

DOI: 10.1002/ev.20467

ORIGINAL ARTICLE

WILEY

Using systems methods to elicit complex program theories

Daniel Reeders¹ | Graham Brown²

¹ Australian National University College of Asia and the Pacific

² La Trobe University

Correspondence

Daniel Reeders, Australian National University College of Asia and the Pacific
Email: daniel.reeders@anu.edu.au

Abstract

Contextual interventions pose unique challenges for implementation studies and evaluation research. By contextual, we mean programs that are tightly integrated with the social systems they are seeking to influence or transform. This integration challenges foundational assumptions of intervention science and necessitates the development of innovative approaches to monitoring and evaluation. The W3 ('What Works and Why') project piloted the use of complex systems theory and methods (CSTM) to develop program theories for peer and community-led (PCL) programs in HIV and viral hepatitis prevention and health promotion. As a class, these programs have struggled to articulate and theorize their approach to justify their funding and to build an evidence base. CSTM allowed us to articulate detailed program theories based on 'mental models' held by program staff and stakeholders. This chapter describes our reflection and adaptation experience as we piloted the CSTM approach in—and *as*—social research.

INTRODUCTION

Peer-based programs have a long history in national responses to epidemics of HIV and viral hepatitis. The peer approach is inherently relational and embedded in local cultures and communities, with practitioners drawing on their own experience to establish rapport and work effectively with contacts and clients. These epidemics affect historically marginalized communities, including gay and bisexual men (cis and trans) and people who use drugs (PWUD). The peer approach creates a more egalitarian basis for encounters between practitioners and clients/contacts. This takes on particular importance when addressing stigmatized conditions in vulnerable groups. In practice, volunteers and paid staff work as peer educators, facilitate one-to-one and group-based peer support, help clients navigate the health system and social services, participate in developing projects and services, and advocate for policy reforms.

1

2

2 | WILEY

USING SYSTEMS METHODS TO ELICIT COMPLEX PROGRAM THEORIES

3 Research on the peer approach as a mode of intervention has typically concentrated
4 on individual behavior effects, overlooking its relational and contextual elements (Brown
5 et al., 2013). Consequently, peer-based programs have faced ongoing difficulty in justifying
6 their funding, articulating their theory of change (Chen, 2005), and building an evidence
7 base to support quality practices. Indeed, the approach has been described as ‘a method
8 in search of a theory’ (Turner & Shepherd, 1999). By theory, Turner and Shepherd intend
9 ‘capital T’ social psychological theories of health education, such as the Health Belief Model
10 (Taylor et al., 2006). In the W3 (What Works and Why) Project, we took a different tack, con-
11 ceptualising the peer approach as a contextual intervention and using complex systems
12 theory and methods (CSTM) to inform the development of ‘lower-case T,’ program theo-
13 ries (Weiss, 1997) for three peer-based programs working at different levels, from frontline
14 service provision to federal policy advocacy, in the Australian context.

15 Early frameworks for evaluating complex interventions focused on interventions with
16 lots of moving parts (Craig et al., 2008). Contextual interventions are complex in a differ-
17 ent way, being tightly integrated with the social systems they seek to influence or trans-
18 form. This challenges foundational assumptions of intervention science, particularly the
19 ability to measure changes in the context in separation from the intervention (Green &
20 Higgins, 2008). As the Cochrane Collaboration acknowledges (ibid), this challenge neces-
21 sitates the development of innovative approaches to evaluating contextual interventions.
22 The W3 project piloted the use of CSTM to articulate and theorize how peer-based pro-
23 grams engage with, influence, and respond to their diverse and dynamic social and pol-
24 icy contexts. We hoped the systems perspective might help us identify what matters when
25 evaluating PCL interventions. Focusing on individual behaviors overlooks how making a
26 difference in policy advocacy is dependent on close engagement with the community and
27 *vice versa*.

28 Indeed, Finegood and colleagues argue that evaluators of complex systems’ interven-
29 tions must shift their focus from attributing causal effects to evaluating (and supporting)
30 the adaptation of interventions to their context (Johnston, Matteson, & Finegood, 2014).
31 Accordingly, we used the program theories we elicited to develop a monitoring, evalua-
32 tion, and learning (MEL) framework supporting and evaluating adaptation. The framework
33 identifies key functions that allow long-running contextual interventions to maintain rele-
34 vance and influence under the conditions of uncertainty that are ever-present in a contin-
35 ually changing social and policy environment. Therefore, monitoring these functions can
36 inform the confidence of stakeholders (which includes communities, partners, funders,
37 policy-makers, and implementation scholars) in the adaptive capacity of the intervention.
38 Making these functions explicit allows ‘flexible fidelity’ when peer-based approaches are
39 implemented with new issues and communities, focusing on what needs to be happen-
40 ing rather than reproducing the intervention’s specific details as it was originally designed
41 (Chambers & Norton, 2016; Hawe, Shiell, & Riley, 2004).

42 This chapter provides a case study of our experience of piloting CSTM as participatory
43 research methods within an evaluation research project conducted as an equal partner-
44 ship with three different PCL programs. This experience reflects the same phenomenon
45 we describe in our findings: a process of learning and adaptation based on insights from
46 our participants and reflection on our practice as researchers. Consistent with the values of
47 the peer approach itself, we prioritized our participants’ insights, adapting our approach
48 in response to their feedback on the process. Originally, we intended to develop computer
49 simulations from the models we elicited from our participants; as we recount here, partic-
50 ipants pushed back, and we developed alternative approaches.

51 At times, this meant adapting CSTM methods, which, in turn, required critical enquiry
52 into their onto-epistemological assumptions, which are not always explicit. The result is
53

'impure' as CSTM but has greater depth as social research. Pointers are provided for readers wanting specific instructions on CSTM methods, and additional detail on the W3 project can be found elsewhere (Brown et al., 2018). First, we describe the case study, including our three partner programs. We describe our use of participatory methods to develop *causal loop diagrams* (CLDs) and reflect on the process of analyzing the outputs. We conclude with a reflection on the strengths and difficulties encountered in our use of CSTM.

CASE STUDY

This chapter offers a case study of the initial phases of the W3 project (2014-15). We held participatory workshops and developed CLD-style program theories with three partner programs, each one working at a different level of the Australian response to HIV and hepatitis C virus (HCV).

Our first partner was Peer Based Harm Reduction Western Australia, which operates a peer-based needle and syringe exchange program (NSEP) integrated with harm reduction services and advocacy. This chapter describes our experience of the initial workshop, how we adapted CSTM methods in response to participant feedback, and what that enabled us to show. Subsequent partners included a health promotion initiative targeting networks and cultures among 'sexually adventurous' men who have sex with men, and a federal advocacy network among federal and state-based organizations for people living with HIV.

The project then developed a monitoring, evaluation, and learning (MEL) framework, based on a mid-level theory developed via cross-case analysis (Pawson & Tilley, 1997) of the three program theories. Finally, we assessed the validity of the framework in consultation with nonpartner programs and stakeholders. (Findings from the pilot implementation of the MEL framework can be found at w3project.org.au.) The project was strategic research funded by the Australian Government Department of Health and Ageing and undertaken with ethics approval from La Trobe University Human Research Ethics Committee (FHEC14/155)

METHODS

We drew from the 'toolkit' of systems thinking and modeling (STM) (Maani & Cavana, 2000). STM is often described as a method for integrating multiple perspectives to achieve a consensual model of a complex problem (Meadows & Wright, 2009; Senge, 2010). Proponents of STM advocate the use of interactive workshops with stakeholders who hold multiple perspectives. In our project, we aimed to address two problems: the limitations of the public health literature and, more locally, the lack of understanding of the role and contribution of PCL programs in the Australian prevention system. Thus, our challenge was two-fold: we wanted to elicit and integrate our participants' various mental models of their work and to visualize how these programs—as exemplars of the larger class of contextual interventions—are embedded in diverse and dynamic social and policy systems.

Causal loop modeling offers a precise set of conventions to translate implicit understandings into explicit diagrams that visualize the cyclical relationships among key variables in a problem situation (Maani & Cavana, 2000, p. 7). Whereas 'program logic' describes a linear causal flow of determinants, activities, and effects (Kellogg Foundation, 2004), by developing program theories in the CLD format, we were able to demonstrate how programs emerge out of the *interaction* of many different elements, many of them operating beyond each program's formal boundaries (Carvalho, 2013). This section will

Box 1

Jess is 28 years old and has been injecting heroin for about 5 years. She has a partner Greg who also uses, somewhat more frequently than she does. They sometimes argue about this. He does casual work as a laborer and his income is unreliable. Jess takes the pill for contraception, except when she forgets, and then she usually gets Greg to use a condom. Jess is usually the one who visits an NSEP to pick up a new kit, usually in a hurry after work, so they can use on her weekend. Greg scores for both of them from his friends and doesn't like going to the NSEP.

describe how these methods are intended to work and how we experienced their use in practice. We describe how we validated and analyzed the models to generate research findings and share our reflections on the strengths and challenges we identified.

Participatory workshops

Participatory workshops were appealing because our work aimed to center and reflexively engage with peer experience. Commitment to the 'phenomenology of frontline practice' is a key principle of community-based responses to HIV/AIDS (Mykhalovskiy, Patten, Sanders, Bailey, & Taylor, 2009). The situated judgment of expert practitioners has been historically neglected as a site of knowledge generation (Flyvbjerg, 2001). We developed workshop activities to elicit participants' *theories-in-use*, which often differ from the explicit theories encapsulated in research and policy-making (Argyris, 1991). We describe this as tacit knowledge acquired through—and encoded in *experienced practice in context* (EPIC). This framing elaborates on Flyvbjerg's notion of *phronesis* and acknowledges that contextual interventions are always, by definition, knowledge-generating (*pace* Rhodes & Lancaster, 2019).

Initial workshops took place over 2 days. We began with a hypothetical narrative prepared beforehand with a program staff member (Box 1). The narrative was written to relate the immediate aspects of service provision ('a fit pack across a counter') with broader contextual issues. In small groups, we discussed the narrative, asking participants to describe what aspects of the program and its context, in their understanding, are reflected in the narrative. We asked probe questions, such as 'how would a nonpeer service respond differently?' We invited participants to write their responses on sticky notes. We used the affinity method (Britz et al., 1999) to group related insights on craft paper. Giving each group a name, we used markers to begin tracing relationships between groups, forming an early prototype CLD. Overnight, the facilitators (GB and DR) transferred the prototype into a web-based diagramming tool (LucidCharts). This allowed real-time revision, extension, and refinement of the diagram in dialogue with participants on the second day of the workshop. A simplified version of the CLD developed for the peer-based NSEP and harm reduction service is shown in Figure 1 below.

Causal loop diagrams

Causal loop diagrams provide a visual language for tracing multiple one-way relationships between variables in the situation being mapped. Lines are labeled according to the

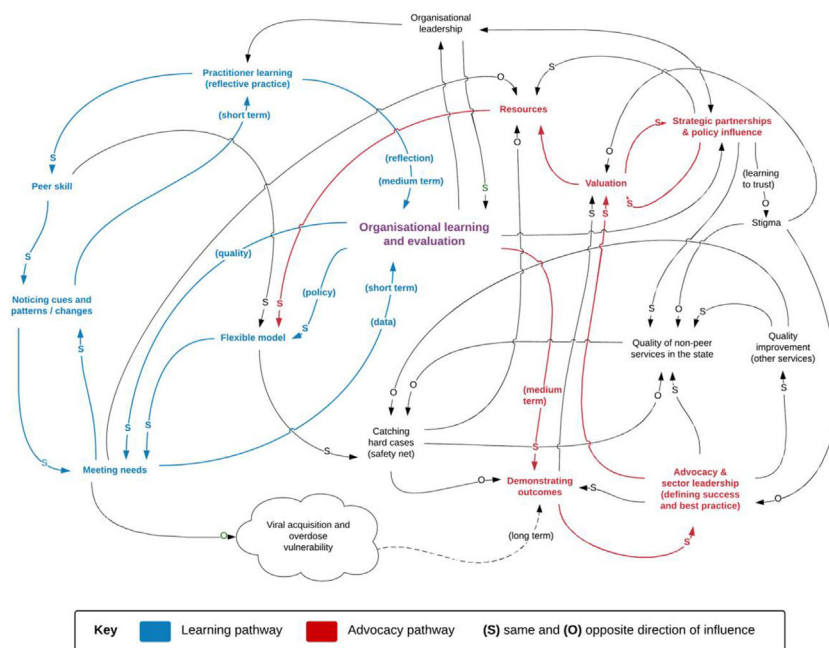


FIGURE 1

correlation of the linked elements over time: 'S' if the destination changes in the *same* direction as the source, and 'O' if it changes in the *opposite* direction. As the diagram fills out, cycles begin to appear, where the output from one variable flows through other variables back to its starting place. The resulting feedback loops can be labeled according to their effects, whether positive (accelerating) or negative (balancing). As we discuss below, decisions we made to accommodate our participants' feedback made it necessary to adapt these conventions and develop new but equally effective strategies for validating and simulating our models.

These simple conventions work to visualize emergent structures and functions of the system represented in the diagram as it plays out over time. It may also be possible to identify leverage within the system, where small changes in the strength of a local variable may drive larger changes across the system overall. We originally intended to develop proxy measures for the key variables and explore computer simulation of the models. By identifying quantitative proxy measures for each variable, it is possible to write equations that allow computer simulation of the system represented in the diagram (Maani & Cavana, 2000).

From the first workshop, however, our participants indicated that strict adherence to CLD conventions made it impossible to represent their understanding of what matters *most* to evaluate PCL interventions. Participants made two important objections. First, some variables which they viewed as crucial could not be quantified, nor could any meaningful proxy be identified. To illustrate, we can give the example of *peer skill* as it emerged in our first workshop. This concept turned out to be foundational in all three program theories, including at the highest level of federal policy advocacy. Second, not all linkages could be labeled S/O, as some represent flows that are contingent on their effects. Participants were adamant there *was* a link, but when we asked what the direction of influence was, they were equally adamant, 'it depends.' Some linkages in Figure 1 have annotations in

brackets that signal the nature of the contingent social and material process that translates influence from one place in the map to another.

At our first workshop, NSEP staff and stakeholders agreed that the program's strength was its unparalleled access to real-time insights into emerging changes in illicit drug use, such as changes in street drug markets policing. They described how the program verified and 'packaged' these insights to influence the policy arena (the red pathway in Figure 1). None of this worked, they argued, if program staff did not possess *and refine* peer skill (blue pathway in Figure 1). Whereas the literature on peer interventions sees 'a peer' as a person with similar personal attributes and experiences, our participants defined peer as a verb—the skilled ability to use personal experience (even when different) to establish rapport and work effectively with a contact or client. Peer skill allowed program staff to gather subtle insights into the broader system experienced by their clients/contacts.

This two-way exchange of knowledge makes the encounter more egalitarian and transforms peer practice into a crucial site of EPIC knowledge generation. This insight was foundational for our project, and it is one we would have missed had we insisted on sticking to the 'rules' of STM. For instance, peer skill enables developing a collective perspective that informs effective peer leadership in our program theory of a federal policy network. The discourse of consumer representation suggests individual people can contribute to policy-making by sharing their own personal lived experience as committee members, conference speakers, and consultation participants (CHF, 2018). Our participants noted this practice could assign outsize influence to charismatic speakers with powerful personal narratives. By contrast, peer leadership is grounded in a detailed understanding of the diverse and dynamic needs, identities, and experiences within the broader HIV-positive community in Australia.

Adapting the CLD method meant finding alternative ways to achieve later steps in the STM process, as sketched out by Maani and Cavana (2000). In particular, computer simulation makes it possible to identify system behaviors that are otherwise difficult to predict. As well, quantitative outputs from a simulation can be compared to real-world data, providing a way to validate the model. However, our maps contained qualitative elements and contingent relationships—correlations, neither positive nor negative. We, therefore, developed alternative methods to validate the models and identify relevant but nonobvious strategic dynamics.

Following initial workshops, the first and second authors (DR and GB) drafted accompanying text for the CLDs, writing definitions for each element that captured workshop discussions. Where possible, causal loops were identified and labeled. In contrast to the neat cycles that appear in conventional CLDs, we identified extended casual *pathways* representing 'strategic dynamics'—latent issues and challenges that might emerge over the longer term in the systems we had mapped. Two pathways are highlighted in Figure 1. The learning pathway, drawn in blue, shows how insights captured via peer skill are vital for practitioner learning and service provision refinement. An advocacy pathway, in red, shows how formal organizational learning allows insights from peer practice to be translated into influence in policy networks and the prevention system.

To validate the models, we held follow-up workshops with initial participants. We used a 'gaming' method (cf. Cleland, 2017), picking an issue from practice and a starting place on the map, and then following the pathways laid out by arrows and items—much like 'snakes and ladders'—discussing how that issue would 'play out' on the map. We tested hypotheticals, asking 'if this element suddenly stopped working, what might happen elsewhere in the system?' We made changes where participants, drawing on their shared and personal EPIC knowledge, felt the outcomes predicted were not true to life. The strategic dynamics were widely felt to be credible, with one organization leader calling them 'horrifyingly accurate'

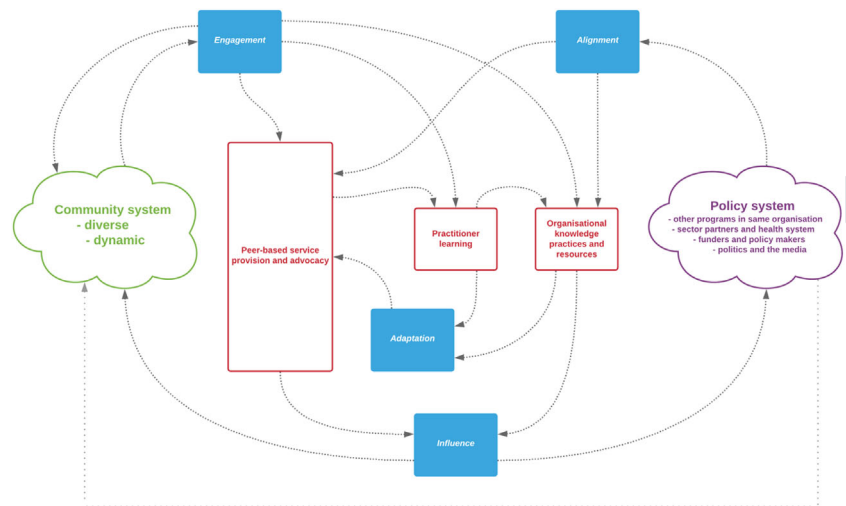


FIGURE 2

issues that kept him awake at night. This approach offered an effective and equally valid alternative to the more quantitative system dynamics approach's simulation methods.

Developing the mid-level theory and monitoring and evaluation framework

Following the three program theories' development and validation, we set out to analyze the outputs, hoping to develop a mid-level theory of the peer approach more generally. A critical, reflexive inquiry is the hall-mark of social research, yet STM models are viewed as straightforward representations of reality (Maani & Cavana, 2000, p. 20). Just as complexity thinkers argue, we found that conventional analytic methods destroyed the relational context that gave meaning and materiality to each element of our CLDs (Luke & Stamatakis, 2012). Success only followed when we adopted Pawson and Tilley's (1997) recommendations for analyzing the models developed across our multiple cases to develop 'mid-level' theory. After diving into complexity theory, we asked a theoretical probe question drawing on ecological systems theory (Holling, 1973). Imagining the program as a system, we asked: what functions would it need to have to survive under the conditions of continual uncertainty created by the diversity and dynamism of its environment? This is a question faced by all contextual interventions in public health and clinical practitioners, funders and policy-makers, and even social researchers, as they all seek to understand how networks, cultures, and communities transform themselves in response to health issues and emerging prevention opportunities. The answer is a mid-level theory (see Figure 2) that identifies four key functions—*engagement*, *adaptation*, and *alignment*, which permit sustained *influence* on community and policy systems.

Based on these functions (see Table 2), the authors developed a monitoring, evaluation, and learning (MEL) framework, which permits identifying program-specific indicators, data types and sources, and collection methods. Rather than seeking to isolate and measure an intervention's effects upon its context (i.e., the attribution paradigm described by Finegood), the framework identifies two goals for evaluation. First, it aims to inform observers' *confidence* that the program can fulfill the four key functions, enabling the

Q5

TABLE 1 Elements of the mid-level theory

Community system	A complex adaptive system (Holland, 1995) in which agents interact according to broad 'rules' generating diversity and dynamism.
Policy system	A complex system of multiples streams including problem, policy and politics (Kingdon, 1995).
Red boxes	Activities of the contextual intervention.
Blue boxes	Four essential functions for maintaining influence and relevance in a diverse and dynamic environment (see Table 2 for definitions).

TABLE 2 Four essential functions

Function	Definition
Engagement	How the program maintains up-to-date mental models of the diverse and dynamic needs, experiences, and identities in target communities.
Alignment	How the program picks up signals about what is happening in its policy system and uses them to reorient service or advocacy priorities.
Adaptation	How the program changes its approach based on mental models refined according to new insights from engagement and alignment
Influence	<i>On communities</i> —how the program taps into communities' existing ways of doing things to promote new ways of doing things. <i>On policy</i> —how the program achieves or mobilizes influence on processes and outcomes within its policy environment.

program to respond effectively to its changing context. Second, both the program itself and the evaluation elicit substantive knowledge that can guide adaptation in the broader prevention system. Indeed, if PCL programs are not playing this role, generating knowledge that can inform partnerships, program development, and policy-making, we contend that governments are not getting the full value. In subsequent phases of the project (see Brown et al., 2018), the framework has been piloted with additional organizations that did not participate in the original workshops, providing additional support for validity and opportunities to refine the approach.

REFLECTIONS

In this section, we reflect critically and appreciatively, on our experience of using CSTM as research methods in a participatory evaluation research project. Acknowledging there is an enormous diversity of methods and approaches in the systems field (Byrne, 1998), these remarks focus on how STM such as CLD modeling can fit into scholarly research projects and whether it can function as a bridge between academia, policy and practice.

Strengths

With necessary adaptations (see below), the approach produced outputs with very high fidelity, as judged by participants in our initial workshops, and surface validity, as judged by nonparticipant stakeholders. The resulting program theories offered clear and useful insights for policy, funding, planning, and evaluation. Our hypothesis that program staff and EPIC knowledge would have more sophisticated PCL theories than are found in peer-reviewed research was borne out. The subtlety and contingency of the models we obtained

suggest that CSTM offers an alternative to intervention science approaches that overlook, distort, or extinguish the relational mechanisms that make the peer approach effective.

Difficulties

Advocates of STM, such as Peter Senge and Donella Meadows, present systems thinking as a characteristic mode of cognition induced by using STM techniques in groups of practitioners and policy-makers. If it happened at all in our workshops, the ‘a-ha!’ moment came very late in the process—once the maps had been mostly constructed. We relied on enormous trust from our participants, who were willing to keep going until it made sense.

At different times it seemed the onto-epistemology that underlies STM is either thin, or deeply implicit, or perhaps even nonexistent. We read widely across the CSTM tradition (Cabrera, Colosi, & Lobdell, 2008; Checkland, 1998; Kirkwood, 1998; Maani & Cavana, 2000; Meadows & Wright, 2009; Richardson, 1991; Waring, 1996; Maani, 2009; Senge, 2010; Ulrich & Reynolds, 2010). As distinct from the broader field of complexity science, the STM approach characterized by Senge and Meadows is onto-epistemologically minimal (or perhaps just atheoretical). This is intentional, allowing its application across vastly different substantive fields (Herbert Simon, 1962). However, it became an acute challenge for our project at two points in particular. First, although we adopted a systematic guide (Maani & Cavana, 2000), when we needed to adapt the methods to accommodate our participants’ feedback, we lacked guidance on how to make changes while maintaining the method’s integrity. Second, as mentioned above, we found little guidance in the STM toolkit on analyzing the program theories we had articulated in the CLD format, and there was a significant pause in the project. At the same time, we undertook a ‘deep dive’ into the field of complexity studies to identify relevant theoretical resources.

Ways forward: CSTM as an evaluation research methodology

As any postgraduate research methods course will emphasize, the methodology is more than methods; it includes a theoretical framework that outlines the ontological and epistemological assumptions that guide their application. This section offers some brief methodological pointers for evaluation researchers and practitioners who might seek to follow in our footsteps.

Rather than viewing CLDs as straightforward representations of reality, we assumed that our participants were relating *partial* insights, from diverse and particular perspectives, into actually-existing, complex social systems—consistent with the critical realist tradition of social explanation (Archer, 1996; Elder-Vass, 2010; Little, 2012). Verifying the characterizations offered by our participants and partner programs was beyond our project’s scope, but we still required a minimal conception of what it was these programs were engaging with. We conceptualized the ‘community system’ as a complex adaptive system (Holland, 1995; Ostrom, 1999), and for the ‘policy system,’ we drew on Kingdon’s multi-stream model (1995). To conceptualize the linkages made visible by our methods and the practices by which knowledge is translated into influence, we used the sociology of translation (Callon, 1986). This approach recognizes that knowledge and influence do not simply flow, unmodified, within and between systems, but are instead mediated and translated via chains of association between human and nonhuman actors. These resources allowed us to adapt our methods when required by feedback from our participants and achieve the rigor and reflexivity that are characteristic requirements of social research.

CONCLUSION

In the W3 project, we piloted methods from STM and drew on social theory and scientific studies of complexity to help us adapt the methods and analyze their outputs for the purposes of theory-building. Although not without its challenges, this allowed us to articulate explicitly how staff and stakeholders in peer-based and community-led programs tacitly understand how this class of interventions works. This offers an innovative solution to a long-standing problem in implementation studies: evaluating contextual interventions that are enmeshed and embedded in the social systems they are seeking to influence and even transform. Our findings are program theories that can inform the theory-based evaluation. They highlight the value of experience practice in context. Similarly, they may suggest mechanisms for studies using realist evaluation and synthesis (Pawson & Tilley, 1997; Pawson, Greenhalgh, Harvey, & Walshe, 2004).

Our findings emphasize the social processes through which tenuous, partial, and emergent insights, acquired through experienced practice in context, are made to matter in public health policy-making and community-based health promotion. These insights are only possible through the patient and generous participation of the staff and stakeholders of our partner programs, who stayed the course even as we fumbled through our first experiences of CSTM in practice. If our findings demonstrate one thing, we hope they highlight the quiet brilliance of practitioners and the value of research-practice-policy partnerships.

REFERENCES

- Archer, M. S. (1996). *Culture and agency: The place of culture in social theory*. Cambridge University Press.
- Argyris, C. (1991). Teaching smart people how to learn. *Harvard Business Review*, 69, 99–109.
- Britz, G. C., Emerling, D. W., Hare, L. B., Hoerl, R. W., Division, A. S., for, Q. A. S., ... Shade, J. E. (1999). *Improving performance through statistical thinking*. ASQ Quality Press.
- Brown, G., Johnston, K., Ellard, J., & Carman, M. (2013). Evidence synthesis and application for policy and practice project: Full report. Australian Research Centre in Sex, Health and Society, La Trobe University
- 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, 231. <https://doi.org/10.3389/fpubh.2018.00231>
- Byrne, D. S. (1998). *Complexity theory and the social sciences: An introduction*. London ; New York: Routledge.
- Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning*, 31, 299–310. <https://doi.org/10.1016/j.evalprogplan.2007.12.001>
- Callon, M. (1986). Some elements of a sociology of translation: Domestication of the scallops and the fishermen of St. Brieuc Bay. *Power, action, and belief: A new sociology of knowledge*, 32, 196–223.
- de Carvalho, J. (2013). *Program logic for the twenty first century: A definitive guide*. Xlibris Corporation.
- Chambers, D. A., & Norton, W. E. (2016). The adaptome: Advancing the science of intervention adaptation. *American Journal of Preventive Medicine*, 51(4 Suppl 2), S124–S131. doi: [10.1016/j.amepre.2016.05.011](https://doi.org/10.1016/j.amepre.2016.05.011)
- Checkland, P. B. (1998). *Systems thinking, systems practice*. Wiley.
- Chen, H. (2005). *Practical program evaluation: Assessing and improving planning, implementation, and effectiveness*. Thousand Oaks, Calif: SAGE.
- CHF, (2018). Consumer representation—the big picture. Consumers Health Forum of Australia. <https://chf.org.au/guidelines-consumer-representatives/consumer-representation-big-picture>
- Cleland, D. (2017). A playful shift: Field-based experimental games offer insight into capacity reduction in small-scale fisheries. *Ocean & Coastal Management*, 144, 129–137. <https://doi.org/10.1016/j.ocecoaman.2017.05.001>
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: The new Medical Research Council guidance. *BMJ*, 337, a1655. <https://doi.org/10.1136/bmj.a1655>
- Elder-Vass, D. (2010). *The causal power of social structures: Emergence, structure and agency*. Cambridge; New York: Cambridge University Press.
- Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge University Press.
- Green, S., & Higgins, J. P. T. (2008). Reviews in public health and health promotion. In *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley.

- Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: How “out of control” can a randomised controlled trial be? *BMJ*, 328, 1561–1563.
- Holland, J. H. (1995). *Hidden order: How adaptation builds complexity*. Perseus Books.
- Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4, 1–23. <https://doi.org/10.1146/annurev.es.04.110173.000245>
- Johnston, L. M., Matteson, C. L., & Finegood, D. T. (2014). Systems science and obesity policy: A novel framework for analyzing and rethinking population-level planning. *American journal of public health*, e1–e9. Q6
- Foundation, K. (2004). *Logic model development guide*. Michigan: W.K. Kellogg Foundation.
- Kingdon, J. W. (1995). *Agendas, alternatives and public policies*. New York: Harper Collins.
- Kirkwood, C. W. (1998). *System dynamics methods: A quick introduction*. Tempe, AZ: Arizona State University.
- Little, D. (2012). Analytical sociology and the rest of sociology. *Sociologica*, 6. Q7
- Luke, D. A., & Stamatakis, K. A. (2012). Systems science methods in public health: Dynamics, networks, and agents. *Annual Review of Public Health*, 33, 357–376. <https://doi.org/10.1146/annurev-publhealth-031210-101222>
- Maani, K. (2013). *Decision-making for climate change adaptation: A systems thinking approach*. National Climate Change Adaptation Research Facility, Brisbane: University of Queensland.
- Maani, K. E., & Cavana, R. Y. (2000). *Systems thinking and modelling: Understanding change and complexity*. Pearson Education.
- Meadows, D. H., & Wright, D. (2009). *Thinking in systems: A primer*. London: Earthscan.
- Mykhalovskiy, E., Patten, S., Sanders, C., Bailey, M., & Taylor, D. (2009). Beyond buzzwords: Toward a community-based model of the integration of HIV treatment and prevention. *AIDS Care*, 21, 25–30. <https://doi.org/10.1080/09540120802068753>
- Ostrom, E. (1999). Coping with tragedies of the commons. *Annual Review of Political Science*, 2, 493–535. <https://doi.org/10.1146/annurev.polisci.2.1.493>
- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). *Realist synthesis: An introduction*. ESRC Research Methods Programme. Manchester: University of Manchester.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. SAGE.
- Richardson, G. P. (1991). *Feedback thought in social science and systems theory*. University of Pennsylvania Press.
- Rhodes, T., & Lancaster, K. (2019). Evidence-making interventions in health: A conceptual framing. *Social Science & Medicine*, 238, 112488. <https://doi.org/10.1016/j.socscimed.2019.112488>
- Senge, P. M. (2010). *The fifth discipline: The art & practice of the learning organization*. Crown Publishing Group.
- Taylor, D., Bury, M., Campling, N., Carter, S., Garfield, S., Newbould, J., & Rennie, T. (2006). *A Review of the use of the Health Belief Model (HBM), the Theory of Reasoned Action (TRA), the Theory of Planned Behaviour (TPB) and the Trans-Theoretical Model (TTM) to study and predict health related behaviour change* (pp. 1–215). London, UK: National Institute for Health and Clinical Excellence.
- Turner, G., & Shepherd, J. (1999). A method in search of a theory: Peer education and health promotion. *Health Education Research*, 14, 235–247.
- Ulrich, W., & Reynolds, M. (2010). *Critical systems heuristics. In Systems approaches to managing change: A practical guide*. London: Springer Science & Business Media.
- Waring, A. (1996). *Practical systems thinking*. Cengage Learning EMEA,
- Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. In D. J. Rog & D. Fournier (Eds.), *New Directions for Evaluation: No. 76. Special issue: Progress and future directions in evaluation: perspectives on theory, practice, and methods* (pp. 41–55). Jossey-Bass. <https://doi.org/10.1002/ev.1086>

How to cite this article: Reeders, D., & Brown, G. (2021). Using systems methods to elicit complex program theories. *New Directions for Evaluation*, 1–11. <https://doi.org/10.1002/ev.20467>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

UNCORRECTED PROOFS