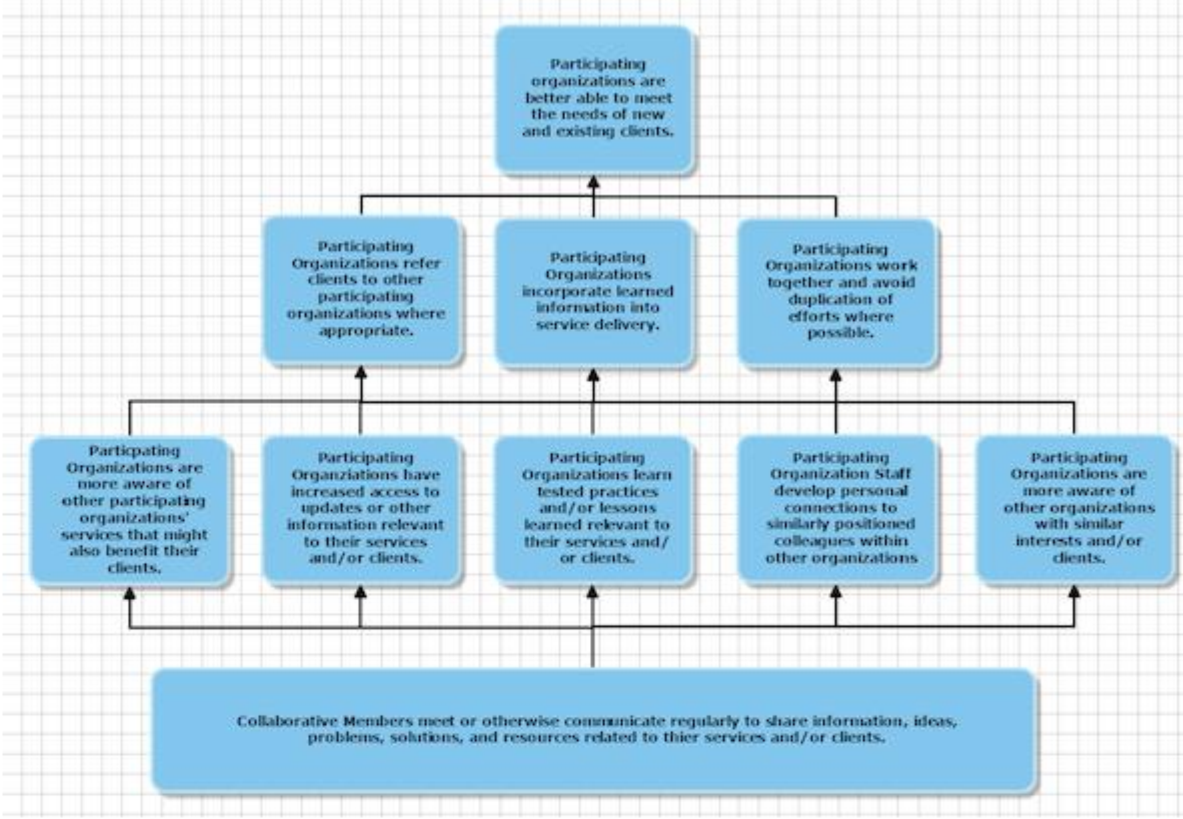
The scale of the problem present in chains of lists should not be underestimated. For example, in Figure 3 there are nine Outputs feeding into six categories of users, within the "Reach" column. There are 9 x 6 = 54 possible links that could exist here, any one (or more) of which could be the subject of attention by an evaluation. If a multiple conjunctural causation perspective (Rihoux and Regan, 2009) was adopted by an evaluator then there are 2^54 different combinations of these connections which could be important. Clearly, there is a lot of work to be done articulating the details of this programme before it would be transparent where it would be best to invest in evaluation resources. As Weiss (1995:69-70) argued, a good Theory of Change "concentrates evaluation attention and resources on key aspects of the program… No evaluation, however well-funded, can address every question that might be of interest to someone."

This representational problem is not "academic". As more emphasis and attention is being given to adaptive and flexible programming, it is likely that the menu of outputs of such programmes will become more varied and more changeable than the fabled blueprint projects of the past. This development will present two types of problems. One is the "curse of dimensionality" – that as the number of programming variables increases the number of ways they can combine grows exponentially. The other is the limited ability to identify *in advance* the nature of the expected connections between particular outputs and expected outcomes.

Problem 3: Symmetric connections

Figure 4 provides an example of a style of representation that is surprisingly common. In these Theories of Change, it is aesthetics which seems to be the primary design consideration. The connections between events provide no more information than to say, "This lot of events leads to this lot of events" (i.e. Problem 2). The combinatorial problem remains.

Figure 4: Theory of Change for evaluating community coalitions and collaboratives



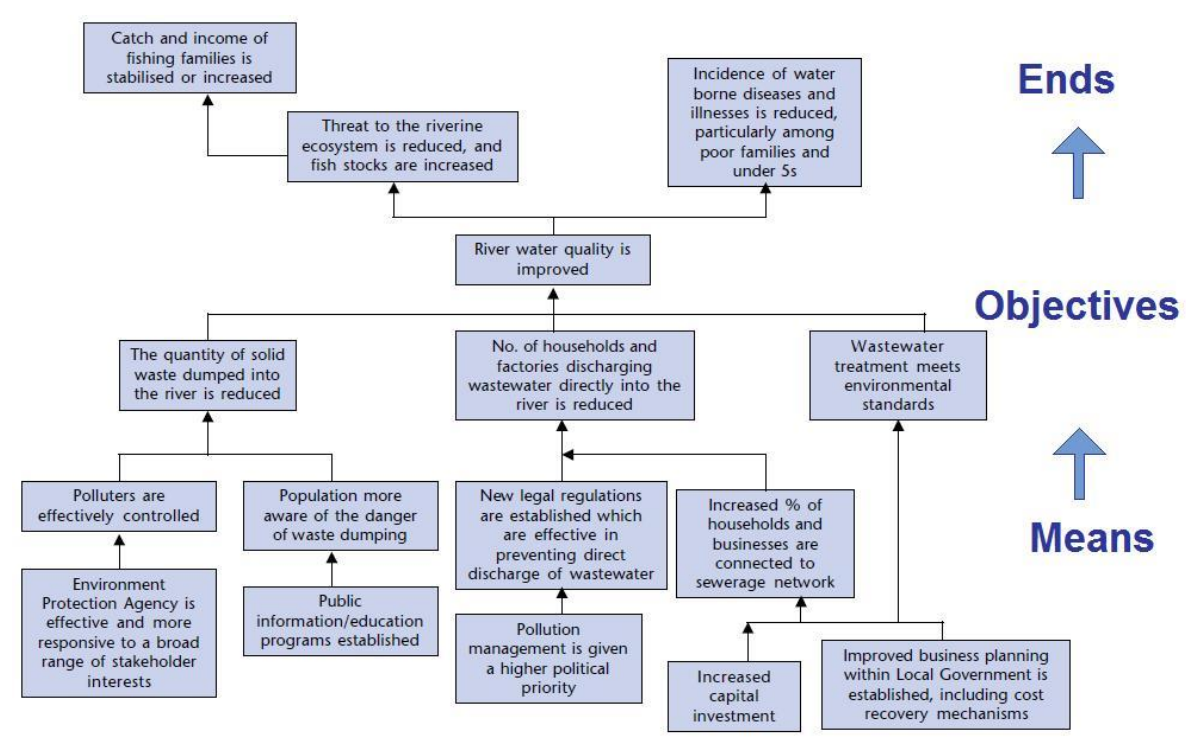
Levinton (2015)

While this problem is common in Theories of Change used as teaching examples (Levinton, 2015), it can also be seen in Theories of Change representing large and complex real-life programmes, such as DFID's $48 million Humanitarian Innovation and Evidence Programme (ITAD, 2016).

Problem 4: Numerous pathways

There are at least two versions of Theories of Change representations that are relevant here. The first are often derived from a "problem tree" analyses, which in the past at least, were used at the early stage of the design of a Logical Framework. After Problem Trees are identified they are reconstructed into objective trees, with antecedents and consequences. Sometimes these are in the shape of an inverted pyramid, others are in the shape of an hourglass. Figure 5 is an example of the latter. In this Theory of Change, there are seven different pathways to the mid-level objective, but the diagram provides no information to an evaluator about the relative significance of each of these pathways. However, given the limited scale involved in this particular example, this would not necessarily be an unmanageable problem for an evaluation team.

Figure 5: Theory of Change about the control of water pollution



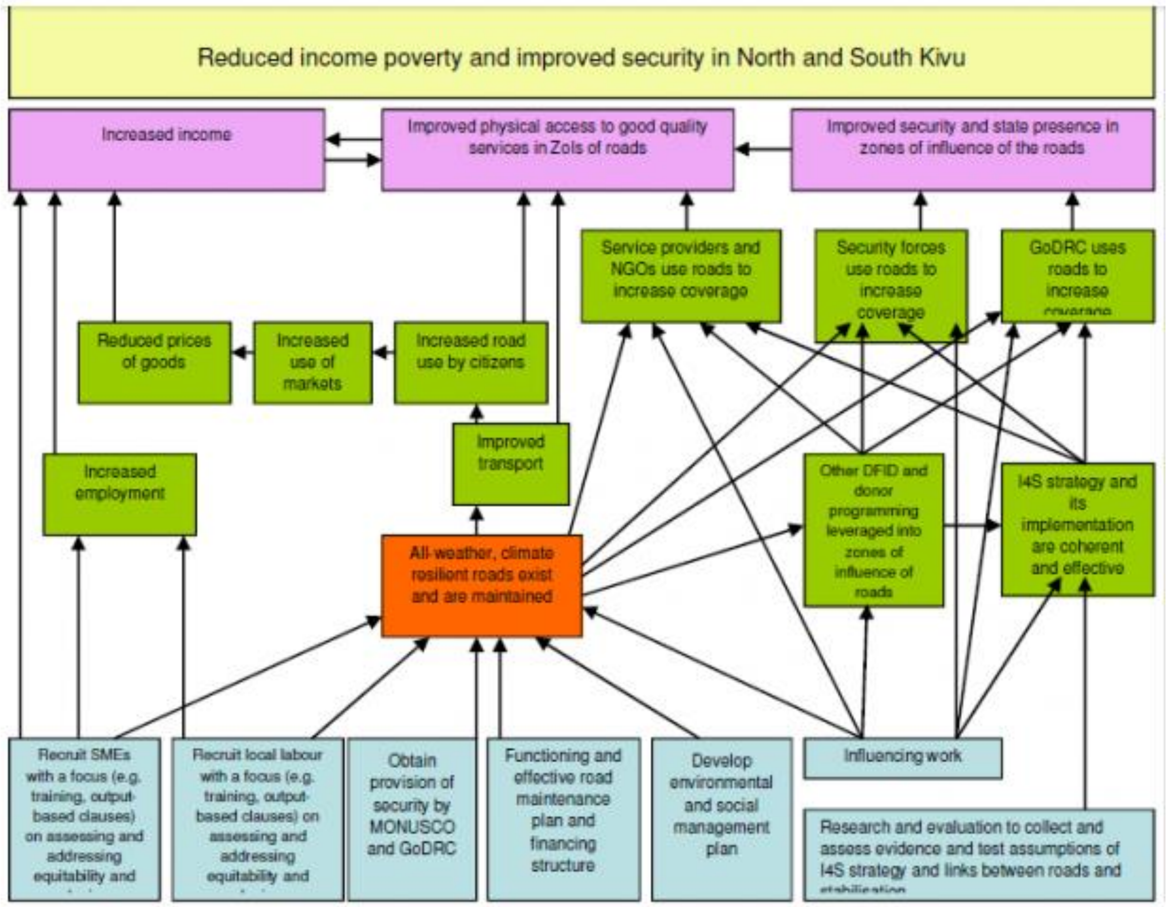
Sustainable Sanitation and Water Management (2018)

The second type, shown in Figure 6 below, is more complicated and challenging. It is a heterarchy rather than nested hierarchy – i.e. a given event can contribute to more than one outcome. In network terms this kind of structure has a higher network density, there are more interconnections between the various events (boxes). This means there are many more potentially important pathways through which causal influences can work.

This type of representation reproduces on a larger scale the same combinatorial explosion problem seen earlier on a small scale, in Figure 3, where the focus was on relationships connecting events in two boxes. In Figure 7 there are upwards of 50 distinct pathways which may be at work, and 2^50+ possible combinations of these, if each are treated as binary options.

The situation is potentially more complicated still. This, and most other diagrams like it, do not tell us anything about timing requirements, of what inputs into an event need to precede other inputs. Yet when it comes to evaluation planning, expectations of likely outcomes are likely to be affected by expectations about timings of relevant inputs. In reality, there is both a combination and permutation problem, it’s not only the combination of events but also their sequencing, which needs to be clarified.

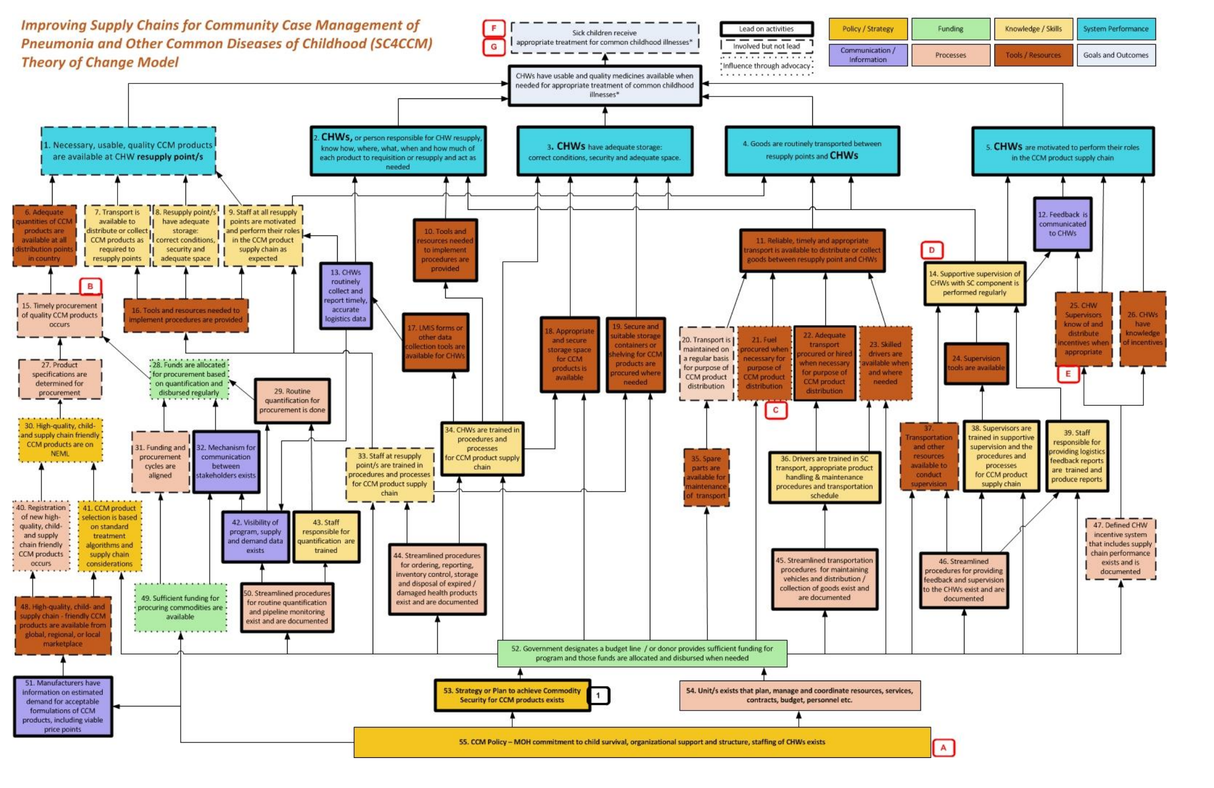
Figure 6: A heterarchy with highly interconnected pathways



DFID Programme Theory of Change: Roads in East DRC (DFID, 2012)

Figure 7 represents another real programme that is much more complicated. Theories of Change developed using participatory processes involving different stakeholders can be especially complicated. Figure 8 is an example of the "Post-It Note" stage of the development of a Theory of Change for the IFAD funded Agricultural Services Support Programme (ASSP) and Agricultural Sector Development Programme – Livestock (ASDP-L) in Tanzania.

Figure 7: Theory of Change for the Supply Chain for Community Case Management project



(JSI Research & Training Institute, Inc, 2016)

Figure 8: Theory of Change for the IFAD funded Agricultural Services Support Programme (ASSP) and Agricultural Sector Development Programme – Livestock (ASDP-L) in Tanzania

Pabari, M, (2008)

Through the discussion of the examples shown above, we can see that the diagrams used to represent Theories of Change leave understated the huge range of possibilities that may be taking place. Chain models, nested hierarchies and heterarchies all have their limitations, although the latter are probably a better approximation of the real world. A lot more clarification would be needed before any of these can be evaluated within a realistic time frame.

In addition, the combinatorial possibilities of the causal connections described in this small sample of Theories of Change highlight the humbling fact that while evaluations can in practice only test a few theories at a time, there are likely to be many more untested but potentially important causal pathways out there which may have a better fit with the data, if and when it becomes available.

All the above has been concerned with just how complicated some Theories of Change can be. The task of evaluators becomes more challenging when we look at Theories of Change that describe complex, rather than complicated programmes. It is the presence of feedback loops in diagrams that make the difference, as will be explained.

Problem 5: Feedback loops

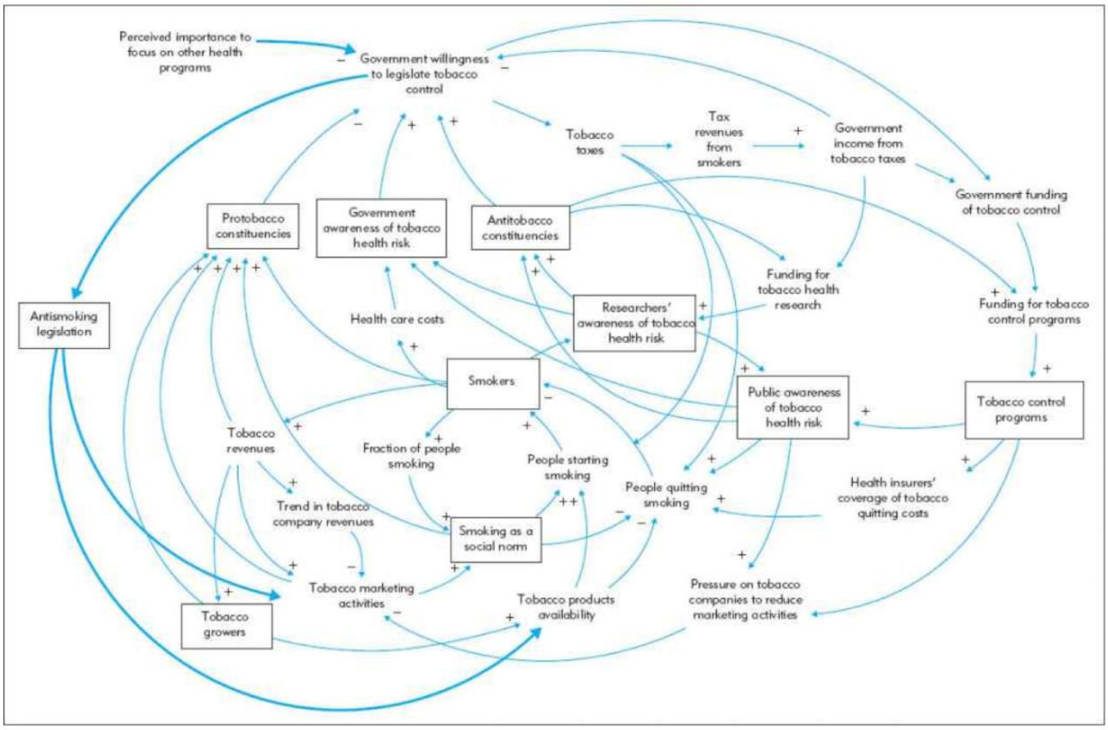
Theories of Change with feedback loops can sometimes be found in chain models, nested hierarchies and heterarchies, but they tend to be uncommon. In the Figure 1 sample, 14% of the diagrams had some form of feedback loop. In a search for “Logic Model” images, they are much rarer. Where there are no feedback loops this implies a process of change that has a linear trajectory, with a constant rate of change projected into the future. In principle, this would be evaluable, because expected outputs at a given point of time could be predicted. In practice, one test of evaluability would be stakeholders’ willingness to own the ambitious predictions from such a theory.

Where feedback loops have been included the next most common problem is lack of information as to whether they are positive or negative feedback loops, though this can sometimes be inferred from the contents of the connected events.

Where feedback loops have been labelled, positive feedback loops are the most common. In the absence of any negative feedback loops, this implies an exponential trajectory of change, which is arguably much less plausible than a linear trajectory. On the other hand, negative feedback effects have dampening effects, reducing the scale on which change can be achieved. In a world of conflicting stakeholder interests, and other things being equal, these are likely to be the more plausible theories.

The presence of both positive and negative feedback loops brings models closer to being real-world approximations. But they also create new technical challenges. Firstly, in cases like Figure 9, the resulting trajectory of change is no longer so evident from visual inspection only.

Figure 9: A Theory of Change about government control of tobacco

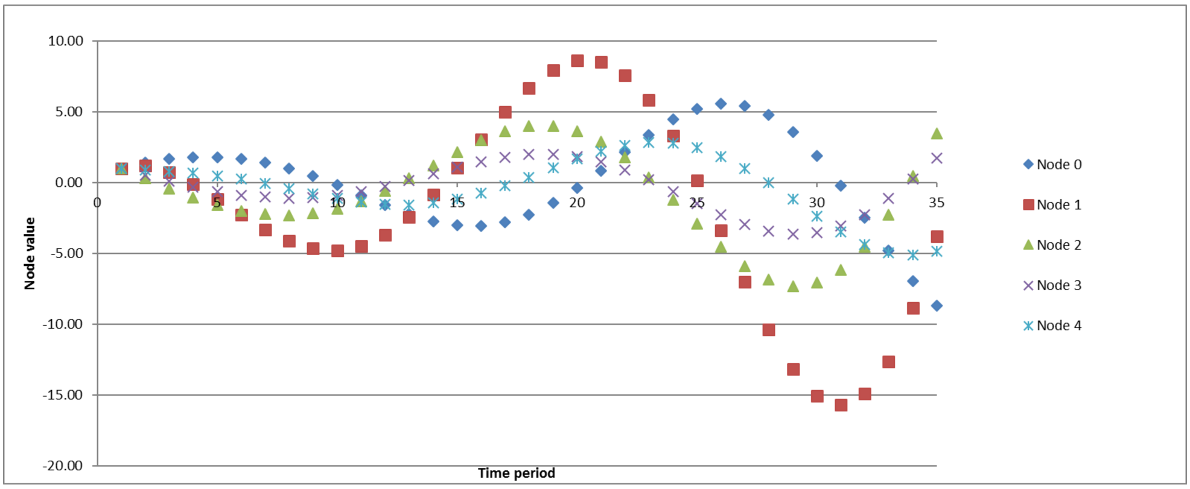


Haiku Analytics (2018)

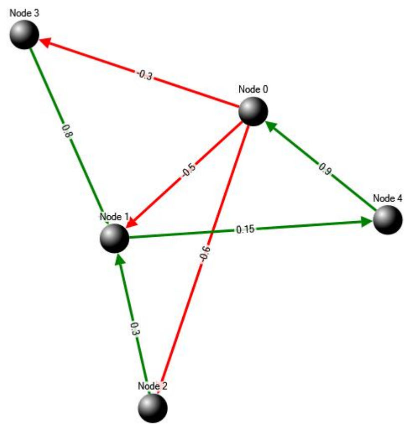
Secondly, where there are multiple types of feedback loops simple modelling in an Excel spreadsheet suggests that the consequences for the states of various events in a network will be less stable and predictable. This presents a much greater challenge for any evaluation, raising questions about what to expect to find in a network of events at a given point in time and/or what sort of time period needs to be the focus of the evaluation. This is the territory covered by Woolcock's (2009) seminal paper on impact trajectories.

Figure 10 below shows how outcomes can vary dramatically over time when events in a network are interconnected. The fictional network has a simple structure, of five nodes connected by three negative (red) and four positive (green) nodes, and where each relationship has a different strength. In the graph above the network diagram, the values of each node are shown changing over time. Their values are dependent on the values of the nodes they are connected to in the previous point in time T-1 but weighted by the value given to those links.

Figure 10: Outcomes over time in a network of inter-connected events



The network structure used as the basis of an Excel simulation:

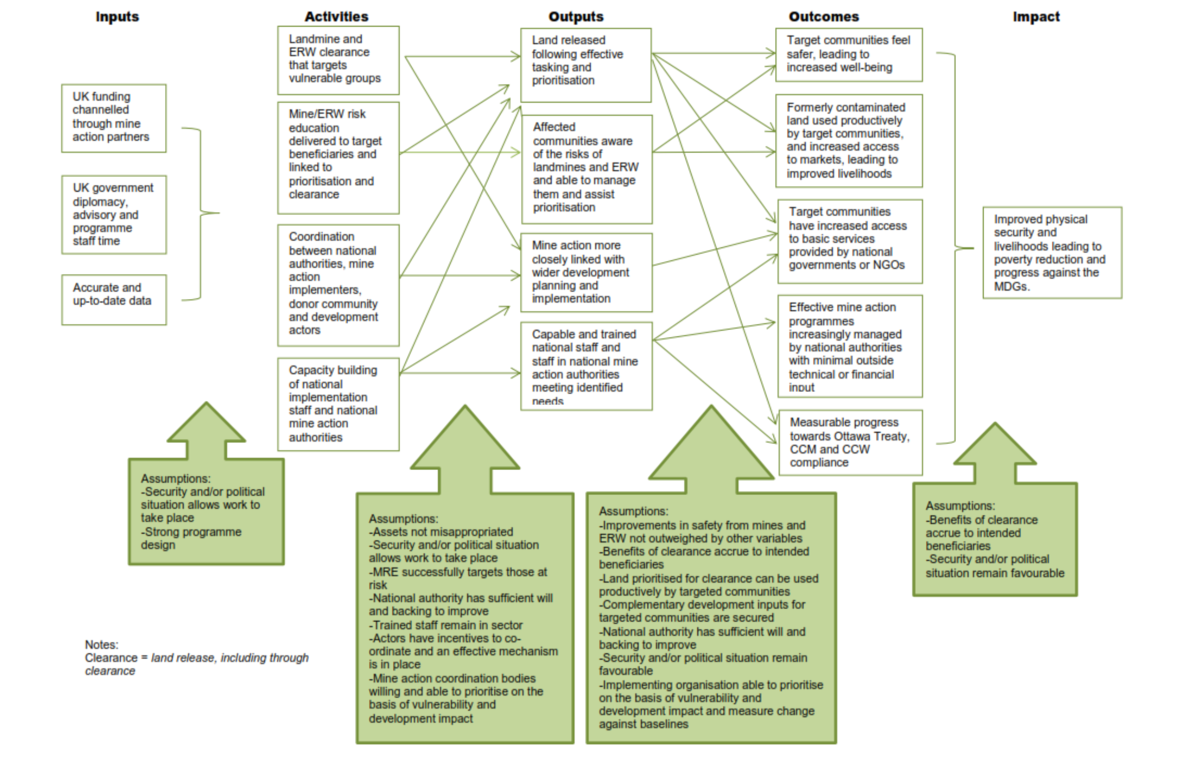


Problem 6: Wider connections

All Theories of Change are circumscribed as well as simplified visions of the world. Some representations are worse than others at acknowledging the wider context. The Google Image search for "Logic Models" suggests that chain models are especially weak. Typically, with wider contextual influences summarised in adjacent boxes with generic titles such as External Factors, Assumptions, or Risks. Figure 11 is one example.

At best these may be accompanied by more specific lists, often describing what could be described as scoping conditions, conditions under which the Theory of Change will work as expected. In some representations, like the post-2011 DFID Log Frame guidance (DFID, 2011), Assumptions are listed in effect as hopes, while the Risks are relegated to a less specific "Risk Rating". In the worst case, "Organizations [i.e. their Theories of Change] imply that change in a society revolves around them and their program, rather than around a range of interrelated contextual factors, of which their program is part." (Valters, quoted by Alford, 2017).

Figure 11: UK Theory of Change for Mine Action



DFID (2013)

One consequence of limited articulation of connections to the wide context is a lack of real-life constraints on the typically optimistic vision at the core of a Theory of Change. Some approaches to programme design have the potential to address this problem, such as context-in (Roche, 1999) or context-centred (Cartwright, 2017) approaches. Network diagrams representing these wider views, especially those populated by actors rather than abstract processes have the potential to show more numerous and more specific connections with the surrounding context. But this will be at the price of increased network complexity, which can accentuate the combinatorial problems of choosing which impact pathways to prioritise during an evaluation.

Section 3

A summary of the problems….

Theory of Change representations frequently fail to adequately describe the expected causal connections in the most basic way i.e. events are left unconnected, or only connected at a macro level by being part of a group of activities. Where connections are made the nature of these linkages is inadequately described. Most often, there is no colour or shape coding or text annotation. Where linkages are described there can often be more causal pathways than are practically evaluable, sometimes astronomically so. Feedback loops are uncommon whereas in reality, these are ever-present, both in dyadic relations between actors and in larger social structures. When feedback loops are present in Theories of Change they convert complicated models into complex models and make evaluation planning more challenging. Theory of Change diagrams partly because they are circumscribed and intentional simplifications tend to have few linkages to the wider surrounding world of other actors who could potentially constrain what is often an optimistic view of what can be achieved. This lack of constraining feedback and wider connections can weaken another aspect of evaluability, which is the plausibility of the Theory of Change working as described.

Section 4

And a word in defence….

Theories of Change are not expected to be perfect and complete the day they are born. For two reasons. One is that theories are models, and models are intentional simplifications that necessarily leave out many features of the real world, they are not supposed to be one-to-one scale mappings. The other is that most programme designs are "works in progress" involving a lot of unknowns and uncertainties, which at best might be reduced over time, as implementation proceeds, and progress is evaluated.

Typically, Theories of Change undergo various iterations at different stages of programme design, then also during implementation and during evaluation. A MandE NEWS (2007) online survey of 99 self-selected users of Logical Frameworks found 63% of respondent's programmes altered their Logical Framework at least once a year. As implementation proceeds, we might expect causal connections and pathways to become more clearly identified and characterised if there is sufficient expectation of this happening.

In the next section, some solutions are proposed for the problems identified so far. Some of these can be applied from the design stage onwards, others are dependent on some progress with implementation, and some may be technically more suited to use by evaluation teams.

Section 5

Six possible ways forward

1. Better descriptions of the connections

Two types can be distinguished. The first are generally applicable categories of causal effects, the second are more customised and specific categories. Both could be identified by appropriate colour or shape coding of links or limited text annotations.

In the philosophy of logic, a distinction is made between necessary and sufficient causes. In a Theory of Change diagram, some connections may be necessary for the occurrence of the event they connect to. Or they may be sufficient, or they may be neither. These possibilities have different consequence for evaluations. Necessary connections are points of vulnerability for a theory, which would normally deserve priority attention. If they fail to work, the theory fails. Sufficient connections are a source of resilience in a theory, and the failure of one will not doom the theory as a whole. All sufficient connections would need to be tested. Coding of which connections in a Theory of Change have necessary or sufficient causal status would be very useful for an evaluation.

The third possibility is that a connection may be neither, but it may be a necessary part of a combination of connections which is sufficient but not necessary for an event to happen (known as INUS[7]). Coding these expected combinations of connections would also aid an evaluation. Testing the workings of such combinations of connections will be more demanding, but a casual reading of recent Qualitative Comparative Analysis studies suggests these are more prevalent than single necessary or sufficient causes (Compasss bibliography, 2016). The importance of such "causal packages" has been highlighted by Cartwright and Hardie (2012) and Mayne (2015).

Mayne has also argued for a different use of logic statements to that described above, where they describe the wider circumstances impinging on a causal pathway. He proposes their use as textual explanations of the causal mechanisms connecting any two events in a Theory of Change. While this is practically possible in a chain model, as per his own example (Mayne, 2015:127) it is not a realistic possibility when working with more complex heterarchical or network models. Other representational devices described below are more useful.

There are other facets of connections that could be recognised by appropriate coding and annotation. One is the expected sequence of any set of inputs connecting into an event. The other is the relative "causal weight" of two or more connections. Two connections might each be sufficient for an outcome to occur, but one may make a bigger difference than the other[8]. Weightings can be assigned to connections to reflect those differences and coded via line thickness; participatory design process as discussed below can be one source of such weightings.

It is likely that the influence of some connections will be variable - a matter of degree, rather than categorical. For example, the impact of different degrees of access to credit, or to training. Where connections are matters of degree there may be value thresholds that govern when an input starts to have an effect, and other thresholds beyond which it has no additional effects (Gough et al, 2018). Here the relationship would be a type of sigmoid function and could define the boundaries of what was a necessary or sufficient input. Or, there may come a point where additional inputs no longer improve outcomes and they start to decline, i.e. a type of parabolic function. Excessive training might be one example of such an input. As Klein (2018) has suggested, a simple graphics menu of possible relationship functions could be developed for use by designers of Theories of Change.

The second class of connection descriptors, mentioned above, are those which are customised to the specific Theory of Change and its users. Degrees of Change's (2014) Theory of Change of the Act Six program is a complicated network of inter-connected sequences of events, presenting the same evaluability problems as discussed earlier. However, the causal connections have been type-coded, distinguishing between those which are "project facilitated" and others which are a "natural effect". Evaluators are likely to be most immediately interested in the first type.

---

[7] Insufficient but Necessary part of a combination that is Unnecessary but Sufficient.

[8] For example, by well exceeding a targeted outcome value

Chris Dunford's (2013) Evidence Project blog goes further. Using a Freedom from Hunger Theory of Change he clarifies the status of various causal connections by combining different forms of type coding: "*Colors represent the confidence we can have in the evidence that certain household- and individual-level impacts do in fact occur in response to participation by women in Credit with Education and Saving for Change programs (in one or more of their variations). The width of an arrow pointing from cause to effect represents the likelihood the impact will occur; that is, the relative frequency of the impact's occurrence. The evidence that an impact actually occurs may be very strong even though the circumstances which create the impact may be relatively rare*". Both distinctions could inform the planning of an evaluation.

The RISE Theory of Change (2008) illustrates a more narrative form of coding of relationships that could also aid evaluability. Linkages are annotated with different text labels e.g. "leads to", "improves", "is foundation for", "informs choice of", etc. While not all these labels have immediate evaluation consequences it is possible to imagine a wider range of text labels that would.

Another option, highlighted by Roche (2018), is for Theories of Change representations to be more explicit about what connections or pathways are more provisional and open to doubt than others. This status can be signalled through confidence codings, as used by Dunford above, or by text annotations of the kind used by DCED below.

All the above proposals raise the real possibility that there may be many more facets of the connections in a Theory of Change that need representation that can actually be captured by shape and colour coding and limited text annotation. One way of resolving this problem has been illustrated by the Donor Committee for Enterprise Development (DCED), who have taken connection annotation to another level of detail altogether. At first glance, the DCED Evidence Framework appears to be another heterarchically structured Theory of Change diagram with similar weaknesses to those above. However, each of the connections between events are "clickable" and take the viewer to a supporting web page, where the evidence for the particular causal link is described in detailed text. This supporting web page can be seen as a nested sub-Theory of Change, where assumptions or knowledge about connecting causal mechanisms can be described in evaluable detail. This is a much more flexible approach to nesting than the use of hierarchically nested Logical Frameworks as proposed in the past (Wiggins and Shields, 2012) and which no longer seems to be in use.

2. Better software for drawing Theory of Change diagrams

If expectations are raised about the amount of detail required in a Theory of Change diagram, especially about the connections, then the ability to clearly visualise these in a usable form becomes more important too. Most Theory of Change diagrams are drawn using some form of software, typically using the draw functions in MS Word, Excel or PowerPoint. The exceptions will be Theories of Change developed through participatory means, which will probably initially use Post-It Notes or the like, then later reproduce these results on a computer.

There are two alternatives which should get more attention. The first is software specifically designed for the representation of Theories of Change. The options here are limited but expanding. They include Do-View (n.d), TOCO (Centre for Theory of Change, n.d), Theory

Maker (n.d) and Changeroo (n.d). DoView and Changeroo appear to be the most widely use and have the most options for describing boxes and arrows and be the most user-friendly. Their features include web and desk-based use, text annotation, colour and shape coding of connections. They also enable the visualisation of nested models. Despite these options, many of the examples shown on all three websites, seem to replicate many of the problems discussed above (grouped connections, unlabelled connections, lack of feedback loops), and chain models seem to the main type in use. Complex models are uncommon.

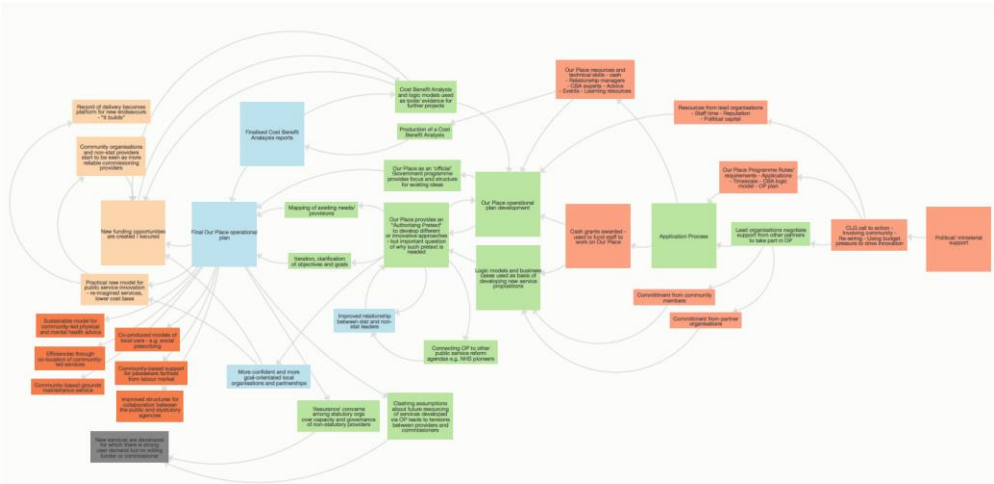
The second alternative is software designed specifically for the representation and analysis of network structures. One example, among the many software packages available[9], is yED and its online equivalent - yWorks (2018). This type of software enables the same text annotation of network connections colour and shape coding and nesting of models as are available in the Theory of Change software above. Network analysis packages also have capacities not available within software designed specifically for representing Theories of Change. When dealing with larger and more complex network structures users can create filtered views of those networks, according to types of connections and/or types of nodes (events) of specific interest. They can use different layout algorithms to highlight clusters of events in a network with similar causal connections. They can also carry out different forms of quantitative network analysis to identify particular kinds of nodes and links, including simple measures which can inform evaluation planning, as discussed below. There are also online versions that can be used collaboratively (yWorks, n.d; KUMU, n.d). In contrast to the software specific to Theory of Change design, there are large communities of users, bodies of theory and research associated with social network analysis software (Kadushin, C. 2013).

3. Basic forms of network analysis

Complicated network models stretch our capacity to analyse their structure by visual inspection only. See for example Figure 9 above or the Figure 12 example of a UK community services programme.

[9] For a review of 30 packages available in 2015, https://www.kdnuggets.com/2015/06/top-30-social-network-analysis-visualization-tools.html

Figure 12: Theory of Change of DCLG "Our Place' programme



Shared Intelligence (2018)

The position of individual events (aka nodes) in such a network can be measured using different "centrality" measures[10], which are all based on the numbers of incoming and outgoing links with other events. In a Theory of Change an event with high "betweenness centrality" will be an event that is part of multiple causal pathways. The failure of that event could have major consequences for the likely success of the Theory of Change. Identifying such events would be useful. The square brown node in Figure 6 is an easily identifiable example.

Desk experiments can be carried out by deleting such events from a Theory of Change and observing the consequences for the connectedness of the other events. Does this leave some events without a cause or cause without an effect? Does it lengthen the causal chain through which some other events have their effects, potentially limiting their effects? The consequences are likely to inform where an evaluation directs its attention within the Theory of Change as a whole.

If connections within a network have known values (e.g. weightings representing the strength of their expected influence) then potentially more important pathways through networks can be identified using simple algorithms that connect adjacent connections according to which has the highest value (e.g. Kruskals or Prims (Wikipedia, n.d)). The result is a "spanning tree", i.e. a tree structure that connects all nodes in a network, without creating any loops. This can be done either by computer or manually in smaller networks. Such a spanning tree could highlight the main channel of expected causal influence in an otherwise complex network representation of a Theory of Change. It would provide a point of focus for evaluation efforts.

[10] https://en.wikipedia.org/wiki/Centrality

4. Participatory network mapping

As seen above in Figure 3, sometimes the problem within a Theory of Change is missing connections rather than lack of information about such connections. One solution is the use of a participatory process using a matrix representation of network structure. In 2009 I worked with GTZ staff in Indonesia to clarify the expected causal relationships between 16 outputs and 11 outcomes of a complex maternal health systems strengthening project operating in multiple districts. A matrix representing the 16 x 9 possible relationships was projected on a screen in a workshop and staff were asked to consider one outcome (column) at a time (Figure 13). They were asked to allocate 100 points across the different outputs (rows), according to the extent to which they expected each of these to contribute to the achievement of the column purpose. This involved a facilitated discussion leading to an agreed allocation. The same process was repeated for each of the other outcomes in the adjacent columns. The weightings allocated for each output, across all outcomes, were then summated, to provide an indication of the relative importance of each output within the project Theory of Change. This process helped rationalise the allocation of time spent by the evaluation team. By focusing on linkages with above-average strength this exercise helped reduce the number of relevant causal links down from 176 to 26, and down further to 17 if attention needed to be confined to the most important outputs. Although not considered at the time, it would also have been useful to ask participants to identify the cells representing relationships which they thought were necessary or sufficient for a given outcome column of interest.

Figure 13: A matrix view of expected output-outcome relationships

	1	2	3	4	5	6	7	8	9	10	11	Output Weight	As %
1.1	10	5	40	10	10				5	40	15	135	12%
1.2	20	5	20	10	5				5		10	75	7%
1.3	5	5	15	5	20				5	10	10	75	7%
1.4	5	5	20	10	5	15	10	5	5	50	10	140	13%
1.7			5		5						5	15	1%
2.1	5	10		20			20		10		10	75	7%
2.2	10	10		10	40		10		5		10	95	9%
2.4	10	10		10	10	10	15	30	10		10	115	10%
3.1	10	20		15	5	20	20	30	10			130	12%
3.2		10		5			10	25	10			60	5%
3.3	5					10	5		0			20	2%
3.4				5		5			5		10	25	2%
3.5													
3.6						15			5			20	2%
4.1	20	20				20	10	10	15		10	105	10%
4.2						5			10			15	1%
	100	100	100	100	100	100	100	100	100	100	100	1100	100%

Rows = outputs, columns= outcomes. Red cells had above average values. IMET (2009:12).

There are also other forms of participatory network mapping that can serve similar purposes. The Net-Map toolkit developed by Eva Schiffer (Campbell et al, 2014) has been used for baseline mapping on influence relationships between stakeholders associated with Community Health Workers in Malawi and then for subsequent evaluation purposes.

5. Predictive modelling

There may be circumstances where stakeholders' expectations about causal connections between large numbers of outputs and one or more outcomes cannot be readily identified in advance, and so it is not possible to narrow down the focus of an evaluation by participatory means. However, if during programme implementation data is being collected on all the outputs delivered and the outcomes being experienced by multiple cases (e.g. individuals, households, villages or other entities) then it is possible to use search algorithms to find the strongest associations between one or more outputs and a specific outcome of interest. When the numbers of relationships with and without the expected outcome are aggregated in a Confusion Matrix[11] it is also possible to identify outputs, or combinations of outputs, that are sufficient and/or necessary for the outcome. These are known as predictive models, and potentially provide important meso-level detail within a larger scale Theory of Change involving other outputs and outcomes. This kind of analysis can be done using free and user-friendly predictive analytics (aka machine learning) software packages, such as BigML (n.d), Rapid Miner Studio (n.d) or EvalC3 - an Excel application (Davies,2018). Evaluation resources can then be used to do carefully selective within-case investigations[12] to identify what kinds of causal mechanisms are at work if any, underneath any association that has been found. Search algorithms in effect provide an initial filtering mechanism that combs through combinatorial space to identify where evaluators can then focus their expensive time and attention.

---

[11] https://en.wikipedia.org/wiki/Confusion_matrix

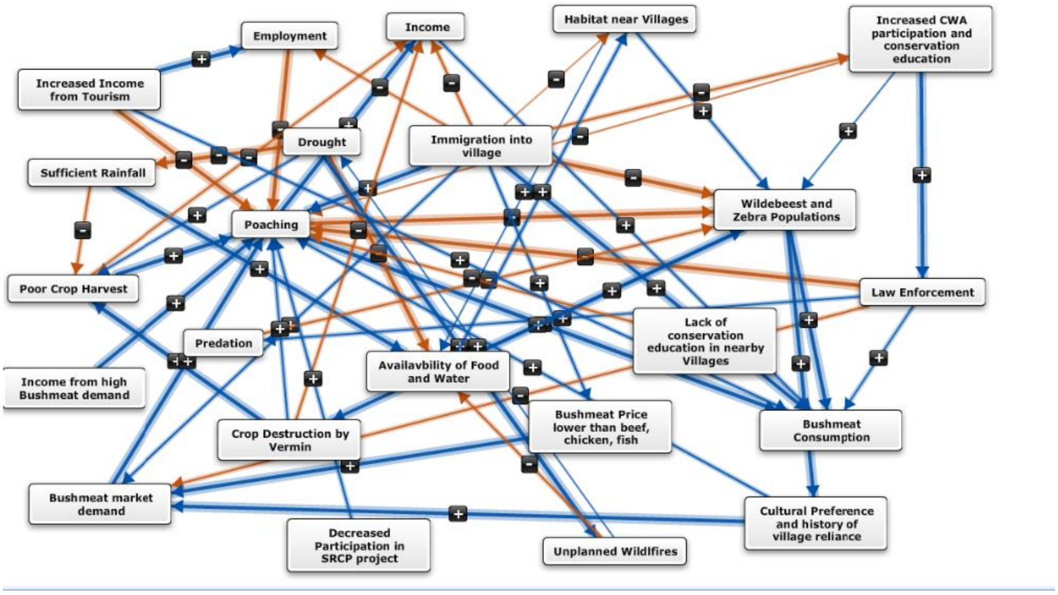
[12] Selective as in informed by a clear case selection strategy e.g. https://evalc3.net/how-it-works/within-case-analysis/

This inductive approach is especially suited to more adaptive programmes with necessarily more "loose" Theories of Change (Davies, 2016b), ones whose component activities, and mixes thereof, may need to change as circumstances change. It can also be used where Theories of Change are more informal and less explicit if combined with ethnographic tools such as card/pile sorting methods (Harloff and Coxon, 2007). Card sorting can be used as a participatory method of generating data on project attributes that may be good predictors of outcomes of interest. Predictive analytics tools can also be useful where *exceptions* to a theory are of more interest. A good model that accurately identifies conditions where outcomes do not occur may still have some False Positives - where outcomes have been achieved despite circumstances which are otherwise conducive to failure. These cases are well worth investigating for their potential replication, using established participatory methods (Positive Deviance Initiative, n.d).

6. Dynamic models

Predictive modelling of the kind described above is a search for stable associations. But if a Theory of Change has a mix of different kinds of feedback loops, then the status of outcomes of interest may vary over time. The timing of evaluations will need to consider expectations of change in the measures of outcomes as generated by a dynamic theory. This will require some form of simulation, since simply eyeballing a network structure, such as Figure 14, will not be sufficient. Ideally, the simulation would be done with user-friendly software where evaluation stakeholders could vary the parameters of the Theory of Change, both the scale of individual events and the strength of their causal connections to each other. Such simulations would help identify the sensitivity of the model parameters and the consequences of design changes, and perhaps even prompt revisions to the Theory of Change prior to an evaluation - if the expected outcomes were no longer plausible.

Figure 14: A dynamic network model of community wildlife management programme



[Gray et al (2015)]

Fortunately, outside the world of evaluation practice and development aid programme design, there has been some innovation in this area. Since the 1980s a body of methods have been developed known as "Fuzzy Cognitive Maps" (FCMs)[13]. These are graphical representations of relationships between concepts, where the connections have both a direction and value. Their structure can be developed through participatory processes engaging model stakeholders, or through expert consultations, or they can be learned from historical data sets. Different software packages have been developed to enable both the development, aggregation and manipulation of these models. Manipulation includes altering model parameters to identify immediate neighbourhood and network-wide consequences. Free software packages include Mental Modeler (n.d), FCMapper (n.d), FCM Expert (Gonzalo N, et al., n.d.), FCM Wizard (Papageorgiou E, n.d.), PC Mapper (n.d) and FSDM (Gregor, 2017).

FCMs are not without their problems. There are debates within the field as to how to interpret model behaviour over time as well as how to appropriately represent causal links in terms of mathematical functions (Gregor, 2017). However, notwithstanding these issues, a wide literature has now accumulated on their use across many fields of applied social and physical sciences (See reviews by Papageorgiou and Salmeron, 2013; Felix et al, 2017).

In parallel to this paper on the representation of Theories of Change, CEDIL has commissioned another paper on the use of structural equation models (SEMs) by Attanasio and Blair (2018). Equations have an advantage over static network diagrams in that the characteristics of the connections can be specified in more detail. On the other hand, they are less "user-friendly" when results need to be communicated to non-specialists. In contrast,

13 https://en.wikipedia.org/wiki/Fuzzy_cognitive_map

FCM software has equations built into their software code but also generate diagrammatic models which display some connection attributes.

There is now a small and emerging literature on the combined use of SEMs and FCMs (De Maio, 2015; Huang et al, 2013; Wang, 2011; Lee, 2009). SEMs are being used to identify the parameters for an FCM, using an existing data set. The FCM is then being used to do what-if analyses, manually testing different input configurations – for example, those corresponding to different management strategies - in order to simulate implications for output variables (De Maio et al, 2015).

Regardless of how they are constructed (i.e. using participatory, expert or data-based analysis) FCMs generate two kinds of predicted behaviour that can and should be evaluated. One is the predicted versed observed results of different what-if scenarios, which may be implemented in different locations within a large-scale programme and evaluated at a given point in time. The other is a predicted trajectory of any of the variables over time. In the short term, this may be a line with one kind of shape or other (e.g. linear, exponential, sigmoid, parabolic, etc). In the longer term the variable values may stabilise at a fixed point, move in a cycle (as in predator-prey numbers) or have no stability (i.e. be chaotic).

Given the above descriptions, the exploration and adaption of FCM and SEM modelling methods clearly represent the more complex end of a spectrum of ways of addressing the representational problems discussed earlier in this paper. For many, any issues associated with these methods may be "a problem we would like to have", given the more common and elementary problems associated with chain, hierarchical and heterarchical models without any feedback loops. When it comes to the representation of Theories of Change it seems that many are still struggling within a complicated rather than complex world. But the existence of dynamic modelling options may help lift expectations.

Section 6

Why so little progress?

In the last five years, there have been seven new publications on the use of Theory of Change in development aid programs. The HIVOS (van Es, 2015) guide is one of the more recently published practical advice that is available. On the first page, it recognises some of the more complex processes of change touched upon above: "*Change emerges as a result of the simultaneous push and pull of multiple political, cultural and social forces involving many individuals and entities. Social change processes are complex and characterised by non-linear feedback loops: our own actions interact with those of others and a myriad of influencing factors. This triggers reactions that cannot be foreseen and makes outcomes of change interventions unpredictable*". In a later section, it is noted that "*Pathways of change, or causal pathways, can be pictured as a series of intermediate changes realised, often called 'results chains', or in the form of a less linear representation, such as a flow chart, web or system map. It is essential to indicate the inter-relations between elements, feedback mechanisms, and how the process is expected to evolve over time, although in real life that will never be linear: think of backlashes and recurrent processes*."

Nevertheless, the nearest thing to a Theory of Change diagrammatic template, reproduced three times in the document, is an hourglass model, where changes flowing upwards through time, initially converging on project objectives then diverging on wider and more distant outcomes, like Figure 5. The hourglass model does have explicit causal linkages between events, both horizontal and vertical. But there are no feedback loops of any kind. Wider contextual factors are indicated but without any causal links being involved. It seems that despite their good intentions the authors are still constrained by a dominant mental model of a Theory of Change that reduces the representation of real-life complexity to a simple linear perspective, where time is the main organising axis.

Why is this so? Practitioners' ability and interest in using the technical options that are available to them is likely to be influenced by their surrounding social and political context, both within their own organisations and other surrounding organisations. One aspect of that context is the nature of demands for different kinds of information, including the demand being expressed for Theories of Change in one form or another. One form of that demand is for *communicable* Theories of Change, which put a premium on simplicity. Simple communicable narratives are needed to convey overall strategic direction within organisations and to key stakeholders outside, to audiences who can't be assumed to have specialist knowledge.

Another form of demand is for *evaluable* Theories of Change, where detail is essential. More detailed theory means a more falsifiable theory. Here I propose a testable meta-theory i.e. a theory about Theories of Change The problems noted in this paper are to be found where there is a lack of demand for evaluable Theories of Change. More specifically, that lack of demand can be seen in the contents of evaluation questions listed in the Terms of Reference for development project evaluations. In my own experience, evaluation questions are more often couched as open-ended questions rather than specific testable hypotheses. While they may refer to the occurrence of specific outcomes or interventions, their phrasings do not include expectations about the particular causal pathways that are involved. In effect these open-ended questions imply either that those posting the questions either know nothing, or they are not willing to put what they think they know on the table as testable propositions. Either way this is bad news, especially if the stakeholders have any form of programme funding or programme management responsibilities. While programme managers are typically accountable for programme implementation it seems they and their donors are not being held accountable for accumulating testable knowledge about how these programmes actually work. Given the decades-old arguments for more adaptive programme management, it's about time this changed (Rondinelli, 1993; DFID, 2018).

Section 7

Implications for CEDIL and DFID

CEDIL's stated mission is to develop and test innovative methods for evaluation and evidence synthesis. Most evaluations are based on some form of Theory of Change, and many of those necessarily involve some form of diagrammatic representation[14]. This paper

[14] The exceptions being varieties of goal free evaluation (Scriven, 1991)

has detailed an endemic set of problems with those representations, all which concern the nature of the connections between events represented in Theories of Change. Clarifying the nature of these pathways is essential if impact evaluations are to shed light on the question of *how* impact is achieved, as well as *if* the impact has been achieved. Six different ways of addressing these problems have been introduced. But more work is needed yet. Four strands of work could be pursued by CEDIL. The first three are largely technical, the fourth is more political.

Firstly, to produce a draft guidance note that could be made available to DFID Evaluation Advisers, with the intention of raising their expectations of what might be expected in a Theory of Change in order to make it more evaluable. Although based on the analyses made in this paper, this note should be informed by a review of a wider set of Theories of Change diagrams, focusing on those found in a sample of DFID funded evaluations. Three guidance components may be needed, addressing different evaluation stakeholders. One is designed to help improve the evaluability of the Theories of Change as initially developed by programme designers. Another to help programme managers to further articulate the details of a Theory of Change in the light of their implementation experience. And a third to help evaluators to do the same if managers have failed to adequately do so. In doing so, more emphasis needs to be placed on programme managers' responsibility to develop testable theories about *what has worked*. They should be accountable for accumulating testable knowledge.

Secondly, good practice examples need to be documented for two of the six possible ways forward: (a) Participatory approaches to the design of evaluable Theories of Change, which provide more details on the connections between events, (b) The use of predictive analytics algorithms to identify ex-post the possible causal connections between events described in a Theory of Change. Useful participatory approaches are likely to already exist but need to be identified and documented. The use of predictive analytics in evaluations is likely to be much rarer. Opportunities to illustrate the usefulness of these methods may be available within upcoming DFID evaluations that will involve CEDIL research teams. In the longer term, the most desirable outcome would be where both approaches were being used by programme managers during implementation, not just evaluation teams.

Thirdly, there needs to be further exploration of ways of developing and using dynamic representations of Theories of Change. This can take two forms. One is to explore the settings in which FCMs and related simulation software have already been used, and the opportunities and constraints associated with those uses to date. The other is to examine the software used to develop FCMs and the like, with a limited number of evaluation criteria in mind: (a) usability by evaluators, (b) flexibility in parameter settings, (c) correspondence between model equations and real-world processes, and (d) adaptability of the underlying code. Subject to the result of this work, a preferred software package would be adapted and supported by guidance on its use, and with appropriate examples of use. Testing of the adapted software could then be carried out in selected DFID funded programmes, via collaborating consultancy firms and selected other CEDIL supported evaluations of DFID programmes

Together, these three strands of work, plus the advice already provided in this paper, should provide a range of solutions that address a range of representational problems, from the most basic (no identifiable connections between events in a theory of change) to the more sophisticated (no means of identifying the iterated consequences of the connections within a theory).

The fourth strand would involve CEDIL investigating and testing the meta-theory proposed in this paper, in the course of evaluations that it becomes involved in. To re-capitulate, that theory argues that the limitations of many Theories of Change noted in this paper arise from *a lack of demand for* testable Theories of Change. More specifically, that this lack of demand is evident in the content of the evaluation questions posed in evaluation Terms of Reference: the dominance of open ended questions and scarcity of testable claims about how interventions are having their effects. As well as investigating this meta-theory CEDIL supported evaluation teams should then test interventions aimed changing the content of evaluation questions, by assisting and challenging programme funders and managers to put their current understandings on the table, ready and available to be tested.

References

Alford C (2017) Two approaches for combining theories of change and system maps. In: In Too Deep. Available at: https://blog.kumu.io/two-approaches-for-combining-theories-of-change-and-system-maps-5b556091c880 (accessed 23 March 2018).

Attanasio, O., Blair, D. and Krutikova, S. (2018) Structural modelling in policy making. CEDIL. Available at: https://cedilprogramme.org/structural-modelling-policy-making/ (accessed 27 March 2018).

Borgatti, Stephen P., Martin G. Everett, and Jeffrey C. Johnson. (2013) *Analyzing Social Networks*. Los Angeles, Calif.: SAGE Publications Ltd.

BigML (n.d) Retrieved from https://bigml.com/

Campbell, N., Schiffer, E., Buxbaum, A., McLean, E., Perry, C., & Sullivan, T. M. (2014). Taking knowledge for health the extra mile: participatory evaluation of a mobile phone intervention for community health workers in Malawi. *Global Health: Science and Practice*, 2(1), 23–34. http://doi.org/10.9745/GHSP-D-13-00141

Cartwright N (2017) Evidence for Policy Prediction: Intervention-centred or Context-centred? In: LSHTM. Available at: https://www.lshtm.ac.uk/newsevents/events/evidence-policy-prediction-intervention-centred-or-context-centred (accessed 23 March 2018).

Cartwright, N., & Hardie, J. (2012). Evidence-based policy: Doing it better. A practical guide to predicting if a policy will work for you. Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:osobl/9780199841608.001.0001

Center for Theory of Change. (n.d.). *TOCO Software | Theory of Change Community*. Retrieved June 25, 2016, from http://www.theoryofchange.org/TOCO-software/

Changeroo (n.d) Retrieved 16 March 2018, from https://www.changeroo.com/

Coleman, G. (1987) The Logical Framework Approach to the Monitoring and Evaluation of Agricultural and Rural Development Projects. *Project Appraisal* 2(4): 251-259.

Compasss bibliography. Retrieved February 1, 2018, from http://www.compasss.org/bibliography/allApp.htm

Davies, R. (2004). Scale, Complexity and the Representation of Theories of Change. *Evaluation*, 10(1), 101–121. http://doi.org/10.1177/1356389004043124

Davies, R. (2005). Scale, Complexity and the Representation of Theories of Change Part II. *Evaluation*, 11(2), 133–149. http://doi.org/10.1177/1356389005055528

Davies, R. (2013, October). *Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations. Report of a Study Commissioned by the Department for International Development*. DFID Working paper 48. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf

Davies, R (2016a) "…and then a miracle happens (or two or three)" Rick on the Road blog. Retrieved from http://mandenews.blogspot.com.au/2016/10/and-then-miracle-happens-or-two-or-three.html

Davies, R. (2016b). *Evaluating the Impact of Flexible Development Interventions Using a 'loose' Theory of Change: Reflections on the Australia-Mekong Ngo Engagement Platform*. ODI Methods Lab. Retrieved from https://www.odi.org/sites/odi.org.uk/files/resource-documents/10361.pdf

Degrees of Change (2014) *Theory of Change*. Retrieved from https://www.degreesofchange.org/approach/theory/

De Maio CD, Botti A, Fenza G, et al. (2015) What-if analysis combining Fuzzy Cognitive Map and Structural Equation Modeling. In: *2015 Conference on Technologies and Applications of Artificial Intelligence (TAAI)*, November 2015, pp. 89–96. DOI: 10.1109/TAAI.2015.7407094

DFID (2011) *Guidance on using the revised Logical Framework*. London. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/253889/using-revised-logical-framework-external.pdf

DFID (2012) Roads in the East in the Democratic Republic of Congo Phase I. Business Case and Summary 203164. Retrieved from http://iati.dfid.gov.uk/iati_documents/3716982.odt

DFID (2013) Clearing a path to development: The UK government's approach to landmines and explosive remnants of war in developing countries (2013) DFID. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/260365/mine-action-policy.pdf

DFID (2017) *Logical Framework 204019, South Sudan Humanitarian Programme (HARISS) 2014 - 2020* [GB-1-204019]. Retrieved http://iati.dfid.gov.uk/iati_documents/5327724.xlsx

DFID (2018) Global Learning for Adaptive Management (GLAM) [GB-1-205148]. Retrieved from https://devtracker.dfid.gov.uk/projects/GB-1-205148

Dhillon L and Vaca S (2018) Refining Theories of Change. Journal of Multi-Disciplinary Evaluation 14(30): 64–87.

DoView – *Visualising Outcomes* (n.d). Retrieved June 25, 2016, from http://doview.com/

Dunford, C. (2013, June 28). *Building a Theory of Change* | Evidence Project. Retrieved from http://microfinanceandworldhunger.org/2013/06/wrap-up-of-the-evidence-project-building-a-theory-of-change/

EDPQS (2018) EDPQS Theory of Change Retrieved from http://prevention-standards.eu/theory-of-change/

Davies R (2018) EvalC3. Cambridge, UK. Available at: https://evalc3.net/ (accessed 21 August 2018).

Gray, S, Gray, S, Kok D, et al. (2015) Using fuzzy cognitive mapping as a participatory approach to analyze change, preferred states, and perceived resilience of social-ecological systems. Ecology and Society 20(2). DOI: 10.5751/ES-07396-200211. Retrieved from https://www.ecologyandsociety.org/vol20/iss2/art11/

Harloff J and Coxon APM (2007) How To Sort; A short guide on sorting investigations. Available at: http://methodofsorting.com/HowToSort1-1_english.pdf (accessed 24 October 2012).

Huang S-C, Lo S-L and Lin Y-C (2013) Application of a fuzzy cognitive map based on a structural equation model for the identification of limitations to the development of wind power. *Energy Policy* 63: 851–861. DOI: 10.1016/j.enpol.2013.09.012.

ITAD (2016) Evaluation of Humanitarian Innovation Evidence Programme (HIEP). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/590702/Evaluation-of-Humanitarian-Innovation-Evidence-Programme-HIEP.pdf

FC Mappers (n.d.). Welcome to FCMappers! Available at: http://www.fcmappers.net/joomla/ (accessed 15 February 2018).

Papageorgiou E (n.d.) FCM Wizard. Available at: http://epapageorgiou.com/fcm-wizard/ (accessed 15 February 2018).

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. John Wiley & Sons.

Félix G, Nápoles G, Falcon R, et al. (2017) A Review on Methods and Software for Fuzzy Cognitive Maps. Artificial Intelligence Review. DOI: 10.1007/s10462-017-9575-1.

Goertz, Gary, and James Mahoney. (2012) *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton University Press, http://www.amazon.co.uk/Tale-Two-Cultures-Qualitative-Quantitative/dp/0691149712/ref=sr_1_1?ie=UTF8&qid=1353850106&sr=8-1.

Gonzalo N, Leon M, Grau I, et al. (n.d) FCM Expert - Experimenting tool for fuzzy cognitive maps. Available at: http://www.fcmexpert.net/ (accessed 15 February 2018).

Gough, D. (2018) *Causal chain analysis in systematic reviews of international development interventions*. Draft paper presented at CEDIL Inception Paper Launch, London, 25th January 2018.

Green, D. (2011*) From Poverty to Power blog postings* (10) Search result retrieved from: http://oxfamblogs.org/fp2p/?s=%22theory+of+change%22

Green, D. (2012, May 14). *Theories of change = logframes on steroids? A discussion with DFID* | From Poverty to Power. Retrieved from http://www.oxfamblogs.org/fp2p/?p=10071

Gregor B (2017) Using Fuzzy Cognitive Maps to Model Policy Issues in the face of Uncertainty and Limited Data. Oregon Systems Analytics. Available at: https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1114&context=trec_seminar (accessed 14 February 2018).

Haiku Analytics (2018) Logic models and complex system mapping for evaluation, performance measurement and planning. Retrieved from http://haikuanalytics.com/dynamic-system-mapping/

IMET (2009) *A review of SISKES (Improvement of the District Health System in Nusa Tenggara Timur (NTT) and Nusa Tenggara Barat (NTB)).* Unpublished

James, C. (2011, September). *Theory of Change Review. A Report Commissioned by Comic Relief*. Retrieved from http://www.theoryofchange.org/pdf/James_Theory_of_Change.pdf

JSI Research & Training Institute Inc. (2016) Theory of Change for the Supply Chain for Community Case Management project. Retrieved from http://1rqxbs47ujl4rdy6q3nzf554.wpengine.netdna-cdn.com/wp-content/uploads/2016/07/Theory-of-Change-Model.pdf

Kadushin, C. (2013) *Understanding Social Networks. Theories, Concepts And Findings*. Oxford University Press. Oxford

Klein, M (2018) Personal communication re Changeroo.com, 09/03/2018

Kumu (n.d.) Available at: https://kumu.io/ [Accessed Date Accessed].

Lee S and Ahn H (2009) Fuzzy cognitive map based on structural equation modeling for the design of controls in business-to-consumer e-commerce web-based systems. *Expert Systems with Applications* 36(7): 10447–10460. DOI: 10.1016/j.eswa.2009.01.070.

Levinton, R (2015) Using theory of change in evaluating community coalitions and collaboratives. Retrieved from http://aea365.org/blog/cp-tig-week-rachel-leventon-on-using-theory-of-change-in-evaluating-community-coalitions-and-collaboratives/

MandE NEWS (2007-17) Responses to the Logical Framework: An Online Survey. Results available at: https://www.surveymonkey.com/results/SM-LQG6XCM58/

Marks P (2015) How creative computers will dream up things we'd never imagine. New Scientist. Available at: https://www.newscientist.com/article/mg22730360-400-how-creative-computers-will-dream-up-things-wed-never-imagine/ (accessed 15 February 2018).

Mayne J (2015) Useful Theory of Change Models. Canadian Journal of Program Evaluation 30(2): 119–142.

Mental Modeler (n.d) What is Mental Modeler. Available at: http://www.mentalmodeler.org

OECD (2010) Glossary of Key Terms in Evaluation and Results Based Management. Retrieved June 24, 2016, from https://www.oecd.org/dac/evaluation/2754804.pdf

Pabari M (2008) Logic beyond the Logical Framework? In: Managing for Impact. Available at: https://mande4mfi.wordpress.com/2008/04/26/logic-beyond-the-logical-framework/ (accessed 21 August 2018).

Papageorgiou EI and Salmeron JL (2013) Review study on Fuzzy Cognitive Maps and their applications during the last decade. IEEE Transactions on Fuzzy Systems 21(1): 66–79.

Positive Deviance Initiative (n.d) https://positivedeviance.org/

Rapid Miner Studio (n.d) Retrieved from https://rapidminer.com/products/studio/

Rihoux, B., & Ragin, C. C. (Eds.). (2009). *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks, Calif. [u.a.]: Sage.

RISE Theory of Change. (2008). Retrieved June 24, 2016, from http://cmap.cinted.ufrgs.br/servlet/SBReadResourceServlet?rid=1JB01RQKZ-4JMHT9-2L7&partName=htmltext

Rondinelli DA (1993) Development projects as policy experiments: an adaptive approach to development administration. Routledge.

Roche, C J R., (1999) Impact Assessment For Development Agencies: Learning To Value Change. Oxfam.

Roche, C J R, (2018) Personal communication, 23/08/2018

Rogers P (2014) *Theory of Change.* Methodological Briefs Impact Evaluation No. 2. UNICEF. Available from: http://devinfolive.info/impact_evaluation/img/downloads/Theory_of_Change_ENG.pdf.

Scriven, M. (1991). Prose and Cons about Goal-Free Evaluation. *American Journal of Evaluation*, 12(1), 55–62. http://doi.org/10.1177/109821409101200108

Shaw, A. (2018) Personal communication, 23/04/2018

Shared Intelligence (2018) Evaluation of DCLG's 'Our Place' programme. Retrieved from https://sharedintelligence.net/our-work/locality/

Stein, D., & Valters, C. (2012). *Understanding Theory in Change in International Development.* The Asia Foundation. Retrieved from http://www2.lse.ac.uk/internationalDevelopment/research/JSRP/downloads/JSRP1.SteinValters.pdf

Sustainable Sanitation and Water Management (2018). Toolbox. Retrieved from http://www.sswm.info/sites/default/files/toolbox/EC%202004%202%20Example%20of%20a%20Objective%20Tree.jpg

Switchback (2018) SWITCHBACK'S THEORY OF CHANGE. Retrieved from https://www.switchback.org.uk/what-we-do/theory-of-change/

Theory Maker (n.d) Retrieved from http://theorymaker.info/

TOCO Software (n.d). Retrieved January 1, 2018, from http://Theory of Changeo.actknowledge.org/aboutus.php

Toulmin, Stephen. (1953) *An Introduction To The Philosophy Of Science*. London.

UNDG (2017) *Theory of Change - UNDAF Companion Guidance*, UNDG. Available from: https://undg.org/wp-content/uploads/2017/06/Theory-of-Change-UNDAF-Companion-Pieces.pdf

Valters C (2015) *Theories of Change: Time for a radical approach to learning in development*. ODI. Available from: https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9835.pdf.

Valters, C. (2014). *Theories of Change in International Development: Communication, Learning or Accountability?* The Asia Foundation. Retrieved from http://www.lse.ac.uk/internationalDevelopment/research/JSRP/downloads/JSRP17.Valters.pdf

Van Es M, Guijt I and Vogel I (2015) *Theory of Change Thinking in Practice*. HIVOS. Available from: http://www.theoryofchange.nl/sites/default/files/resource/hivos_Theory of Change_guidelines_final_nov_2015.pdf.

Vogel I (2012) *Review of the use of 'Theory of Change' in international development*. Review Report. DFID. Available from: http://www.dfid.gov.uk/r4d/pdf/outputs/mis_spc/DFID_Theory of Change_Review_VogelV7.pdf (accessed 11 July 2012).

Vogel I (2012b) ESPA guide to working with Theory of Change for research projects. LTS/ITAD for ESPA. Available at: http://www.espa.ac.uk/files/espa/ESPA-Theory-of-Change-Manual-FINAL.pdf

Wang C-H (2011) Using fuzzy cognitive map and structural equation model for market-oriented hotel and performance. *AFRICAN JOURNAL OF BUSINESS MANAGEMENT* 5(28). DOI: 10.5897/AJBM11.197.

Weiss C (1995) Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families. In: *New Approaches to Evaluating Community Initiatives*, Washington DC: Aspen Institute, pp. 65–92. Available from:

http://www.seachangecop.org/sites/default/files/documents/1995%20Aspen%20New%20Approaches%20to%20Evaluating%20Community%20Initiatives.pdf.

Wiggins S and Shields D (1995) Clarifying the 'logical framework' as a tool for planning and managing development projects. Project Appraisal 10(1): 2–12. DOI: 10.1080/02688867.1995.9726968.

Wikipedia (2018) Minimum spanning tree. Available at: https://en.wikipedia.org/w/index.php?title=Minimum_spanning_tree&oldid=845161708 (accessed 21 August 2018).

Woolcock M (2009) *Towards a Plurality of Methods in Project Evaluation: A Contextualised Approach to Understanding Impact Trajectories and Efficacy*. SSRN Scholarly Paper, Rochester, NY: Social Science Research Network. Available from: http://papers.ssrn.com/abstract=1407981 (accessed 28 June 2016).

yWorks, the diagramming company. (n.d.). Retrieved June 25, 2016, from http://www.yworks.com/yed

Contact us

Centre of Excellence for Development Impact and Learning

London International Development Centre

36 Gordon Square

WC1H 0PD

www.cedilprogramme.org

@CEDIL2018