Nottingham Trent University

08 December 2022

Evaluative Options Review

Jon Hoskin-Campion
Centre for Student and Community Engagement (CenSCE)

Summary & Key Findings:

In attempting to understand impact across a range of more complex interventions there is a commensurate need to adopt, or at least consider, a broader range of evaluative strategies. It can also be noted that evaluation can be used as a means to develop the intervention – an important shift in emphasis that should not be undervalued.

A number of approaches have been considered, and their strengths and weaknesses have been analysed, in relation to several schemes within CenSCE. In particular, **a Realist evaluative approach for the Mentoring scheme and either a Developmental or Participatory approach, to place based evaluation(s), would have considerable merit.**

Contents

Summary & Key Findings: 1
1. Introduction and Context 2
2. Stakeholder Based Approaches 3
2.1 Collective Impact (CI) Approaches 4
2.2 Developmental Evaluation (DE) 6
2.3 Most Significant Change (MSC) 10
2.4 Narrative Storytelling 12
2.5 Participatory Evaluation (PE) 14
2.6 Realist Evaluation (RE) 16
2.7 Ripple Effect Mapping (REM) 19
2.8 'Other' Approaches for Further Consideration 21
3. Commonalities 21
4. Conclusion & Recommendations: Potential Application across CenSCE Programmes 22
Appendix 1: Principles of DE (Patton, 2016 p.256) 23
Appendix 2: The Narrative Interview 25
Bibliography 26

1. Introduction and Context

The aim of this paper is to examine a range of evaluative approaches (seven) and to judge their applicability for use in how CenSCE undertakes programme based evaluation(s). The aim is not to undertake a detailed analytical review of each approach (but references are given for those wishing to expand their knowledge).

In evaluation terms CenSCE adopts a multi-faceted approach to programme[1] evaluation and monitoring. As a basic standard all programmes adopt a Theory of Change (ToC) framework to produce both map and narrative of programme structure and core outcome. As a result, all interventions should at least be commensurate with the OfS Type One evidence standard (OfS, 2019). The process entails close incorporation of, in particular, intervention leads and broader stakeholders within the approach. The balance in the approach favours attempts (and emphasis) to establish *impact* (usually through correlation) but *process* based support to programme improvement is also given.

In addition, we would expect activities to have before (to establish a baseline) and after testing utilising appropriate qualitative and quantitative methods (and validated scales where appropriate). In general too, though perhaps more informally, there is an initial search and exploration for available comparator groups and preferability for the randomised counterfactual. This later stage has invariably proved problematic due to, primarily, the difficulty in altering relationships within established schemes and the small number of participants with some programmes. This can to some extent be mediated though the use of a quasi-experimental design, to create a non-randomised control group, which shares similar characteristics to the treatment group and reduces (although not necessarily eliminates) selection bias.

The above process is not entirely appropriate to programmes where stakeholder involvement and / or the propagation of social change takes place within the modification (or at least greater understanding) of complex systems. It can be explored in greater depth when we examine realist approaches, that causality (Gates and Dyson, 2017) is often understood from a very different perspective to the dominant positivist[2] approach adopted (so far in much of the department's evaluative output).

The following literature review looks at several methodological approaches that could have merit in how particular programmes are evaluated. The report references, in particular, the following sources:

- Impact Evaluation With Small Cohorts: Methodology Guidance (TASO, 2022)

[1] Programme, intervention, scheme and activity are used interchangeably

[2] i.e. *"The basic affirmations of positivism are (1) that all knowledge regarding matters of fact is based on the "positive" data of experience and (2) that beyond the realm of fact is that of pure logic and pure mathematics"* See: https://www.britannica.com/topic/positivism

- Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework (White and Phillips, 2012)

In assessing ease of applicability a narrative and scoring system will be used. The system is based on the following where the present process (described above) scores 5 on a range of 0 (difficult) - 10 (easy).

1. *Ease of understanding (stakeholders).* Throughout CenSCE's evaluation cycle we place considerable emphasis on working **with** a broad range of stakeholders - therefore conceptual clarity is important to encourage involvement.
2. *Ease of Application (practitioners).* The present model (of ToC / narrative / data collection framework / data collection and analysis / data dissemination) is fairly straightforward. However, some models may require the learning of new skills either in framework construction, data collection or analysis.
3. *Degree of Fit.* The scope of this review is predicated on the weaknesses of the present model in light of programme development(s). This item will consider degree of fit in both conceptual (where appropriate) and practical terms.

The reason for developing a scoring system (over the degree of applicability and fit) is to initiate discussion on the literature and how the various models could be used within CenSCE – they are not offered to make fixed conclusions.

2. Stakeholder Based Approaches

The following approaches all place significant emphasis on the role participants play within the research process (commonly as partners rather than subjects). In addition, the use of narrative and causal generation (as opposed to attribution) are often key elements. The following are outlined in alphabetical order and a brief discussion is given with regard to programme applicability within CenSCE (where relevant). In each case the outline and method is summarised, strengths and weaknesses are discussed and utility assessed for potential applicability within CenSCE.

2.1 Collective Impact (CI) Approaches

Outline

'Collective Impact' is *"a network of community members, organizations, and institutions who advance equity by learning together, aligning, and integrating their actions to achieve population and systems level change"* (Collective Impact Forum, 2015). This is achieved through implementing five conditions, aligned to a particular context:

- Establishing a common agenda and shared vision
- Establishing a shared measurement framework
- Foster mutually re-enforcing activities
- Continuous communication
- Development of a strong backbone organisation to coordinate activity (see (Kania and Kramer, 2011)

Evaluation therefore, at least in part and at the initial stage of development, is a consideration as to whether the above five elements exist and their subsequent effectiveness.

Methodology

A number of evaluative principles are central. The role of continuous learning is highlighted in order to shape responses within an ever changing context. The adoption of a common measurement system is also highlighted but perhaps most interestingly, the evaluative process is broken down into stages commensurate with the initiatives' lifecycle:

1. **Initial adoption**: Here the focus is on understanding the context (in order to implement effective solutions) and establishment of the five conditions (see above) and in particular a shared measurement system. The recommended evaluative approach is **Developmental** (see 2.2).
2. **In middle and later years** more formative and summative evaluative processes can take place based on a well-developed joint measurement framework (Preskill et al., 2014)

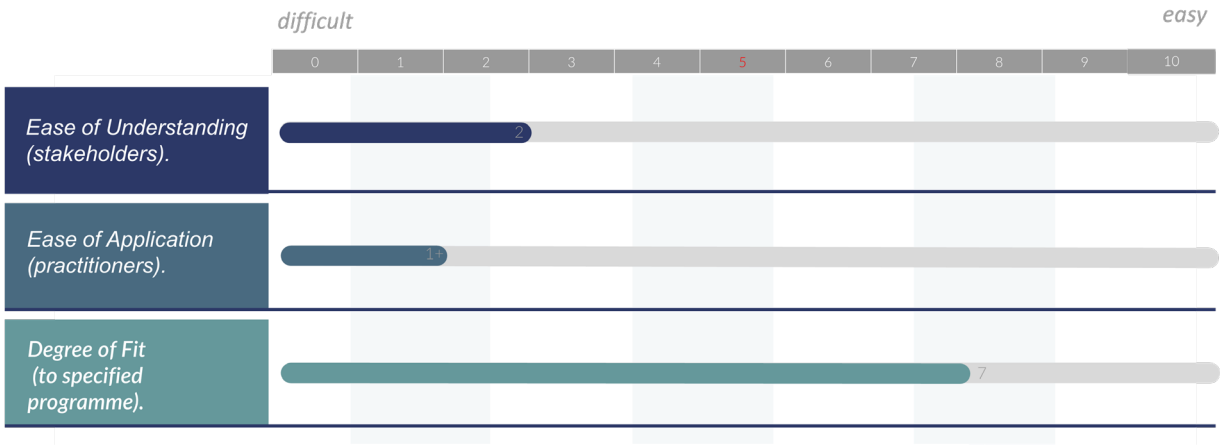
Strengths and Weaknesses

If we separate out the evaluative process from the CI approach it becomes evident that timelines are longer (than in single programme based evaluations to establish impact) **and that *learning* from the evaluative process is *the* critical element** (at least in the early stages) rather than the establishment of 'what works'.

Utility

We can see in Figure 1 (below) that both ease of understanding and application are problematic when compared to a single programme (before and after), ToC based evaluative approach. The evaluation needs to take place in multiple stages and there would be considerable difficulty in establishing a shared measurement framework (both an initiative and evaluative aim). Whilst the methodology is entirely consistent with a collective impact model, schemes developed by CenSCE could be described as CI, systems or participatory in approach (or a combination of all three).

Fig 1: Collective Impact Evaluation



Potential application:

- Bulwell Going Places ? *(systems thinking based approach aimed at tackling educational disadvantage within Bulwell).*
- Getting School Ready ? *(multiple stakeholder programme designed to build the life skills of young people. It aims to empower parents and communities to become key agents of change in supporting children from their early years through to GCSE level in 'getting ready' for each of the three main stages of their educational and work journeys).*

2.2 Developmental Evaluation (DE)

Outline

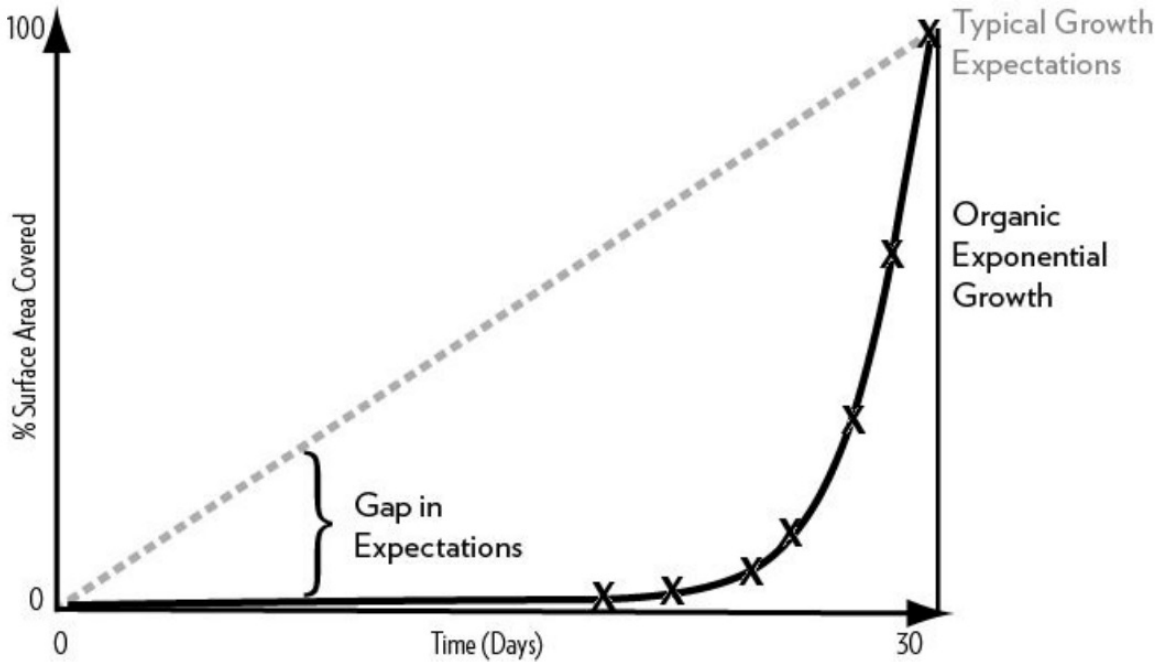
As the name would suggest, developmental evaluation places emphasis upon how the evaluation process can support aspects of programme development (Patton, 2011). The approach can be seen to be particularly relevant where goals are emergent and where the, *"purposes are innovation, change, and learning rather than external accountability (summative evaluation) or getting ready for external accountability (formative evaluation)"* (Patton, 1988 p.318).

This leaves open the possibility of differing forms, of perhaps more summative evaluation types, taking place at a later stage in the programme development cycle. This would be entirely wrong in DE terms as the evaluative need (based on an understanding of programme development) is centred around an ongoing adaptation and development in complex systems. DE is neither a form of summative or process based evaluation but rather has its own particular form.

An initial question for those with an outline for what developmental evaluation can offer, at least for those with some evaluative experience, is how developmental evaluation differs in scope from a 'process' evaluation (see Fox et al., 2016 Chap 4). The major difference between the two approaches is that developmental evaluation supports the process of innovation development especially in the context of complex systems. For a review of 'Systems Thinking' see Stroh (2015).

An important insight when considering change for evaluation purposes (in a complex system) is that change is less likely to occur in a linear fashion and more likely to happen over a longer time frame. **The point being that a differing evaluative perspective needs to be adopted** – see below Figure 1 (*where the example refers to the time taken for a lily to cover a pond in 30 days if the lily doubles in size every day – surprisingly, only on day 29 is the pond half covered*):

Figure 2 (*taken from Stroh, 2015 figure 4.2*)



To re-enforce the above, the *evaluative approach* within a systems model has to be agile enough to capture and explore the non-linear, unpredictable and dynamic aspects of a *systems approach* (Nobles et al., 2022).

Method

Patton is clear that 'situational recognition' needs to take place within every evaluative environment (Patton, 2016, 2011). This means that particular methods have to be considered relevant to the programme in question (though close links between evaluator and other stakeholders is critical). However, the concepts in question; complexity, innovation and emergence are treated, or rather understood, from a qualitative perspective. The issue with fidelity[3] has led Patton to adopt; *"essential principles of developmental evaluation (DE)* [to] *provide high-inference sensitizing guidance that must be interpreted and applied contextually"* (Patton, 2016 p250). The principles are listed in Appendix 1.

Strengths and Weaknesses

The approach emphasises a form of internal rather than external accountability where central questions resonate around emergence. This form of 'utilisation' focused evaluation is primarily intended for the use of social innovators (rather than

[3] *"Fidelity concerns the extent to which a specific evaluation sufficiently incorporates the core characteristics of the overall approach to justify labelling that evaluation by its designated name"* (Patton, 2016 p.250)

external funders concerned with cost application). The approach is particularly appropriate within systems based approaches where:

> *"non- linearity, emergence, dynamical systems, adaptiveness, uncertainty, and coevolutionary processes"* occur. *"Developmental evaluation likewise centres on situational sensitivity, responsiveness, and adaptation, and is an approach to evaluation especially appropriate for situations of high uncertainty where what may and does emerge is relatively unpredictable and uncontrollable"* (Patton, 2011 p7).

Utility (see Fig 3)

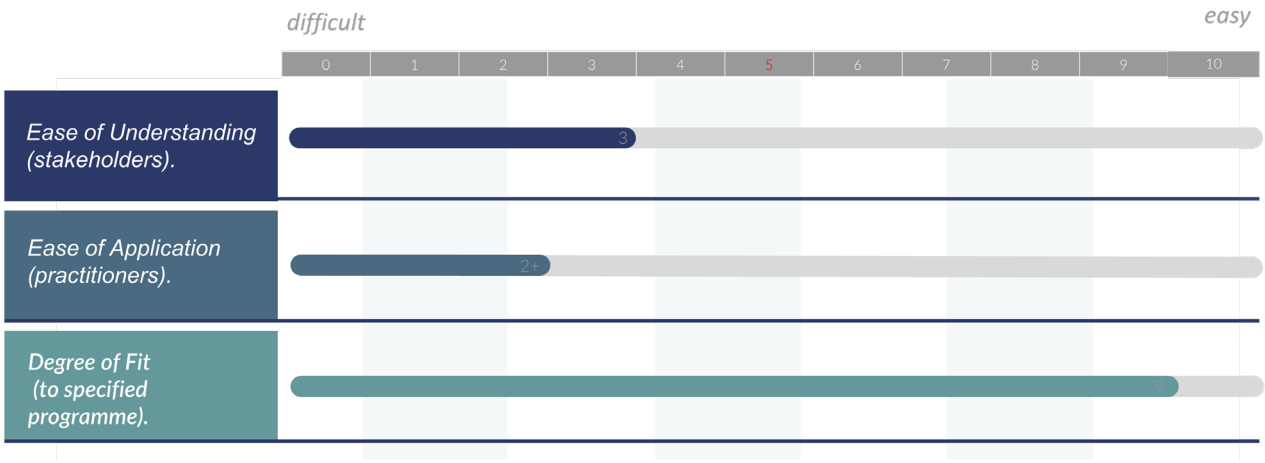
Understanding the approach, or rather the principles involved, could be viewed as relatively complex (see Appendix 1). Indeed, a deeper understanding of the broader systems literature could prove problematic especially when combining the two (evaluative approach and systems thinking) at a conceptual level.

In implementation terms, actioning the 'principles' pose considerable challenges in the establishment of (in particular) trust in order to elicit insight from programme participants. By association, the time and commitment necessary to develop effective stakeholder relationships is considerable. The (sociological) language used, such as 'sensitising concepts' and the problematisation of taken for granted assumptions about what we understand 'context' or 'concept' to actually mean (see Bowen, 2019) places the approach within the realm of fairly advanced qualitative enquiry, as stated by Patton (2016 p36):

> *"The observer moves between the sensitizing concept and the real world of social experience, giving shape and substance to the concept and elaborating the conceptual framework with varied manifestations of the concept. Such an approach recognizes that although the specific manifestations of social phenomena vary by time, space, and circumstance, the sensitizing concept is a container for capturing, holding, and examining these manifestations to better understand patterns and implications".*

We can also note that whilst there is a high degree of symmetry between the approach and CenSCE programmes (listed below) the approach would not fulfil summative questions of 'what works'. Additional monitoring and/or evaluative approaches may also need be necessary.

Fig 3: Developmental Evaluation



Potential application:

- Bulwell Going Places
- Getting School Ready

2.3 Most Significant Change (MSC)

Outline

Basically, MSC is a participatory device for collecting the most significant change stories from the field, as defined by the participant. The story articulates who did what, when and why and why the event was important. The process is heavily structured, to afford rigour, with *"the systematic selection of the most significant of these stories by panels of designated stakeholders or staff"* (Davies and Dart, 2005 p.8). The process utilises a systematic ten step process:

1. How to start and raise interest
2. Defining the domains of change
3. Defining the reporting period
4. Collecting SC stories
5. Selecting the most significant of the stories
6. Feeding back the results of the selection process
7. Verification of stories
8. Quantification
9. Secondary analysis and meta-monitoring
10. Revising the system.

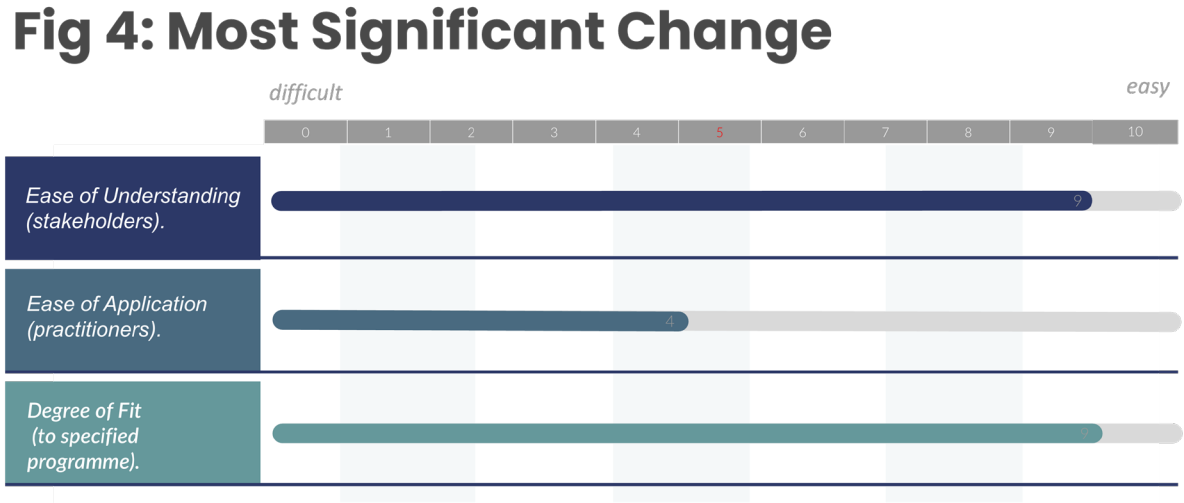
Method

The approach can be used as both a monitoring and evaluation mechanism. However, the MSC technique is best used in combination with other evaluative approaches and not in isolation – to provide a richer understanding through participants stories (Davies and Dart, 2005).

Strengths and Weaknesses

Telling stories is a more engaging and participatory process centred on the subject – so unexpected outcomes can be explored. The approach is particularly relevant when exploring emergent and complex change (though perhaps less relevant in single programme evaluations with defined outcomes). The strength in richness and detail is also a weakness in the potentially large amounts of data (for a broad a range of stakeholders) to assimilate and judge. Any participatory approach requires considerable time and input, to foster the relationships, that ultimately provide insight.

Utility



Potential application:

- Getting School Ready
- Parent Power (Mansfield): *Brilliant Club delivered scheme based on KCL programme*

2.4 Narrative Storytelling

Outline

As the name suggests, narrative storytelling is a qualitative method focusing on understanding a phenomenon or experience. It does not formulate logical or scientific explanations but rather changes the question:

- *'how do I know the truth?*' to…
- *'how do we come to endow experience with meaning?*' (Richard Rorty, philosopher)

As stated by Connelly and Clandinin; *"narrative and life go together and so the principal attraction of narrative as method is its capacity to render life experiences, both personal and social, in relevant and meaningful ways"* (1990 p.10).

Method

Whilst many 'narratives' can be taken from many different sources, narrative interviewing is a particular method distinct from the semi structured interview (see Appendix 2).

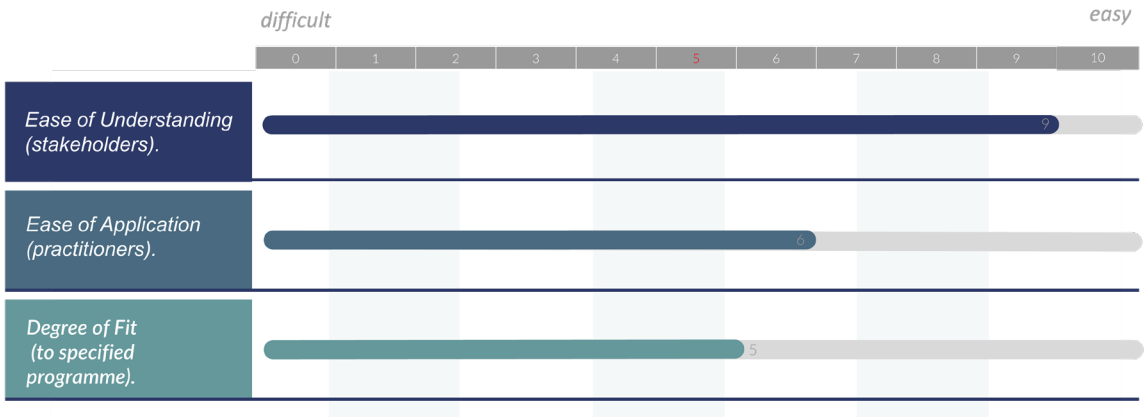
Strengths and Weaknesses

There is a particular power in stories that can reflect human understanding and meaning. However, a narrative relies on criteria other than validity, reliability, and generalizability – we are not striving for causality. The approach is used where we require rich and detailed insight into the 'how' and 'why' - moving beyond statistics.

Utility

Whilst there is little difficulty for participants, the interviewer requires considerable skill, experience and relatability. Whilst the degree of fit is potentially quite high with a range of programmes – there is the likelihood that more concrete (pre-defined) outcomes and outputs would (still) need to be assessed.

Fig 5: Narrative Storytelling



Potential application:
START Project *(drama based transition support scheme based in Mansfield)*

2.5 Participatory Evaluation (PE)

Outline

The key focus within a PE is the active *subject* and their involvement within all stages of the evaluative process. As with participatory types of research, utility is not judged only on the results but also through how the process is able to empower the subject (Plottu and Plottu, 2011). There are several differing forms of PE each with variations in emphasis (Pollitt, 1999). There are also differing forms of participatory research (i.e. Participatory Action Research (PAR) see Funk et al., (2016), Participatory Action Learning (PAL) see Zuber-Skerritt (2018) and Participatory Action Learning and Action Research (PALAR) (Wood and Zuber-Skerrit, 2013). We can note that the particular difference between PAR and Action Research (AR) is that PAR goes beyond understanding and actively seeks to support change. Participatory Appraisal is also of utility when considering community consultation (NESEP, 2014) and Participatory Systems Mapping could be very useful when mapping out complex systems through a range of stakeholders (Penn and Barbrook-Johnson, 2019). Participatory evaluation;

> *"like participatory research, is not just a matter of using participatory techniques within a conventional setting. It involves radically rethinking who initiates and undertakes the process, who learns or benefits from the process and how the causal consequences of a process are recognized: who decides what findings matter"* (Abma et al., 2019 p216).

The below will draw out some key principles.

Method

The main question centres around whose voices to include or exclude rather than on particular methods, per se (Abma et al., 2019). In effect, methods choice comes from community participants themselves in the context of what is expedient within the programme and particular context. It should be noted that when evaluating participatory research, by implication, questions concerning the following need to be applied:

- to capture what enables participants to be aware of what drives their own thinking and acting
- how their own thinking and the thinking within their setting has affected change
- how changes in thinking have translated into changes in practice
- the process must be participatory, dialogical, embedded and dynamic with critical reflection at its heart (Abma et al., 2019)

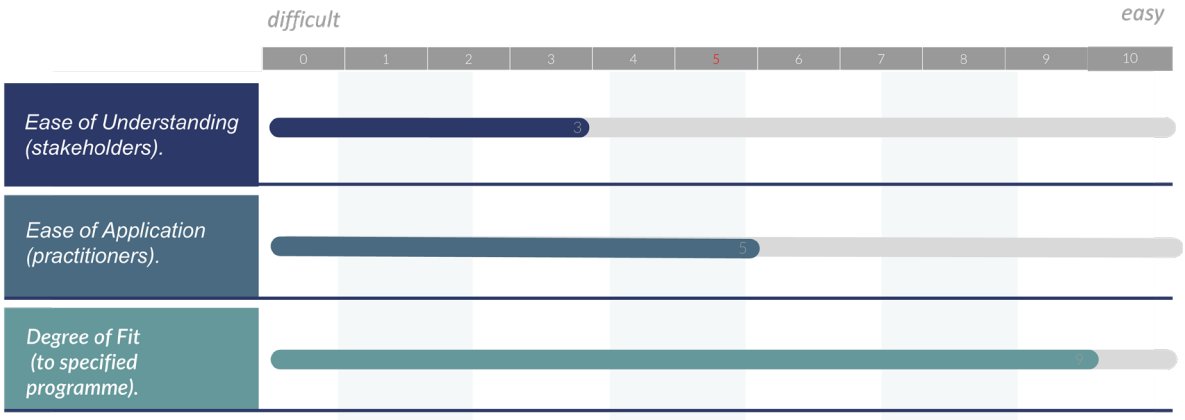
Strengths and Weaknesses

It is perfectly possible to use participatory techniques without necessarily employing a holistic approach to participatory evaluation, as not all the underlaying values of agency, democracy and/or collaboration need be applied. A particular challenge within the participatory evaluative process is to ensure a trustworthy *"evaluation rather than creating a particular pathway that favours one way of thinking"* (Abma et al., 2019 p.217)

Utility

The approach has to match the intervention type. A process based evaluation might be perfectly feasible when considering the evaluation of a participatory appraisal (consultation) type process. Consideration also needs to be given as to whether the evaluative approach needs to be considered as a change agent within its own right – or whether (mere) participatory techniques could be effectively employed.

Fig 6: Participatory Evaluation



Potential application:

- Getting School Ready
- Parent Power (Mansfield)

2.6 Realist Evaluation (RE)

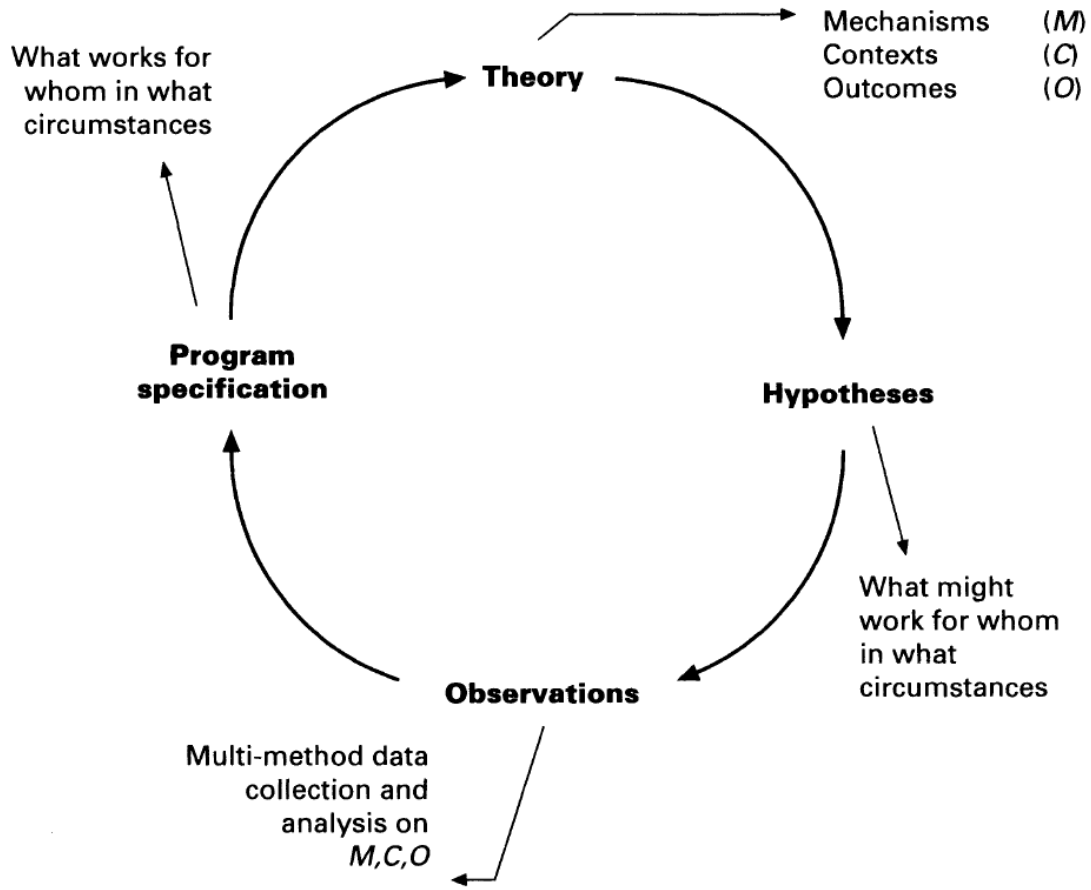
Outline

Realist evaluation is founded on a realist philosophy where the *understanding* of a phenomenon is a central endeavour – and not simply the acquisition of *facts and knowledge* about a phenomenon (which we can term empiricism) (Whewell, 2007). Scientific realists understand the world as separate from and independent of one's perception of it (Pawson, 2002). The goal, in evaluation terms, is to examine regular patterns that exist within the social world and to offer greater understanding of these regularities through in depth exploration of generative causal mechanisms (Salter and Kothari, 2014).

Method

The central aim in adopting a realist approach, across a range of differing interpretations and emphasis, is to develop an understanding of causal mechanisms where the *"mechanisms are underlying causes of the changes that may not be possible to observe but that we can use as indicators in looking for patterns in the data that allow an explanation to be ruled out or supported as consistent with the data."* (Henry, 2004 p.360). Through a process of hypothesis generation (similar to a ToC approach) a series of Context-Mechanism-Outcome (CMO) configurations are developed in order to test out how the mechanisms work in particular contexts to produce outcomes for particular groups of people. In summary, to understand; *"what works for whom in which circumstances…and why"* (Pawson and Tilley, 1997 p.77). The essential elements within the RE cycle are set out (Figure 7) by Pawson and Tilley (1997 p.85):

Figure 7 – RE Cycle



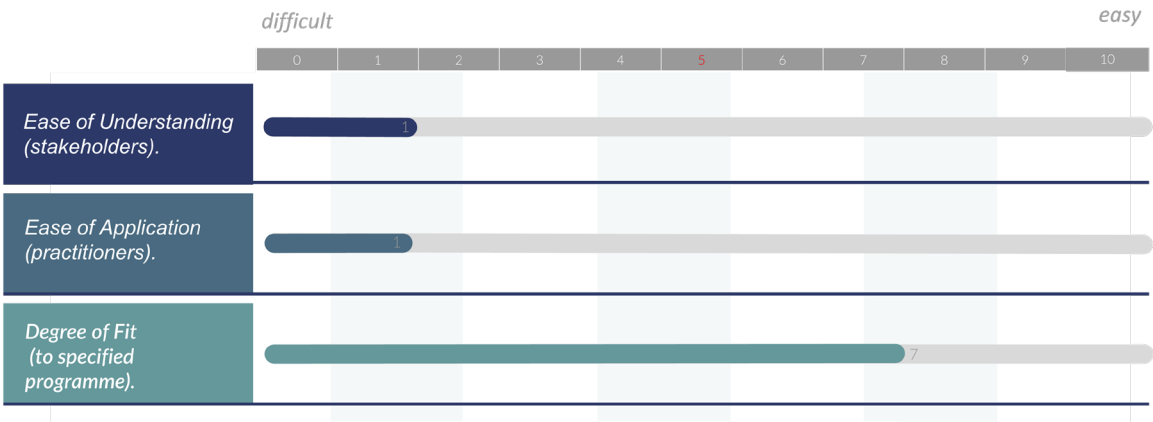
Strengths and Weaknesses

In comparison with some other approaches, the literature in the subject area becomes dense and difficult very quickly. It is easy to become involved in broader philosophical, ontological and epistemological debates. However, the approach is particularly appropriate in understanding the causes (or mechanisms) of observed regularities i.e. *"to examine not only outcome, but also the possible causes and contextual factors associated with change"* (Salter and Kothari, 2014 p.2)

Utility

The approach could have potential application within the CERT mentor scheme due to the (presumed) complex interplay between the various elements within the scheme – where the understanding of 'what works' or rather what components within the scheme drive change, is problematic. However, understanding the approach, its application and the political issues related to the emphasis on causal generation rather than attribution mean implementation is not straightforward.

Fig 8: Realist Evaluation



Potential application:

- CERT Mentee / Mentor *(Support for first-year students to enhance a sense of belonging and settle into university life (see here)*
- Black Leadership Programme (A *programme designed to empower second year undergraduate students of black heritage to gain confidence and become resilient leaders through a series of interactive events and community-building activities throughout the academic year (see here)*

2.7 Ripple Effect Mapping (REM)

Outline

As with DE, because we are dealing with complex problems inside complex systems, the evaluative response is inherently problematic due to the complex interplay of an any number of factors. The main emphasis within REM is to capture the wider effects, and adaptive nature, of a systems based programme (Nobles et al., 2022).

Method

One approach is to gather, through a series of workshops, visual maps set against a timeline to understand system change effects (Nobles et al., 2022). The underlaying principles have been outlined by Chazdon et al., (2017):

- Use of **appreciative inquiry**
- Adoption of a **participatory approach**,
- **interactive group interviewing** and reflection,
- use of **mind mapping** to visualise the impacts

Strengths and Weaknesses

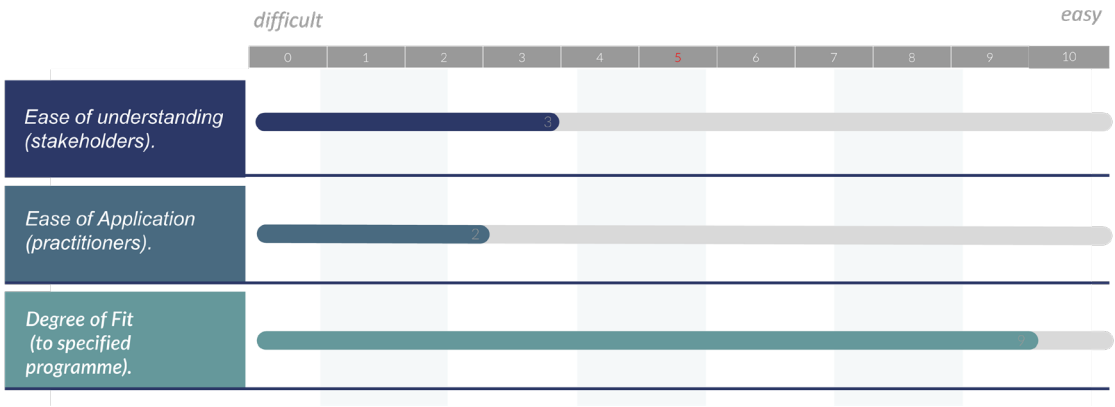
Even in situations where a systems based approach has failed (in some sense) to deliver change relevant to the original purpose, it is not difficult to imagine a broader range of 'effects' that might emanate from a systems approach – therefore the process could (also) point to how the respective intervention could adapt into areas where it is working (or perceived to be). However, if the primary aim(s) has not been met it is unlikely that a broad range of scheme participants would involve themselves i.e. do you need a stone in the water to produce a ripple?

The emphasis again is on contribution rather than attribution in a causal sense. This in an important point with political implications (Abma et al., 2019) – are those funding the scheme aligned with this type of evaluative output?

Utility (see Fig 9)

The process has a high degree of applicability for use in systems thinking approaches (see above in DE utility comments) – in order to find out broader systems impacts than those originally considered. Whilst the sessions and workshops are not problematic, REM analysis and the identification of 'impact pathways' need to be carefully considered with regard to objectivity.

Fig 9: Ripple Effect Mapping



Potential application:

- Bulwell Going Places

2.8 'Other' Approaches for Further Consideration

The aforementioned only gives an outline to a range of possible methodologies. Other frameworks might also be applicable to, broadly speaking, small n or stakeholder centred evaluations. These include:

- Qualitative Comparative Analysis (QCA)
- Process Tracing
- Outcome Harvesting

In addition, a range of quasi-experimental statistical methodologies might also be applicable to CenSCE programmes – utilising, in part, a highly developed monitoring programme. These include:

- Regression Discontinuity
- Difference-in-Difference
- Propensity Score Matching
- Case-Control Matching

Future work could consider the above (or their increased application). However, whilst perhaps inevitably a subjective judgement, the most applicable in terms of manpower (and existing skill base) and most importantly, methodology match to programme, has been offered above.

3. Commonalities

A number of commonalities become distinct when considering the above:

- All could (potentially) benefit from theoretical underpinning within a (highly participatory) **ToC approach** (i.e. for ToC and Realist approaches see Blamey and Mackenzie (2007) and for ToC and Systems approaches see Wilkinson et al., (2021)**.** For participatory based approaches a ToC framework could be of use once outcomes have been established – after the project initiation phase (which could be lengthy).
- **Causality**, when considered, is understood in generative terms. Some of the approaches do not make evaluative claims at all (i.e. storytelling) but rather afford a different perspective on experience
- Several of the approaches are **employed ideally in conjunction** with other approaches and / or methods (i.e. MSC)
- No approach could be considered an **'easy'** option in its application or delivery
- Most of the approaches employ a rigour and/or formalised process to **qualitative** inquiry (though some employ a range of methods i.e. realist)
- There is considerable underlaying symmetry between **Developmental and Participatory approaches** (i.e. when considering PA, Abma et al., contend there is a need for the, *"construction of concepts of quality that capture the*

changing and multiple perspectives of program development over time" (2019 p218)

- Several approaches emphasise the importance of **context**

4. Conclusion & Recommendations: Potential Application across CenSCE Programmes

There is no simple approach to either understanding or fostering the development of social change. All of the methodologies summarised offer potential solutions but all have inherent strengths and weaknesses. It is therefore critical to properly assess both scheme and evaluation purpose for ***all*** stakeholders.

All schemes and all evaluative processes would benefit from a research informed approach (and potentially a ToC process – but not a logic model approach that emphasises linear relationships). A ToC approach also has less utility when the outcomes are emergent – as with many of the approaches outlined here. However, the respective process would benefit, in all instances, from close stakeholder co-operation and involvement. The approaches mentioned all have, to some degree, differences in emphasis and fluidity in how they are applied within different contexts – they are not straight jackets and should be interpreted and applied as guidelines.

Following this, several of these methodologies could be used in the evaluation of CenSCE programmes – with the large caveat concerning resource implications as some approaches (DE / Realist in particular) are far more time consuming and intensive than the present model.

Appendix 1: Principles of DE (Patton, 2016 p.256)

Table 1. Eight Essential Developmental Evaluation (DE) Principles.

Essential DE principles	What to look for to assess the *degree of manifest sensitivity and sensibility* in DE practice, from design to use of findings	Examples of contextual evidence of the essential DE element being incorporated in practice
1. **Developmental principle**: illuminate, inform, and support what is being developed, by identifying the nature and patterns of *development* (innovation, adaptation, and systems change) and the implications and consequences of those patterns	Something (the innovation) is being *developed*. The evaluation tracks what is being developed and the implications of what emerges. The evaluation itself is developed (emergent design) as the innovation develops	The evaluation's purpose, supporting *development* and adaptation of the innovation, is explicit and that focus is maintained throughout. The evaluation design's emergence and adaptations are documented and their implications discussed
2. **Evaluation rigor principle**: Ask probing evaluation questions; think and engage evaluatively; question assumptions; apply evaluation logic; use appropriate methods; and stay empirically grounded—that is, rigorously gather, interpret, and report data	DE is *empirically driven*, and *evaluative thinking* undergirds all aspects of the engagement	Data are gathered, reported, and interpreted about the implications of what is being developed; DE findings and feedback inform next steps in the adaptive process
3. **Utilization-focused principle**: Focus on intended use by intended users from beginning to end, facilitating the evaluation process to ensure utility and actual use	Intended use by intended users focuses the evaluation	Social innovators and their supporters are the primary intended users of DE and clearly identified as such. The explicit purpose of the evaluation is to support the development and adaptation of the innovation (vs. improvement, accountability, or summative judgment)
4. **Innovation niche principle**: Elucidate how the change processes and results being evaluated involve innovation and adaptation, the niche of developmental evaluation	A commitment to innovate is explicit and authentic: a fresh and effective response to an intractable social challenge or problem or to an emergent one	DE has helped the social innovation develop and adapt within the context where the innovation is occurring
5. **Complexity perspective principle**: Understand and interpret development through the lens of complexity and conduct the evaluation accordingly. This means using complexity premises and dynamics to make sense of the problems being addressed; to guide innovation, adaptation, and systems change strategies; to interpret what is developed; to adapt the evaluation design as needed; and to analyze emergent findings	The characteristics of the *complex dynamic system* in which innovation and evaluation are occurring are described. The complexity characteristics of the innovation being developed and evaluated are also described. The DE design, process, and outcomes reflect these complexity characteristics	The nature and degree of uncertainty, turbulence, nonlinear interactions, and dynamical patterns are highlighted. DE is explicitly aligned with the complexity of the innovation. Sensitivity to and implications of emergence, adaptation, and context are manifest

(continued)

Table 1. (continued)

Essential DE principles	What to look for to assess the *degree of manifest sensitivity and sensibility* in DE practice, from design to use of findings	Examples of contextual evidence of the essential DE element being incorporated in practice
6. **Systems thinking principle**: Think systemically throughout, being attentive to interrelationships, perspectives, boundaries, and other key aspects of the social system and context within which the innovation is being developed and the evaluation is being conducted	Attention to interrelationships, perspectives, and boundaries undergirds and informs both the innovation processes and the developmental evaluation	The design, data collected, findings presented, and use of findings demonstrate systems understandings and systems thinking. Contextual sensitivity is explicit and evident throughout the evaluation
7. **Cocreation principle**: Develop the innovation and evaluation together—interwoven, interdependent, iterative, and cocreated—such that the developmental evaluation becomes part of the change process	The developmental evaluator is close enough to the action to build a mutually trusting relationship with the social innovators. The collaborative process is active, reactive, interactive, and adaptive	In the process of collaboration, adaptations and developments are *cocreated*. DE becomes part of the intervention (cocreation). How this occurs and with what implications and consequences are discussed
8. **Timely feedback principle**: Time feedback to inform ongoing adaptation as needs, findings, and insights emerge, rather than only at predetermined times (e.g., quarterly or at midterm and the end of project)	Feedback of findings is timely and ongoing (not just delivered at predetermined times, like quarterly, or midterm and the end of project)	Evidence is reported about how the DE feedback was engaged, useful, and used in close conjunction with real-time decision-making and adaptations

Appendix 2: The Narrative Interview (Jovchelovitch and Bauer, 2000)

The elicitation technique

Table 4.1 Basic phases of the narrative interview

Phases	Rules
Preparation	Exploring the field
	Formulating exmanent questions
1 Initiation	Formulating initial topic for narration
	Using visual aids
2 Main narration	No interruptions
	Only non-verbal encouragement to continue story-telling
	Wait for the coda
3 Questioning phase	Only 'What happend then?'
	No opinion and attitude questions
	No arguing on contradictions
	No why-questions
	Exmanent into immanent questions
4 Concluding talk	Stop recording
	Why-questions allowed
	Memory protocol immediately after interview

Bibliography

Abma, T., Banks, S., Cook, T., Dias, S., Madsen, W., Springett, J. and Wright, M.T. (2019) *Participatory Research for Health and Social Well-Being*, Springer.

Blamey, A. and Mackenzie, M. (2007) 'Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges?', *Evaluation*, 13, 4: 439–455.

Bowen, G. (2019) *Sensitizing Concepts*, available at: https://www.academia.edu/41101347/Sensitizing_Concepts (accessed 3 December 2022).

Chazdon, S., Emery, M., Hansen, D., Higgins, L. and Sero, R. (2017) *A Field guide to ripple effects mapping*, Scott Chazdon, Mary Emery, Debra Hansen, Lorie Higgins, & Rebecca Sero (eds.), Minneapolis, Minnesota: University of Minnesota Libraries Publishing.

Collective Impact Forum (2015) *What Is Collective Impact?*, available at: http://www.fsg.org/OurApproach/WhatIsCollectiveImpact.aspx (accessed 11 November 2022).

Connelly, F.M. and Clandinin, D.J. (1990) 'Stories of Experience and Narrative Inquiry', *Educational Researcher*, 19, 5: 2–14.

Davies, R. and Dart, J. (2005) *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*, available at: https://www.mande.co.uk/wp-content/uploads/2005/MSCGuide.pdf (accessed 26 July 2022).

Fox, C., Grimm, R. and Caldeira, R. (2016) *An Introduction to Evaluation*, London: Sage.

Funk, J., Worthington, S. and Price-Winter, L. (2016) *Using Evaluation and Participatory Action Research Methods to Develop an eLearning Resource*, available at: https://methods.sagepub.com/.

Gates, E. and Dyson, L. (2017) 'Implications of the Changing Conversation About Causality for Evaluators', *American Journal of Evaluation*, 38, 1: 29–46.

Henry, G.T. (2004) 'Realist Evaluation', in *Encyclopedia of Evaluation*, 359–362.

Jovchelovitch, S. and Bauer, M.W. (2000) 'Narrative interviewing', in *Qualitative researching with text, image and sound: a practical handbook*, London: SAGE Publications.

Kania, J. and Kramer, M. (2011) 'Collective Impact', *Stanford Social Innovation Review*, Winter: 36–41.

NESEP (2014) *Introduction to the Principles of Participatory Appraisal*, available at: https://www.wansbeckcvs.org.uk/files/NESEP_Participatory_Appraisal_Handbook_2014.pdf (accessed 1 July 2019).

Nobles, J., Wheeler, J., Dunleavy-Harris, K., Holmes, R., Inman-Ward, A., Potts, A.,

Hall, J., Redwood, S., Jago, R. and Foster, C. (2022) 'Ripple effects mapping: capturing the wider impacts of systems change efforts in public health', *BMC Medical Research Methodology*, 22, 72: 1–14.

OfS (2019) *Using standards of evidence to evaluate impact of outreach*, available at: https://www.officeforstudents.org.uk/media/f2424bc6-38d5-446c-881e-f4f54b73c2bc/using-standards-of-evidence-to-evaluate-impact-of-outreach.pdf (accessed 1 March 2019).

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

Patton, M.Q. (1994) 'Developmental Evaluation', *Evaluation Practice*, 15, 3: 311–319.

Patton, M.Q. (2016) 'What is Essential in Developmental Evaluation? On Integrity, Fidelity, Adultery, Abstinence, Impotence, Long-Term Commitment, Integrity, and Sensitivity in Implementing Evaluation Models', *American Journal of Evaluation*, 37, 2: 250–265.

Pawson, R. (2002) 'Evidence-based Policy: The Promise of `Realist Synthesis'', *Evaluation*, 8, 3: 340–358.

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

Penn, A. and Barbrook-Johnson, P. (2019) *Participatory Systems Mapping: a practical guide*, 1–4.

Plottu, B. and Plottu, E. (2011) 'Participatory Evaluation: The Virtues for Public Governance, the Constraints on Implementation', *Group Decision and Negotiation*, 20, 6: .

Pollitt, C. (1999) 'Stunted by stakeholders? Limits to collaborative evaluation', *Public Policy Adm*, 14, 2: 77–90.

Preskill, H., Parkhurst, M. and Splansky Juster, J. (2014) *Guide to Evaluating Collective Impact 01: Learning and Evaluation in the Collective Impact Context*, available at: https://www.fsg.org/resource/guide-evaluating-collective-impact/ (accessed 13 December 2022).

Salter, K.L. and Kothari, A. (2014) 'Using realist evaluation to open the black box of knowledge translation: A state-of-the-art review', *Implementation Science*, 9, 1: 1–14.

Stroh, D.P. (2015) *Systems Thinking for Social Change*, Vermont, USA: Chelsea Green.

TASO (2022) *Impact Evaluation With Small Cohorts: Methodology Guidance*, available at: https://s33320.pcdn.co/wp-content/uploads/TASO_-Report_Impact-evaluation-with-small-cohorts_methodology-guidance_Secured-1.pdf (accessed 2 July 2022).

Whewell, W. (2007) *Empiricism Vs. Realism Revisited*, available at: https://link.springer.com/chapter/10.1007/978-1-4020-3838-9_9 (accessed 12 January 2021).

White, H. and Phillips, D. (2012) *Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework*, available at: https://www.3ieimpact.org/evidence-hub/publications/working-papers/addressing-attribution-cause-and-effect-small-n-impact (accessed 3 November 2018).

Wilkinson, H., Hills, D., Penn, A. and Barbrook-Johnson, P. (2021) 'Building a system-based Theory of Change using Participatory Systems Mapping', *Evaluation*, 27, 1: 80–101.

Wood, L. and Zuber-Skerrit, O. (2013) 'PALAR as a methodology for community engagement by faculties of education', *South African Journal of Education*, 33, 4: 1–15.

Zuber-Skerritt, O. (2018) 'An educational framework for participatory action learning and action research (PALAR)', *Educational Action Research*, 26, 4: 513–532.