

Thinking with complexity in evaluation: A case study review

Evaluation Journal of Australasia
1–17

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1035719X211008263

journals.sagepub.com/home/evj

Chris Roche

Institute for Human Security, La Trobe University

Graham Brown

Centre for Social Impact – University of NSW

Australian Research Centre in Sex, Health and Society – La Trobe University

Samantha Clune

Australian Institute of Primary Care and Ageing, La Trobe University

Nora Shields

College of Science, Health and Engineering, La Trobe University

Virginia Lewis

Australian Institute of Primary Care and Ageing, La Trobe University

Abstract

Adopting complexity thinking in the design, implementation and evaluation of health and social development programmes is of increasing interest. Understanding institutional contexts in which these programmes are located directly influences shaping and eventual uptake of evaluations and relevant findings. A nuanced appreciation of the relationship between complexity, institutional arrangements and evaluation theory and practice provides an opportunity to optimise both programme design and eventual success. However, the application of complexity and systems thinking

Corresponding author:

Samantha Clune, Australian Institute for Primary Care & Ageing, La Trobe University, Melbourne, VIC 3086 Australia.

Email: s.clune@latrobe.edu.au

within programme design and evaluation is variously understood. Some understand complexity as the multiple constituent aspects within a system, while others take a more sociological approach, understanding interactions between beliefs, ideas and systems as mechanisms of change. This article adopts an exploratory approach to examine complexity thinking in the relational, recursive interactions between context and project design, implementation and evaluation. In doing so, common terms will be used to demonstrate the nature of shared aspects of complexity across apparently different projects.

Keywords

complexity, evaluation, health service programme, project design, social development programme

Introduction

The linked processes of programme design, implementation and evaluation are integral parts of service delivery for health service and social development projects (Pawson & Tilley, 1997). The intricacies of health and social development projects also require engagement with the project components, context and people (Sheikh et al., 2011). Every project is embedded in social and political realities, influenced by cultural frames, which contribute to project outcomes and related success (Sheikh et al., 2011). This inherent complexity has led to some describing health and social development interventions as complex adaptive systems (Braithwaite et al., 2017), that is, systems with ‘a large number of mutually interactive parts, often open to the environment, which self-organise their internal structure and their dynamics with novel and sometimes surprising macroscopic (emergent) properties’ (Ramalingam, 2013, p. 141). There have been increasing calls for researchers, evaluators and practitioners to adopt complexity thinking when designing, implementing and evaluating health service and social development interventions given their inherent complexity (Braithwaite et al., 2018; Castellani et al., 2015; Greenhalgh & Papoutsis, 2018; Kannampallil et al., 2011; Rutter et al., 2017).

Literature review

The application of complexity thinking in programme design and evaluation is variously understood by researchers and practitioners (Damschroder et al., 2017; Greenhalgh & Papoutsis, 2018; Mowles, 2014; Ramalingam et al., 2008; Walton, 2014). Complexity thinking is an ontological position that understands the world as systemic, path-dependent, sensitive to context, emergent and episodic (Boulton et al., 2015; Mason, 2008). Some argue that many applications of complexity thinking do not adequately capture the effect of multiple interactions between agents and levels. In particular, some agent-based modelling tends to work on the basis of simple rules applied to individual actors which are dependent on the behaviour of their near

neighbours, like birds in a flock (Mowles, 2015). However, in human systems, our ideas, beliefs and interactions evolve as we learn, and the structures and institutions we create shape our subsequent behaviour (Befani & Mayne, 2014). These are not simply relationships of 'rule followers' based on individual, linear micro-interactions with little appreciation for the overall adaptive response of the system to the intervention (Byrne, 2013; Mowles, 2014). Agents 'adapt, interact and co-evolve with other systems' (Greenhalgh & Papoutsis, 2018, p. 2) and remind us that complexity is part of the broader landscape or system(s), not just a feature of a given intervention.

Evidence of the usefulness of adopting a complexity thinking approach to health service and social development practice and research is increasing (Braithwaite et al., 2018; Kannampallil et al., 2011; Mowles et al., 2008; Northridge & Metcalf, 2016; Rutter et al., 2017). The literature focuses on using complexity thinking to (1) understand public health phenomena such as epidemics (Agar & Wilson, 2002; Helbing et al., 2015), (2) evaluate intervention success or failure (Grol et al., 2007; Trenholm & Ferlie, 2013; Waqa et al., 2017) and (3) enhance management techniques for social development projects (Mowles et al., 2008). Within this literature, the application of a complexity thinking approach enhances understanding of systems and implementation and evaluation of strategies intended to create positive changes in them.

This article explores adopting complexity thinking in the design, implementation and evaluation of two case studies, the Pacific Leadership Program (PLP) and the What Works and Why (W3) Project. It recognises growing interest in the evaluation literature in understanding different stages of programme development and how they interact (Tsoukas, 2017; Van Ongevalle et al., 2014), and the importance of exploring how evaluation is shaped by the institutional context in which it is located (Arbour, 2020). Understanding context assists in determining the approaches to, and implementation of, monitoring and evaluation (Eckhard & Jankauskas, 2018), and the degree to which evaluation findings are actually used (Langer et al., 2016; Weiss, 1998). This article contributes to broader debates about the relationship between complexity, institutional arrangements and evaluative practice.

Methods

For the purposes of this article, we have engaged in a collaborative, reflexive, deliberation approach (CRDA; Crabtree et al., 2018). Crabtree and colleagues (2018) describe CRDA as a novel approach that uncovers critical insights and interpretations hidden across varying research contexts. Researchers collaborate, using CRDA to discuss and reflect on published and unpublished research findings to identify patterns and insights. We have done this by adopting complexity thinking as a practice lens by which to examine the similarities, differences and patterns between context and project interventions.

The author team collaborated to examine two seemingly diverse projects to uncover the areas of conjunction (Tsoukas, 2017). Authors G.B. and C.R. were directly involved in the case studies (W3 and PLP, respectively). The synthesis of their learning was developed through a process of analysis and discussion with all authors, using these

and other experiences. These included resourcing issues, the changing ‘authorising environment’ needed to enable adaptation during implementation, the demands of funders for communication of unambiguous short-term outcomes and the political nature of certain aspects meaning that important parts of the story could not be publicly told. However, lessons were also learnt about how complexity can be taken into account in programme design, implementation and evaluation, and about how these complexities, including its politics (and bureaucracy), can be managed through purposeful stakeholder engagement. To best reflect our shared experiences across contexts, we refer to key aspects of complexity as outlined by Boehnert (2018) to provide a framework for understanding the benefits of consistently using complexity thinking, by tying together diverse evidence through common language. From the 16 key characteristics of complexity provided by Boehnert (2018), we have exemplified only those key characteristics shared across the case studies to demonstrate apparent conjunctions. We draw on notions of emergence, self-organisation, levers and hubs, open and nested systems, and non-linearity not as absolutes and acknowledge there may be varying definitions of these features within the diverse complexity literature. In doing so, we hope to help exemplify how complexity thinking can be operationalised and the implications this has for the practice of evaluation.

The article therefore focuses on the *practice* of applying complexity thinking in two case studies. In doing so, we recognise that while there has been much theoretical discussion and debate in this area, as well as calls for a greater application of the concepts associated with complex adaptive systems, there has been much less discussion on how practitioners have attempted to operationalise these ideas and what lessons have been learnt as a result.

Context: overview of the case studies

Case study 1: W3 Project

Peer-led organisations run by and for communities of people who inject drugs, gay men, sex workers and people living with HIV have been a key characteristic of the Australian response to HIV and hepatitis C. These organisations conduct programmes ranging from peer-led needle and syringe programmes, to peer-developed social marketing and education, to peers taking leadership roles in policy and law reform. Peer-led organisations participate in complex communities as well as navigate highly stigmatised and changing contexts around sexuality and drug use (Brown et al., 2014; Madden & Wodak, 2014). However, these peer-led organisations often have difficulties in articulating their role, demonstrating their quality and community connection, and showing their effectiveness as part of a multi-sectoral public health response (Bajis et al., 2017; Brown et al., 2015; Crawford et al., 2013). The real-time community insights that peer-led organisations provide within a complex environment can be undervalued by all stakeholders including the peer organisations themselves, reducing their leverage (Brown et al., 2015; Marshall et al., 2015).

The W3 Project supported peer-led organisations to adapt their programmes to changing community needs, scale-up and demonstrate their impact in a complex system. The project used systems thinking methods (Meadows, 2008; Williams, 2010) to draw together insights of over 90 staff from 10 Australian peer organisations to develop a framework for peer-led organisations to implement activities, monitor and demonstrate their roles and influence (Brown et al., 2018). The framework identified four system-level functions which peer-led organisations needed to evaluate to demonstrate they were effective and sustainable: *active engagement* within their community, *alignment* within the policy system, *learning and adaptation* to the changes in the community and policy system, and *ongoing influence* within the community and policy systems. The framework highlighted important elements of complexity thinking when designing, implementing and evaluating the work of peer-led organisations. These elements included adaptability and feedback as they responded to changes in their community, levers and hubs, non-linearity and disrupting domains of stability as they endeavour to influence policy. The W3 Project subsequently worked with two additional peer-led organisations to apply the Framework to develop practical and sustainable ways to demonstrate their impact as they adapted to rapidly changing policy environments.

Case study 2: PLP

The PLP was an international development programme funded by the Australian Department of Foreign Affairs and Trade (DFAT) from 2008 to 2017. The programme aimed to support ‘developmental leadership’ in the Pacific and to support regional leadership processes. Developmental leadership was understood to involve supporting locally led processes of reform, which usually involve individual and collective action. The programme was designed to account for complexity of such reform by recognising the following:

- The nested, open nature of political and social systems in the Pacific which operate at multiple scales and levels. These provide reformers with multiple levers and hubs to enact change. Simultaneously, distributed control across these systems means no one agency or institution can unilaterally change policies and practices. Such shifts require alliances and coalitions to work across levels and with interest groups.
- Coalitions and collective action cannot be engineered by development agencies. They require locally owned and locally led emergent process based on principles of self-organisation than planning; of testing and trialling activities; and adjusting through feedback and adaptation.
- Power relations, vested interests and political settlements create domains of stability and path dependency, which explain continuity and resistance to change. The programme design took into account that tipping points, could occur and lead to dramatic non-linear change. An economic down-turn, a change in government or a natural disaster, for example, might be opportunities

for reform. Being open and flexible with support *initiatives and being able to deploy support nimbly were important;*

- All of the above led to high levels of unpredictability and unknowability about *which* reform initiatives were likely to be successful and therefore merited support, *what* were likely to be the most effective strategies to achieve reform and *how* these were to be prosecuted. This required a degree of ‘spread-betting’ and a portfolio approach to supporting coalitions, recognising not all ‘bets’ would pay off.

The Institute for Human Security and Social Change at La Trobe University was contracted during Phase 3 (2014–2017) to support evaluation and research and to explore the process of coalition building and associated policy reform, and the degree to which this was supported by the programme. This involved implementing evaluative research initiatives based on an understanding of complexity, non-linearity and emergence, notably, outcome harvesting (Wilson-Grau, 2015), social-network analysis (Durland & Fredericks, 2005), qualitative comparative analysis (QCA; Legewie, 2013) and action research (O’Keefe et al., 2015).

Findings: challenges and enablers

Capacity and resourcing issues

W3 Project. The communities in Australia most affected by HIV and hepatitis C are often marginalised and criminalised communities and constantly navigating stigma, politics and law. A strength of peer-led organisations is their unique relationship *within* their communities. This includes the ownership a community feels towards a peer-led organisation, drawing of staff and volunteers from the community, and the organisation’s participation in the community. This relationship provides peer organisations leverage within their communities and within the system. However, this also adds complexity. Peer staff are drawn from communities under pressure. While these communities have demonstrated resilience and innovation, this nevertheless impacts the flexibility there is across staff and volunteer resources – both in time and emotional energy – to undertake reflexive practice and to initiate new ways of working. Managing these elements can be hindered by the leverage role of other actors in the system, such as funders limiting resources, short-term funding contracts or the sector need for peer organisations to be an active participant in research and policy reform processes.

The W3 Project was a co-design project that required sustained commitment and trust from partners over multiple years. While peer-led organisations were strong advocates of the co-design approach (Joint United Nations Programme on HIV/AIDS, 2007), it took significant resources and energy from these organisations already working under financial and political pressure, and who have a history of participation in research without community benefit (Brown et al., 2019).

The W3 Project identified two key resource enablers which supported sustained participation:

- Flexibility of the research funding allowed the W3 Project to adapt and reorient methods and timelines, incorporate meaningful participation and make a multi-year commitment to peer organisations (Brown et al., 2019).
- Funders of the peer organisations demonstrated trust and flexibility in their commitment to the peer model and valued peer leadership within research and policy processes. This enabled the peer organisations and W3 Project to, for the first time, propose the articulation of research and policy participation and leadership into the scope and language of their funding contracts and reporting.

It was important to identify and leverage these system-level enablers, not just the capacity within individual organisations.

PLP. The programme design attempted to respond to critiques of previous programmes in the Pacific, which were deemed to be premised on linear and technocratic approaches to addressing complexity, and not adequately tailored to the local context. It had four key features (Denney & McLaren, 2016). First, there was a concerted attempt to understand local context and power dynamics and to recruit skilled staff from the Pacific Islands. Second, these staff used their understanding to identify partners and reform coalitions and to work in strategically and politically smart ways to support them. Third, staff worked behind the scenes using existing informal and personal networks to get things done. Fourth, the programme design and implementation, initially, took an open-ended ‘purposive muddling’ through approach (Denney & McLaren, 2016, p. 23) in which partnerships with local leaders did not start with predetermined objectives or solutions.

PLP initially had a relatively well-resourced evaluation and research team, which meant their internal capacity to monitor, manage and engage in research was high. However, budget cuts meant staff were moved on to other organisations and not replaced. This weakened staff engagement and the research outcomes came to be seen as the priority rather than the *process* of involvement.

In addition, finding researchers who could do the work; had the social, political and interpersonal skills to relate effectively to programme staff and coalition members; and the communication skills to write clearly proved challenging. This led to discussions about the ‘unicorn problem’, that is, the difficulty of finding rare beasts.

Genuine coalitions and collective action cannot be engineered by development agencies. They require locally owned and locally led emergent process which result in processes based more on principles of self-organisation than planning; of testing and trialling activities; and adjusting through feedback and adaptation. While it is commonplace to see DFAT programme designs incorporating some principles of complexity, monitoring and evaluating these processes are challenging and require resources and evaluators who are skilled in action research and developmental evaluation.

The system needs to allow/enable adaptation to happen

W3 Project. The HIV and hepatitis C sectors have long navigated the politically and legally fraught issues of sexuality and drug use. Peer organisations are often nimble and flexible within these complex and politically volatile community systems and achieve synergies across limited resources. We found the strength of peer-led organisations was their capacity to mediate between, and navigate within, the complex community and policy/health service systems. For peer-led organisations to implement and maintain a complex system approach (such as constantly adapt, monitor for emergence and pursue leverage), they need an environment that enables them to do this.

The W3 Project found a peer-led organisation that is restricted or constrained due to legal (e.g. preventing peer distribution of clean injecting equipment) or contracting impediments (e.g. short-term funding) is undermined in its ability to maintain connection and relevance with its community; thus, the enabling environment had been affected. Losing connection with community has a flow on impact, reducing the peer organisation's leverage to be effective and the credibility of their advocacy in policy and services. An enabling policy environment allows high adaptability, enhancing the relationship between the peer-led organisation and their community, and the accuracy of the policy advice they can provide.

We found peer organisations were navigating a policy and funding environment with an emphasis on evaluation at the macro level (e.g. whole state/jurisdiction policy to increase testing and treatment) and simple accountability monitoring at the micro level (e.g. outputs of individual projects) with little meso-level complex system focus, such as evaluating how organisations interrelated.

We found this had advantages and disadvantages. While there was a disconnect between the micro project-level reporting and the macro whole of state epidemiological reporting, to some extent this may have allowed for the complexity of the system to operate and adapt quickly without interference. The desire for simplicity encouraged the omitting of the complexity – which perhaps provided a level of autonomy or 'protection' from interference.

However, with little accountability at the meso level, this meant the system-level influence of peer-led organisations, and the barriers and enablers they experience, was left invisible and unrecognised. For example, health or law enforcement organisations could be enablers or barriers to peer organisations fulfilling their potential role. The invisibility within reporting and evaluation structures regarding the inter-organisational interactions in the system left open the capacity for institutionalised stigma regarding communities (e.g. punitive policing of people who use drugs, people with HIV or sex workers) and so undermine the potential to leverage more effective system outcomes.

The W3 Project advocated with peer organisations for inclusion in funding contracts 'alignment' indicators that other organisations in the HIV/hepatitis C response were fulfilling their role (leverage) for the peer programme to be effective (and vice versa).

While peer-led organisations endeavour to demonstrate their capacity and credibility, policy and service organisations also needed to recognise their role to value peer leadership and enable peer-led organisations to achieve their potential (Brown et al., 2018, 2019).

PLP. The programme design and implementation recognised the volatility and complexity of the environment and the non-linear nature of reform processes. This resulted in informal and formal processes for programme adjustment in light of changing circumstances and what was being learnt. It included the introduction of six-monthly Review and Reflection (R&R) sessions designed to reserve a ‘safe space’ for programme review. The intention of these sessions was provoking ‘courageous conversations’ through inviting ‘critical friends’ to provide input (Roche & Kelly, 2012, p. 9). These sessions were designed to complement routine meetings and exchanges with programme partners, and the collection of basic activity and output data.

The resourcing and capacity issue related to evaluation was compounded by the challenge of needing to present a simple, aggregated story of the programme’s achievements, preferably based on measurable outcomes. This pressure increased during the programme in line with changes in Australia’s domestic political environment and growing scepticism about foreign aid (Corbett, 2017). In the programme’s last phase, there was little appetite for a complex rendering of the programme’s nuanced support to 31 coalition partners and the even subtler story of how they had, or had not, made progress.

Communicating success and performance

W3 Project. Demonstrating the impact of multiple peer-led programmes and their cumulative leadership role is challenging (Cain et al., 2014; Collins et al., 2016; Genberg et al., 2016; Kielmann & Cataldo, 2010; Maxwell et al., 2008). Peer-led health promotion is about relationships between the programme, the communities they work with, and the policy and political environment within which they operate (Brown et al., 2018, 2019). Peer-led organisations are simultaneously accountable to their communities and to their funders/policy system. Evaluation that focuses on individual short-term projects can struggle to capture the complexity of the peer-led organisations’ ongoing interaction with and adaptation to the rapidly changing demands of community and socio-political contexts in which they are embedded (Brown et al., 2015).

To maintain credibility with their communities, peer-led organisations need to demonstrate their influence beyond their community. Taking a systems perspective helped peer-led organisations illustrate to their funders that delivering trusted and credible peer services and undertaking peer advocacy for improved policy and structural reform were linked. Evaluating and communicating success was not just about meeting accountability demands to enable sustained funding, it also leveraged reputational and political legitimacy in both community and the policy system that was needed to increase credibility and balance institutionalised stigma.

PLP. Early in the programme, it was agreed the funding agency – the Department of Foreign Affairs and Trade – would locate two staff in the programme which was run by a managing contractor. During this time, the programme leadership was assumed by the funder, but the implementation was the responsibility of the managing contractor. This arrangement enabled responsive decision-making, allowing greater programme adaptation; a greater appreciation and management of political risks and how they could be mitigated, which in turn allowed for less risk-averse decision-making; and higher levels of trust and a degree of protection of staff from reporting demands (Denney & McLaren, 2016; Henderson & Roche, 2012). In addition, the programme retained its independence from the funding agency which was welcomed by programme partners. This configuration enabled the programme to work in politically savvy ways which were congruent with the original programme design and the complexity of its operations.

This arrangement was discontinued in the programme's third phase, with the managing contractor assuming full responsibility. This slowed decision-making, hampered the programme's agility and diminished mutual understanding between programme and funding staff.

Coalitions supported by PLP were involved in delicate political processes, including challenging vested interests – this was inherent in the change theory the programme design had established. However, this meant publicly sharing this information through reports would have diminished the trust necessary for coalitions to work together, exposed individuals and organisations to greater threats, and ultimately have undermined the reform process. Important and timely insights could often not be reported. This meant the complexity could only be partially told, and key political mechanisms in complex generative processes were deliberately or strategically withheld – even if they were informally shared with funding staff through the co-location arrangement. A complex understanding of the specificity of a context and process is sometimes simplified in formal documentation, which can be to the detriment of those seeking to learn or adapt that experience using written sources.

Managing diverse interests

W3 Project. Many early peer education activities began as illegal initiatives conducted at risk of conviction by community and health services, for example, needle and syringe programmes, (Brown et al., 2014). Peer programmes are often funded by governments to 'deal with' communities and health issues deemed complex, taboo and politically difficult.

Providing funds for peer health programmes conducted by communities of people who use drugs, gay men, sex workers and people living with HIV or hepatitis C is a political as well as an evidence-based decision. The political will and funding for peer programmes can be driven by ideological rather than evidence-based origins, and the capacity and autonomy of organisations can ebb and flow within short time frames with limited connection to the needs of communities, synergies with other projects or partnerships with other organisations. This could lead to rhetoric in policy frameworks

about community mobilisation, peer leadership and structural change at a policy level, but projects being funded and managed as standalone and often short-term projects due to the political sensitivities involved. The restrictive framing of the funding contract can limit the capacity of peer-led organisations to take a complex systems approach.

W3 Project worked with peer organisations to advocate with policy and funders to recognise the *leverage* and *self-organising* contribution of peer organisations to achieve the high-level outcomes of the overall HIV and hepatitis C policies (*emergence*) and to identify the indicators that meet the needs of the contract management of individual projects but also show whether the combined investment in peer organisations was fulfilling its system role, rather than just individual projects.

PLP. Programme staff in PLP needed to ‘think and work politically’ (Faustino & Booth, 2014), with their funders and with the reform coalitions they supported. This meant continually navigating between domestically driven policies and bureaucratic agendas in the Pacific Island countries they were supporting. This seemed to mean two ‘layers’ of internal support were generally required in DFAT to protect their ‘authorising environment’ to work in the way they did. This demands a set of skills and high inter-cultural competence to work across different institutional environments and across different cultural values and worldviews. In particular, this required the following:

- A nuanced understanding of the agendas of different actors at different levels and how these agendas affected programme success, as well as developing tailored strategies to meet these diverse interests;
- Nurturing relationships with system advocates who are willing to help explain and translate programme strategies and successes in ways more sceptical colleagues would find convincing;
- Selecting advisors and evaluators who understood the approach being tested and were able to craft innovative means of monitoring and evaluating the programme.

Discussion

While the two case studies explored in this article are in many ways very different, including that they operate in very different contexts, we noted a number of important similarities. That these resemblances occur despite these differences in context suggests that these findings may have more generalisable validity. We propose four specific characteristics that enable the practice of adopting complexity thinking and outline a set of common challenges associated with the design, implementation and evaluation of working in this way, and which are of particular relevance to evaluation practitioners.

First, knowledge of the context and the key actors involved is an essential starting point as is an acceptance of uncertainty and non-linearity in the design of interventions (Befani & Mayne, 2014). Both projects had an intimate understanding of the

environment in which they worked and close engagement with peer networks in the case of the W3 Project and reform coalitions in the case of PLP. They did not impose an external understanding of how change happens in these contexts but rather sought to allow for emergent co-designed processes which recognised practice knowledge (Donetto et al., 2015). Furthermore, both projects recognised the diversity and differences between different groups within them. The W3 Project, for example, developed separate systems maps with different community groups before seeking to identify common patterns and dynamics across the peer programmes and the eco-system they were part of. Similarly, PLP carefully tailored its support and engagement to the specificities of the different coalitions they worked with, noting that some worked at the very local level and some at national or regional level across the Pacific.

Second, both case studies demonstrated the importance of trust and the establishment of relationships and reflexive spaces for knowledge generation and sharing (Tsoukas, 2017). In particular, this required valuing different forms of knowledge and experience, and the importance of uncovering the generative processes which led to outcomes, that is, the combinations of factors that help explain complex change (Befani & Mayne, 2014). In the case of W3, this involved understanding how different levels of systems change interacted and recognising the need for a degree of alignment at the practice, organisational and policy levels. In the case of PLP, this led to trialling monitoring and evaluation activities which were based more on conjunctural and generative logics – such as QCA and outcome harvesting, rather than those based on more counterfactual logic (Schatz & Welle, 2016). At the same time, both projects also engaged peer networks or reform coalition members in regular ‘sense-making’ exercises which sought to bring together experiential learning and knowledge and monitoring and evaluation data. This provided important opportunities to explore not only ‘what’ was, or was not, changing but ‘how’ and ‘why’ things were, or were not, shifting. In essence, this generated important feedback loops, which then led into adjustments in strategy adaptation over time.

Third, both projects had to work in politically savvy ways. This meant recognising the interests and needs of different stakeholders involved, framing things in ways that might resonate with them and exploring how best to meet their needs (Donetto et al., 2015). In some cases, this also meant ‘educating’ different parties so that they were better able to understand the contribution of peer organisations or reform coalitions to less visible processes of systems change and how they went about this, in addition to more visible individual project outputs or outcomes. The W3 Project’s synthesis of systems maps with peer organisations produced a more visible representation of the complex role played by peer organisations, but at the same time recognised the need for funders to have performance indicators to satisfy their accountability requirements. In PLP’s case, this required an investment in highly skilled staff and advisors who had the cross-cultural skills to work across different stakeholder groups to help balance demands in ways that did not distort the iterative and contingent ways of working, which were central to the programme.

Finally, both projects experienced a number of common difficulties and challenges in attempting to support peer organisations and reform coalitions in ways that recognise the complexity of the processes they are involved in and the complexities of the

eco-system of which they are a part (Arbour, 2020). In particular, documenting outcomes and success in ways that are generally preferred by funders or governments proved tricky. In part, this stems from a particular notion and common understanding of accountability based on principal–agent models which seek to assess performance based on pre-defined outcomes and indicators. It also stems from a preference for aggregated, preferably quantified, succinct and unambiguous performance reporting, which tends to elide nuance, diversity and uncertainty. Furthermore, funding for work on HIV/AIDS in Australia or international development in the Pacific is also provided for political or ideological purposes which go beyond the justification for, or evidence produced by, a particular programme or project. The changing nature of that political interest can shift quickly, which in turn can mean that the enabling or authorising environment for work in both domains can, and did, change in short order. This volatility suggests that projects or programmes that seek to engage with the complexity of the context in which they operate need to think hard about how they can create the political space to support their work and build the alliances necessary to maintain it (Arbour, 2020). Engaging stakeholders early in that journey and investigating different ways in which changing demands can be met is perhaps one means of doing so.

Conclusion

This article has attempted to capture, depict and describe the way in which a complexity thinking approach can be applied to design, implementation and evaluation of interventions in practice. While there is a growing interest in, and call for, the adoption of notions of complexity in programmes designed to strengthen health systems and in international development, much less has been written about how this might be done.

We note that working in this way requires adequate capacity and resourcing as it takes time, skill and money. It also requires an enabling or authorising environment for learning and adaptation. This puts an onus on such programmes and their evaluation to have the ability to communicate imaginatively and manage diverse interests in politically savvy ways.

However, both cases demonstrate that with a grounded knowledge of the context, and effective relationships and engagement with the key actors involved, opportunities for knowledge generation and sharing can be established. Much of this is consistent with a recent collection of papers on ‘Understanding Complexity in Health Systems’ (Greenhalgh & Papoutsi, 2018) and the ideas of ‘conjunctive theorising’, that is, the weaving together of diverse forms of knowledge from diverse sources (Tsoukas, 2017). When these learning processes are combined with robust methods of monitoring and evaluation which recognise emergent non-linear change, it seems this action-reflection approach can not only provide the necessary feedback for programme adaptation but also provide some of the evidence required to maintain the authorising environment to operate.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD

Samantha Clune  <https://orcid.org/0000-0003-2553-5776>

References

- Agar, M. H., & Wilson, D. (2002). Drugmart: Heroin epidemics as complex adaptive systems. *Complexity*, 7(5), 44–52. <https://doi.org/10.1002/cplx.10040>
- Arbour, G. (2020). Frameworks for program evaluation: Considerations on research, practice, and institutions. *Evaluation*, 26(4), 422–437. <https://doi.org/10.1177/1356389020920890>
- Bajis, S., Dore, G. J., Hajarizadeh, B., Cunningham, E. B., Maher, L., & Grebely, J. (2017). Interventions to enhance testing, linkage to care and treatment uptake for hepatitis C virus infection among people who inject drugs: A systematic review. *International Journal of Drug Policy*, 47, 34–46.
- Befani, B., & Mayne, J. (2014). Process tracing and contribution analysis: A combined approach to generative causal inference for impact evaluation. *IDS Bulletin*, 45(6), 17–36. <https://doi.org/10.1111/1759-5436.12110>
- Boehnert, J. (2018). *The visual representation of complexity*. <https://www.cecan.ac.uk/sites/default/files/2018-07/JP%20online%20pdf%20The%20Visual%20Communication%20of%20Complexity%20-%20May2018%20-%20EcoLabs.pdf>
- Boulton, J. G., Allen, P. M., & Bowman, C. (Eds.). (2015). *Embracing complexity: Strategic perspectives for an age of turbulence*. Oxford University Press.
- Braithwaite, J., Churrua, K., Ellis, L. A., Long, J., Clay-Williams, R., Damen, N., . . . Ludlow, K. (2017). *Complexity Science in healthcare – aspirations, approaches, applications and accomplishments: A white paper*. Macquarie University.
- Braithwaite, J., Churrua, K., Long, J. C., Ellis, L. A., & Herkes, J. (2018). When complexity science meets implementation science: A theoretical and empirical analysis of systems change. *BMC Medicine*, 16(1), Article 63. <https://doi.org/10.1186/s12916-018-1057-z>
- Brown, G., O'Donnell, D., Crooks, L., & Lake, R. (2014). Mobilisation, politics, investment and constant adaptation: Lessons from the Australian health-promotion response to HIV. *Health Promotion Journal of Australia*, 25(1), 35–41. <https://doi.org/10.1071/HE13078>
- Brown, G., Perry, G.-E., Byrne, J., Crawford, S., Henderson, C., Madden, A., . . . Reeders, D. (2019). Characterising the policy influence of peer-based drug user organisations in the context of hepatitis C elimination. *International Journal of Drug Policy*, 72, 24–32. <https://doi.org/10.1016/j.drugpo.2019.05.025>
- Brown, G., Reeders, D., Cogle, A., Madden, A., Kim, J., & O'Donnell, D. (2018). A systems thinking approach to understanding and demonstrating the role of peer-led programs and leadership in the response to HIV and hepatitis C: Findings from the W3 project. *Frontiers in Public Health*, 6, Article 231.
- Brown, G., Reeders, D., Dowsett, G. W., Ellard, J., Carman, M., Hendry, N., & Wallace, J. (2015). Investigating combination HIV prevention: Isolated interventions or complex system. *Journal of the International AIDS Society*, 18(1), 20499.
- Byrne, D. (2013). Evaluating complex social interventions in a complex world. *Evaluation*, 19(3), 217–228. <https://doi.org/10.1177/1356389013495617>
- Crabtree, B. F., Miller, W. L., Gunn, J. M., Hogg, W. E., Scott, C. M., Levesque, J.-F., . . . & Russell, G. M. (2018). Uncovering the wisdom hidden between the lines: The collaborative reflexive deliberative approach. *Family Practice*, 35(3), 266–275. <https://doi.org/10.1093/fampra/cmz091>

- Cain, R., Collins, E., Bereket, T., George, C., Jackson, R., Li, A., . . . Travers, R. (2014). Challenges to the involvement of people living with HIV in community-based HIV/AIDS organizations in Ontario, Canada. *AIDS Care*, *26*(2), 263–266.
- Castellani, B., Rajaram, R., Buckwalter, J. G., Ball, M., & Hafferty, F. W. (2015). *Place and health as complex systems: A case study and empirical test*. Springer.
- Collins, L. M., Kugler, K. C., & Gwadz, M. V. (2016). Optimization of multicomponent behavioral and biobehavioral interventions for the prevention and treatment of HIV/AIDS. *AIDS and Behavior*, *20*(1), 197–214.
- Corbett, J. (2017). *Australia's foreign aid dilemma: Humanitarian aspirations confront democratic legitimacy*. Routledge.
- Crawford, G., Bowser, N. J., Brown, G. E., & Maycock, B. R. (2013). Exploring the potential of expatriate social networks to reduce HIV and STI transmission: A protocol for a qualitative study. *BMJ Open*, *3*(2), Article e002581.
- Damschroder, L. J., Reardon, C. M., Sperber, N., Robinson, C. H., Fickel, J. J., & Oddone, E. Z. (2017). Implementation evaluation of the Telephone Lifestyle Coaching (TLC) program: Organizational factors associated with successful implementation. *Translational Behavioral Medicine*, *7*(2), 233–241. <https://doi.org/10.1007/s13142-016-0424-6>
- Denney, L., & McLaren, R. (2016). *Thinking and working politically to support developmental leadership and coalitions: The Pacific Leadership Program* (Developmental Leadership Program Research Paper, No. 41). <https://www.dlprog.org/publications/research-papers/thinking-and-working-politically-to-support-developmental-leadership-and-coalitions-the-pacific-leadership-program>
- Donetto, S., Pierri, P., Tsianakas, V., & Robert, G. (2015). Experience-based co-design and healthcare improvement: Realizing participatory design in the public sector. *The Design Journal*, *18*(2), 227–248. <https://doi.org/10.2752/175630615X14212498964312>
- Durland, M. M., & Fredericks, K. A. (2005). An introduction to social network analysis. *New Directions for Evaluation*, *2005*(107), 5–13. <https://doi.org/10.1002/ev.157>
- Eckhard, S., & Jankauskas, V. (2018). The politics of evaluation in international organizations: A comparative study of stakeholder influence potential. *Evaluation*, *25*(1), 62–79. <https://doi.org/10.1177/1356389018803967>
- Faustino, J., & Booth, D. (2014). *Development entrepreneurship: How donors and leaders can foster institutional change*. <https://asiafoundation.org/publication/development-entrepreneurship-how-donors-and-leaders-can-foster-institutional-change/>
- Genberg, B. L., Shangani, S., Sabatino, K., Rachlis, B., Wachira, J., Braitstein, P., & Operario, D. (2016). Improving engagement in the HIV care cascade: A systematic review of interventions involving people living with HIV/AIDS as peers. *AIDS and Behavior*, *20*(10), 2452–2463.
- Greenhalgh, T., & Papoutsi, C. (2018). Studying complexity in health services research: Desperately seeking an overdue paradigm shift. *BMC Medicine*, *16*(1), Article 95. <https://doi.org/10.1186/s12916-018-1089-4>
- Grol, R., Bosch, M., Hulscher, M., Eccles, M., & Wensing, M. (2007). Planning and studying improvement in patient care: The use of theoretical perspectives. *The Milbank Quarterly*, *85*(1), 93–138.
- Helbing, D., Brockmann, D., Chadefaux, T., Donnay, K., Blanke, U., Woolley-Meza, O., . . . Perc, M. (2015). Saving human lives: What complexity science and information systems can contribute. *Journal of Statistical Physics*, *158*(3), 735–781. <https://doi.org/10.1007/s10955-014-1024-9>

- Henderson, S., & Roche, C. (2012). *An independent progress report of the Pacific Leadership Program*. Australian Agency for International Development.
- Joint United Nations Programme on HIV/AIDS. (2007). *The greater involvement of people living with HIV (GIPA): Policy brief*.
- Kannampallil, T. G., Schauer, G. F., Cohen, T., & Patel, V. L. (2011). Considering complexity in healthcare systems. *Journal of Biomedical Informatics*, 44(6), 943–947. <https://doi.org/10.1016/j.jbi.2011.06.006>
- Kielmann, K., & Cataldo, F. (2010). Tracking the rise of the ‘expert patient’ in evolving paradigms of HIV care. *AIDS Care*, 22(Suppl. 1), 21–28.
- Langer, L., Tripney, J., & Gough, D. (2016). *The science of using science: Researching the use of research evidence in decision-making*. EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.
- Legewie, N. (2013). An introduction to applied data analysis with qualitative comparative analysis. *Forum: Qualitative Social Research*, 14(3), 15. <https://doi.org/10.17169/fqs-14.3.1961>
- Madden, A., & Wodak, A. (2014). Australia’s response to HIV among people who inject drugs. *AIDS Education and Prevention*, 26(3), 234–244. <https://doi.org/10.1521/aep.2014.26.3.234>
- Marshall, Z., Dechman, M. K., Minichiello, A., Alcock, L., & Harris, G. E. (2015). Peering into the literature: A systematic review of the roles of people who inject drugs in harm reduction initiatives. *Drug and Alcohol Dependence*, 151, 1–14. <https://doi.org/10.1016/j.drugalcdep.2015.03.002>
- Mason, M. (2008). What is complexity theory and what are its implications for educational change? *Educational Philosophy and Theory*, 40(1), 35–49.
- Maxwell, C., Aggleton, P., & Warwick, I. (2008). Involving HIV-positive people in policy and service development: Recent experiences in England. *AIDS Care*, 20(1), 72–79.
- Meadows, D. H. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing.
- Mowles, C. (2014). Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship. *Evaluation*, 20(2), 160–175. <https://doi.org/10.1177/1356389014527885>
- Mowles, C. (2015). *Managing in uncertainty: Complexity and the paradoxes of everyday organizational life*. Routledge.
- Mowles, C., Stacey, R., & Griffin, D. (2008). What contribution can insights from the complexity sciences make to the theory and practice of development management? *Journal of International Development*, 20(6), 804–820. <https://doi.org/10.1002/jid.1497>
- Northridge, M. E., & Metcalf, S. S. (2016). Enhancing implementation science by applying best principles of systems science. *Health Research Policy and Systems*, 14(1), 74. <https://doi.org/10.1186/s12961-016-0146-8>
- O’Keefe, M., Sidel, J. T., Marquette, H., Roche, C., Hudson, D., & Dasandi, N. (2015). *Using action research and learning for politically informed programming* (DLP Research Paper 29). <http://eprints.lse.ac.uk/60766/>
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. SAGE.
- Ramalingam, B. (2013). *Aid on the edge of Chaos: Rethinking international cooperation in a complex world*. Oxford University Press.
- Ramalingam, B., Jones, H., Toussainte, R., & Young, J. (2008). *Exploring the science of complexity: Ideas and implications for development and humanitarian efforts* (Working Paper 285). Overseas Development Institute.
- Roche, C., & Kelly, L. (2012). *The evaluation of politics and the politics of evaluation* (Background Paper 11). Development Leadership Program.

- Rutter, H., Savona, N., Glonti, K., Bibby, J., Cummins, S., Finegood, D. T., . . . White, M. (2017). The need for a complex systems model of evidence for public health. *The Lancet*, *390*(10112), 2602–2604. [https://doi.org/10.1016/S0140-6736\(17\)31267-9](https://doi.org/10.1016/S0140-6736(17)31267-9)
- Schatz, F., & Welle, K. (2016). *Qualitative comparative analysis: A valuable approach to add to the evaluator's toolbox? Lessons from recent applications* [CDI practice paper]. Institute of Development Studies.
- Sheikh, K., Gilson, L., Agyepong, I. A., Hanson, K., Ssengooba, F., & Bennett, S. (2011). Building the field of health policy and systems research: Framing the questions. *PLOS Medicine*, *8*(8), Article e1001073. <https://doi.org/10.1371/journal.pmed.1001073>
- Trenholm, S., & Ferlie, E. (2013). Using complexity theory to analyse the organisational response to resurgent tuberculosis across London. *Social Science & Medicine*, *93*, 229–237. <https://doi.org/10.1016/j.socscimed.2012.08.001>
- Tsoukas, H. (2017). Don't simplify, somplexify: From disjunctive to conjunctive theorizing in organization and management studies. *Journal of Management Studies*, *54*(2), 132–153. <https://doi.org/10.1111/joms.12219>
- Van Ongevalle, J., Huyse, H., & Van Petegem, P. (2014). Dealing with complexity through actor-focused planning, monitoring and evaluation (PME). *Evaluation*, *20*(4), 447–466. <https://doi.org/10.1177/1356389014551487>
- Walton, M. (2014). Applying complexity theory: A review to inform evaluation design. *Evaluation and Program Planning*, *45*, 119–126. <https://doi.org/10.1016/j.evalprog-plan.2014.04.002>
- Waq, G., Moodie, M., Snowdon, W., Latu, C., Coriakula, J., Allender, S., & Bell, C. (2017). Exploring the dynamics of food-related policymaking processes and evidence use in Fiji using systems thinking. *Health Research Policy and Systems*, *15*(1), 74. <https://doi.org/10.1186/s12961-017-0240-6>
- Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, *19*(1), 21–33. <https://doi.org/10.1177/109821409801900103>
- Williams, B. (2010). Systems thinking and capacity development in the international arena. In N. Fujita (Ed.), *Issues and prospects of evaluation for international development: Beyond logframe; using systems concepts in evaluation* (pp. 35–53). Foundation for Advanced Studies on International Development.
- Wilson-Grau, R. (2015). *Outcome harvesting*. BetterEvaluation. http://betterevaluation.org/plan/approach/outcome_harvesting