

Research article

Embedding evaluation into a policy-practice-research partnership using theory of change: case study of Designing London's Recovery Programme

Gemma Moore^{1,*}  and Julie McLaren² ¹UCL Bartlett Institute of Environmental Design and Engineering, London²UCL IIPP (UCL Institute for Innovation and Public Purpose), London*Correspondence: gemma.moore@ucl.ac.uk

Submission date: 23 January 2024; Acceptance date: 2 July 2025; Publication date: 16 September 2025

How to cite

Moore, G. and McLaren, J. (2025) 'Embedding evaluation into a policy-practice-research partnership using theory of change: case study of Designing London's Recovery Programme'. *Research for All*, 9 (1), 11. DOI: <https://doi.org/10.14324/RFA.09.1.11>.

Peer review

This article has been peer-reviewed through the journal's standard double-anonymous peer-review process, where both the reviewers and authors are anonymised during review.

Copyright

2025, Gemma Moore and Julie McLaren. This is an open-access article distributed under the terms of the Creative Commons Attribution Licence (CC BY) 4.0 <https://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, distribution and reproduction in any medium, provided the original authors and source are credited • DOI: <https://doi.org/10.14324/RFA.9.1.11>.

Open access

Research for All is a peer-reviewed open-access journal.

Abstract

Experimentalism in transformative policy and practice requires a different form of delivery that emphasises collaboration and learning. The Designing London's Recovery Programme was a two-year programme that invited innovators to co-create transformative solutions to address local missions for change. It was delivered through a novel partnership between policy, practice and research. As 'learning partners' embedded in the programme, the research team led on a developmental and democratic evaluation approach which informed the process of developing, applying and refining a programme level theory of change as a form of embedded practice. Our multi-method approach was informed by participatory action research. A theory of change was developed at the outset of the programme that outlined activities, programme outcomes, transformative outcomes and overarching goals. Then it was applied during delivery, at a programme level, and used as a tool for learning and collective reflection. Insights from the reflection led to its refinement. The theory of change was used to convene people, build consensus and identify shared goals at the outset. By using a 'transformative frame' it provoked discussion about wider systems change. The theory of change took on a 'navigational' role

throughout the programme, however it lacked democratic ownership and as a result wasn't used to inform decision making about implementation, where other factors were prioritised.

Keywords policy-research-practice partnership; developmental evaluation; theory of change; transformative change; experimentation and learning

Key messages

- Different approaches to evaluation are needed for transformative practices that bring together different partners. Theory of change can be useful for bringing together different perspectives to identify shared goals and as a navigational tool for reflection. A theory of change focused on the ways of working together rather than specific goals can provide a compass for designing systems change.
- Embedding evaluation presents challenges in practice due to competing and changing priorities between partners. Such a theory of change cannot work unless it is part of a wider learning culture within organisations with the freedom to experiment and innovate.
- A distributed learning model may benefit from structures and formal activities (such as workshops, seminars and case studies) alongside more informal activities (such as mentoring and digital platforms).

Introduction

Navigating complex societal challenges like climate change, health pandemics and social inequalities poses significant challenges for policy makers who are needing to devise policies and shape regulation to guide transformative change, defined as 'deep and sustained non-linear change' (Linnér & Wibeck, 2020). There has been a growing interest in policy making, towards transformative innovation policies (TIPs), which aim to support the required shifts to current socio-technical systems in order to address those complex challenges (Schot & Steinmueller, 2018; Weber & Rohrer, 2012). In practice, this often involves collaboration with a range on stakeholders – including policy-practice-research partnerships – and an emphasis on facilitating societal experimentation and learning. An example of how this approach is being translated at local level is the Designing London's Recovery Programme (DLRP), an open innovation challenge programme designed to address a set of London-wide missions in the aftermath of the Covid-19 pandemic, and the focus of the research presented in this paper.

There remain significant gaps in understanding about how to implement partnerships at local levels to deliver policies aimed at transformation, and the capabilities needed in local government to experiment with these approaches in ways that help adapt policy implementation in real time. In principle, evaluation and learning are an integral part of design and implementing TIP (Molas-Gallart et al., 2021; Rohrer et al., 2023). However, approaches to evaluation can vary, from managerial, technical models of evaluation focused on demonstrating accountability, to more developmental, learning-oriented approaches. In this paper we explore the role of evaluation in a TIP case study, specifically, the use of 'theory of change' (ToC). The research question at the heart of this paper is whether ToC enables the adoption of a more collaborative, accessible, learning approach to evaluation that is more attuned to supporting transformative policy making, delivered through policy-practice-research partnerships. We respond to this question by sharing the process of co-creating a ToC as embedded practice in the DLRP. We discuss the process of developing a ToC at the outset of the programme and the practicalities of using it, alongside examining the ways in which the ToC, as a tool, supported a different approach to evaluation and the broader role and usefulness of ToC in supporting experimentation and learning.

Background: evaluating differently in transformative innovation

Transformative innovation policy and experimentalism

The narrative of 'transformative change' has become popular in both policy and research discourses to indicate a break from the status quo and the need for deep, structural shifts in our systems, institutions and daily practices towards more sustainable, equitable outcomes (Hölscher et al., 2018; Köhler et al., 2019). Transformative innovation calls for a different approach to governing change with an emphasis on collaboration and a more experimentalist mode of action (Sabel & Zeitlin, 2012). This approach places as much emphasis on the *means* of transformative change (i.e. the practices and processes that contribute to catalysing and implementing change), as on the end *goals* or missions that provide the direction.

While national governments have thus far been the dominant players in this shift towards more transformative action frames, there is a growing interest in missions and challenge-based approaches at regional levels, reflecting the idea of a place or context-specific partnership-focused approach to goal setting and problem solving (Pontikakis et al., 2022). However, in the UK context, where the case study for this paper is set, the current political context means local or regional state actors have limited policy instruments to influence the design and implementation of policies in ways that they would like (Wanzenböck & Frenken, 2020). Policy makers are being creative in how they use existing levers – across policy, procurement and regulation – to shape policy for the local context, and who they work with to do this. Examples, including policy-practice-research partnerships, enable the bringing in of different expertise and experience ensuring alignment of research and practice that can also benefit policy, such as the London Research and Policy Partnership.

Towards developmental, democratic and learning-oriented evaluation

As policies become more directional and focused on transformative change, there is a need to reconfigure evaluation practice to meet the needs of complex, often experimental, long-term programmatic approaches (Rohracher et al., 2023). These programmes pilot new ways of working with partners across the system(s) towards system change, challenging linear, output-focused measures of 'instrumental' impact. Much of the conventional wisdom on policy evaluation is framed around 'objectively' assessing 'interventions' and capturing instrumental changes, closely tied with economic growth, behaviour change and product-orientated innovation and scaling (BEIS, 2020). TIP invites us to reframe thinking on evaluation towards understanding the multiple long-term, open-ended potential strategies and outcomes that are undefined at the start (Hargreaves, 2010). As a result, there is a need to take a different approach to thinking about the logics of change, what can be delivered or measured, by whom, and in what time frame.

There are three strands of evaluation approaches that provide an indication of their respective benefits to being applied to TIPs and the experimental nature of their governance.

First, evaluation and learning processes are considered integral within the delivery of TIPs (Molas-Gallart et al., 2021). However, often policy evaluation is undertaken as an 'objective' exercise outside the programme, though researchers point out that in fact evaluations are very political and influenced by different interests and power dynamics (e.g. Hildén et al., 2014). Evaluation can be thought of as within the programme – working with those involved to understand what happened, what changed and why – drawing upon a developmental evaluation approach. Developmental evaluation is 'grounded in Systems Thinking and supports innovation by collecting and analysing real-time data in ways that lead to informed and ongoing decision-making as part of the design, development, and implementation process' (Quinn Patton, 2010, as cited in Coffman et al., 2023, p. 153). As an approach, it is gaining interest, being used in contexts where innovation is taking place and where the outcomes and pathways of change are not entirely clear.

Second, policy evaluation is by definition a *political exercise*. Democratic and participatory approaches to evaluation can bring to the fore where power lies, and different motivations that are at play in deciding who 'judges' what for what 'purpose' (Estrella & Gaventa, 1998; Floc'hlay & Plottu, 1998;

Plottu & Plottu, 2011). A prescriptive model of evaluation, with specific indicators, may not be appropriate or useful for TIPs. Instead, the evaluation process benefits from the inclusion of key actors from across the partnership or sectors, to prioritise what is meaningful to those involved, and thereby contributing to a broader evidence base and deeper understanding of change within the system. As Völker et al. (2020) suggest, the evaluation processes can be considered sites of 'collective imagining'.

Finally, evaluation for critical reflection and co-learning is of relevance for TIPs, which frame and capture learning as an ongoing endeavour, central to ensuring that a programme aimed at addressing the mission or goals stays relevant and adaptive to the challenge (Rohracher et al., 2023). Through evaluative processes and critical reflection, the evaluation approach can play a key role in enabling those delivering the policy to reflect on and improve activity, enabling change within the programme. If this is undertaken by, with and for multi-actors, it can provide the basis for co-learning (Hanberger, 2006; Pineo et al., 2021).

Considering these three strands, the relationship between the evaluation and a TIP programme shifts from that of a management-orientated approach seeking evidence of the impact of policy programmes, to a learning-orientated approach embedded in such programmes and used as a tool to share reflections and set or reset direction with programme actors.

Theory of change

ToC has been a popular model or tool within policy and complex programmes evaluation (see Breuer et al., 2015 and Moore et al., 2021 for examples). A ToC is used as a tool to link programme inputs and activities to a chain of intended (or unintended) outputs and observed outcomes, and then using the model to guide evaluation (Rogers, 2008). This process, usually, works backwards from a clear goal and outcomes to establish factors, preconditions and activities needed to reach that goal and desired outcomes. It requires a discussion of core assumptions that the programme is making about the activities being undertaken and their intended effects. ToCs are typically developed through a structured process involving consultation with partners, alongside problem analysis through drawing upon a range of evidence sources. In collaborative contexts, ToCs are co-created, involving the integration of diverse perspectives (Moore et al., 2021). Within such contexts the ToCs may focus on continuous learning and reflection.

The ToC approach remains a popular and useful framework for evaluating more complex systemic change due to sensitivity to context, values of the policy intervention and ability to accommodate multiple potential pathways of change. Evaluating innovation programmes aimed at systems change benefits from a more flexible approach, ideally one that integrates learning and identifies a broad spectrum of potential transformative outcomes and uses a nested approach to assess multi-level changes (Ghosh et al., 2021; Molas-Gallart, 2015; Molas-Gallart et al., 2021). TIP is still in relative infancy and gaps in understanding remain, including the practical use of evaluation and ToC in the design and delivery of policy that is geared towards long-term, transformative change. Examples are emerging from practice, for example Palavicino et al. (2023) share their experiences of co-creating and testing a 'Transformative Theory of Change' for 'Climate-KIC projects'. They urge the need to explore the lessons learned when applying ToCs. Informed by this knowledge, we identified a need for research to explore how ToCs can contribute to the practical implementation and evaluation of the DLRP.

The Designing London's Recovery Programme

The DLRP was an innovation challenge programme that ran between March 2021 and November 2022 (about 20 months) and was sponsored and led by the Greater London Authority (GLA). The programme was one of several initiatives designed to deliver on the Mayor's London Recovery Strategy. The key aim of DLRP was to co-create design-led solutions to the city's most pressing challenges. DLRP focused in on four of the nine missions set out in the London Recovery Strategy. They were: 1. Helping Londoners into Good Work; 2. Building Stronger Communities; 3. High Streets for All; and 4. Green New Deal. These formed innovation briefs for the DLRP with Good Work and Building Stronger Communities forming the

first two briefs and Green New Deal and High Streets being combined into a third brief. It was delivered via a novel policy-practice-research partnership between the GLA, the Design Council and the Complex Urban Systems for Sustainability and Health (CUSSH) research project and was funded by the London Economic Action Partnership. CUSSH was a consortium that aimed to explore how evidence and research engagement in decision processes can strengthen the envisioning, formulation and implementation of health and sustainability policies in cities.

Together these three 'organisations' formed a programme delivery and coordination layer – it included policy officers and mission leads from the GLA, design experts and project managers from the Design Council and researchers from CUSSH who acted as a 'learning partner'. The policy-practice-research partnership is illustrated in [Figure 1](#).

DLRP was delivered across two phases. Within Phase 1, 20 innovation teams were selected from a competitive open call, which attracted over 120 proposals, to join the first phase of the programme to develop their ideas and receive support from the programme partners via a series of development workshops focused on design thinking and systems mapping. In between workshops projects had additional support from the design experts who acted as coach and guide to individual projects. For both workshops and the coaching, innovation teams were grouped according to the GLA mission to which their project most closely aligned. This was to build more of a cohort effect, encouraging interaction and shared learning among projects addressing similar goals. Phase 1 workshops ran between July and November 2021.

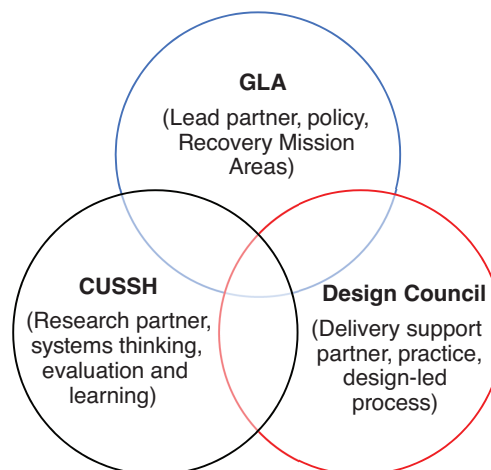
Eleven teams progressed to Phase 2 (workshops ran April–November 2022), which focused on prototyping and testing the solutions out with communities in London. There was a competitive funding call, with the teams applying for funding to support their Phase 2 participation. The timing of the programme during the Covid-19 pandemic meant that the workshops were all held virtually using Zoom and Miro to create interactive sessions. However, the projects participating in the second phase came together to showcase their final prototype designs at an exhibition held at the Victoria and Albert Museum as part of the 2022 London Design Festival.

Research approach

Our role as 'learning partner' evolved over time but essentially started with the challenge of embedding a developmental, democratic, learning-oriented evaluation of the programme, within the partnership. The key aims were to:

- develop, through participatory process, a ToC for a transformative, open innovation programme in London;

Figure 1. Policy-practice-research partnership in DLRP



- apply the ToC, as a tool for integrating evaluation and learning within the delivery of a policy programme; and
- refine the ToC as a form of embedded, evaluative practice.

Our research approach, and the processes adopted to develop, apply and refine the ToC, were informed by participatory action research (PAR). A PAR approach enabled the authors to tap into the intricacies of the dynamics between the stakeholders involved in the programme. As researchers participating and embedded in the policy process, we therefore were participants in the design and delivery of the programme, throughout the length of the programme. Reason and Bradbury (2008) argue that to conduct participatory action research, researchers have to employ a worldview or mindset, which ‘asks us to be both situated and reflexive, to be explicit about the perspective from which knowledge is created, to see inquiry as a process of coming to know, serving the democratic, practical ethos of action research’ (p. 7). Our PAR approach follows the evaluative methodology adopted within the wider transdisciplinary, CUSSH project, in which the researchers were members.

Our methodological stages: development, application and reflection

The scoping, development and application of the ToC for DLRP had a number of key stages which involved both planned, formal mechanisms, and informal responses and opportunities, fitting in with the wider delivery of the programme. The development, application and reflection on the ToC was shaped by wider practices of our role as ‘learning partner’ within the implementation of the policy programme. Table 1 provides details of our overall methodology, framed by the three key stages to (a) co-develop, (b) apply and (c) reflect on the ToC, which informed its refinement. Within Table 1 we also highlight the methods relevant for this paper.

Our methodological processes involved working throughout the programme which ensured that the practice of research and evaluation was active, rather than passive, and gave shape to knowledge creation and the context (Cook & Wagenaar, 2012).

Data collection and analysis

The analysis of the data (i.e. transcripts, survey results and our reflections) identified an overarching thematical framework, drawn from the ToC, which was used to initially group and code the data. This included overarching predefined themes used to sort and group the data into categories (i.e. activities, outcomes). We analysed the data in Nvivo (QSR International Pty Ltd, version 12.6.0, 2019) using thematic analysis (Braun & Clarke, 2006; Nowell et al., 2017). We then used deductive and inductive coding (Fereday & Muir-Cochrane, 2006). We read the transcripts and coded the data, and inductively derived categories and codes were added in both stages. The data were grouped into overarching themes through an iterative process involving regular reflection between researchers regarding the interpretation of the data.

Table 1. The methodological stages

Wider methodological processes as learning partner	Our methodological stages (and dates)	Methods to develop, apply and refine the ToC
<ul style="list-style-type: none"> • Document review and analysis • Participatory workshops • Surveys • Weekly planning meetings • External events • Interviews • Systems sessions • Creation of exhibition • Production of film • Observations 	<p>Stage 1: Develop: co-creating ToC (April–September 2021)</p> <p>Stage 2: Apply: application, reflection and co-learning (September 2021–December 2022)</p> <p>Stage 3: Refine: end of programme evaluation (October–December 2022)</p>	<ul style="list-style-type: none"> • Document review • Three workshops • Survey with workshop participants • Three systems sessions with programme team • Observations (e.g. weekly planning meetings, workshops) • 13 interviews

Findings

Stage 1: Develop – co-developing the theory of change

The development of a ToC was through a collaborative process between members of the CUSSH project from UCL, the GLA and Design Council.

The aims of the 'develop' stage were to:

- build a working relationship across the policy-practice-research partnership;
- reach a shared understanding of the overall vision for the DLRP, the intended changes and work together to identify the impact pathways; and
- contribute different knowledges and evidence to inform the development of the ToC, opening up the research and policy process.

A series of three participatory online workshops and an online survey were undertaken between April and June 2021. The first workshop involved discussing the potential range of evaluation options within DLRP, the second on developing the ToC, and the third on getting feedback on the outline ToC. Between Workshops 2 and 3, a survey enabled participants to prioritise the outcomes generated at Workshop 2 (i.e. ranking the outcomes as high, medium, low).

There were 12 participants in total across the workshops, all of whom were linked to the planning or delivery of the programme. The workshop agendas involved a mix of activities, involving presentations and discussion to enabling partner and group activities. For instance, within Workshop 2 some example ToCs were presented to provoke a conversation on what might a ToC for the programme look like. This was followed by a structured exercise on a Google 'jamboard' to define the scope or boundaries of the ToC (see [Figure 2](#) for an example of the responses).

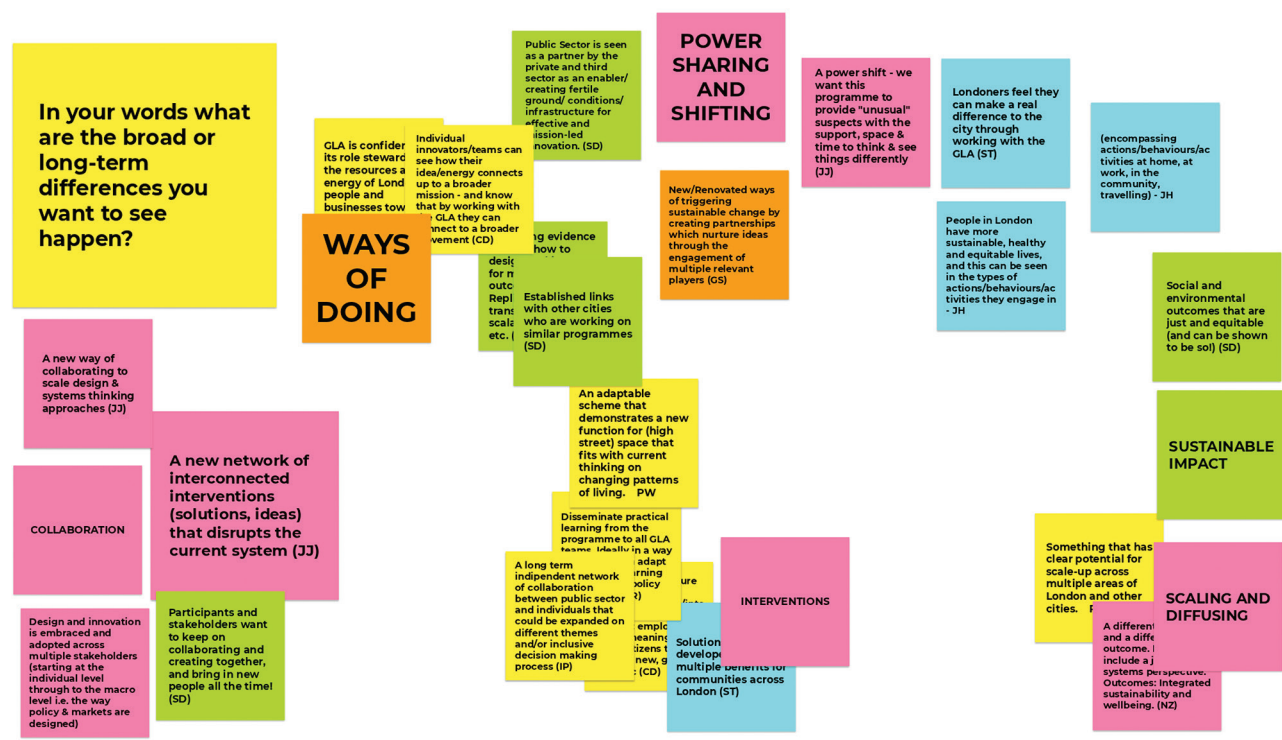
The survey was circulated to all 12 participants involved in developing the ToC, and 7 of the 12 responded. The results of the survey were summarised in a table and shared at Workshop 3. This led to creating a table of outcomes as the first draft of the ToC ([Table 2](#)). The workshop feedback and survey were collated into a single document and synthesised for recurring themes.

As part of the co-design process of developing the ToC, the delivery team decided the purpose of the ToC, and it was agreed that the ToC would:

- represent the dynamics of change, perceived as complex and adaptive;
- acknowledge the different levels and layers in the programme (i.e. missions, projects, teams, individuals);
- focus on transformative outcomes (as opposed to system impacts) to recognise that 'change' as something that isn't implemented or controlled but requires developing the conditions and capacities for change to unfold in the desired direction (i.e. towards the missions);
- capture both the processes and outcomes of the programme; and
- provide a tool for communication and reflection (within the team).

The development of the ToC enabled the team to articulate their aspirations to 'do things differently' in the delivery of the policy programme, for instance there was a strong motivation to change how power was shared within innovation programmes. The ToC captured the *means* or the *how* of transformative change rather than the end *goals* or missions. Different levels of outcomes were represented on the ToC including: transformative outcomes; long-term aspirations for change (e.g. power shifting, changes in ways of doing, network of connected interventions that disrupt the current system, Londoners living healthier and more sustainable lives); and intermediate outcomes (e.g. making connections between disciplines, sectors and communities; learning at project and programme levels including an understanding of what is needed in terms of building a fertile ground for systems change).

Figure 2. Example of responses from an exercise to generate outcomes for DLRP, using a jamboard



The transformative outcome layer set out *means* towards transformation, that is building capacity, sharing and shifting power, changing perceptions and building alternative practices, creating a culture of adoption and adaptation, and fostering learning. The long-term outcomes captured the team's ambitions for what the programme could contribute towards catalysing, and the intermediate outcomes articulated the steps to move towards the larger ambitions. Although the initial draft of the ToC was visualised in a table (shown in Table 2) the team was keen for the ToC not to be presented as a linear model of change, but a dynamic one that recognised the importance of a range of conditions needed for systems change. The participatory process of developing the ToC enabled the team to reach agreement on its role as a navigational tool to help guide DLRP rather than a map that outlines the pipeline from activities to outputs to outcomes.

The sessions to co-develop the ToC worked well to elicit reflections on the values and assumptions of the members of the delivery team; these underlying beliefs influenced and shaped the discussions, helping to surface diverse perspectives. The discussion also revealed varying evaluation needs and expectations, from GLA policy officers' interests, to funder requirements, the Design Council's questions and CUSSH's research priorities. In short there were competing priorities, which were captured within the resulting ToC. The outputs from this process informed a brief to a CUSSH colleague to design a visual illustration of the ToC which would also help to communicate its role as a navigational tool for the programme (Figure 3). Furthermore, an evaluation and learning plan was developed, tied to the outcomes within the ToC.

Stage 2: Apply – applying the theory of change

During the first phase of DLRP which ran from September to December 2021, the programme delivery partners came together in workshops called 'systems sessions'. These were led by the Design Council and took place in between the development workshops (with the innovation projects).

Table 2. Initial theory of change

Goals	Building capacity, sharing and shifting power	Changing perceptions and building alternative practices	Adoption and adaptation	Learning
Transformative outcomes	<p>The public sector is seen as a key partner by the private and third sector, and it is perceived as an enabler, creating the conditions and infrastructure for effective and mission-led innovation.</p> <p>All sectors (public, private and third) see the value of collaborating to enable systems change.</p> <p>New or refreshed ways of triggering sustainable change by creating partnerships which nurture ideas through the engagement of multiple relevant players.</p> <p>Steps taken to share power differently – the ‘unusual’ suspects have been provided with the support, space and time to think and do things differently.</p>	<p>London’s recovery is supported via innovations in design and action, by the way decisions are made and resources are used.</p> <p>There is a shift in societal priorities – the innovation teams create the right conditions for the broader movement: London’s recovery (economic, society and environmental).</p>	<p>Solutions that are developed yield multiple benefits for communities across London.</p> <p>A new network of interconnected interventions (solutions, ideas) disrupts the current system.</p> <p>Design and innovation are embraced and adopted across multiple stakeholders (starting at the individual level through to the macro level, i.e. the way policy and markets are designed).</p>	<p>The practical learning from the programme is shared to all GLA teams; the teams adapt and use the learning for a range of policy challenges.</p> <p>We, collectively, demonstrate a model to recovery that can be replicated by other cities across England.</p> <p>GLA’s role in design missions is recognised.</p>

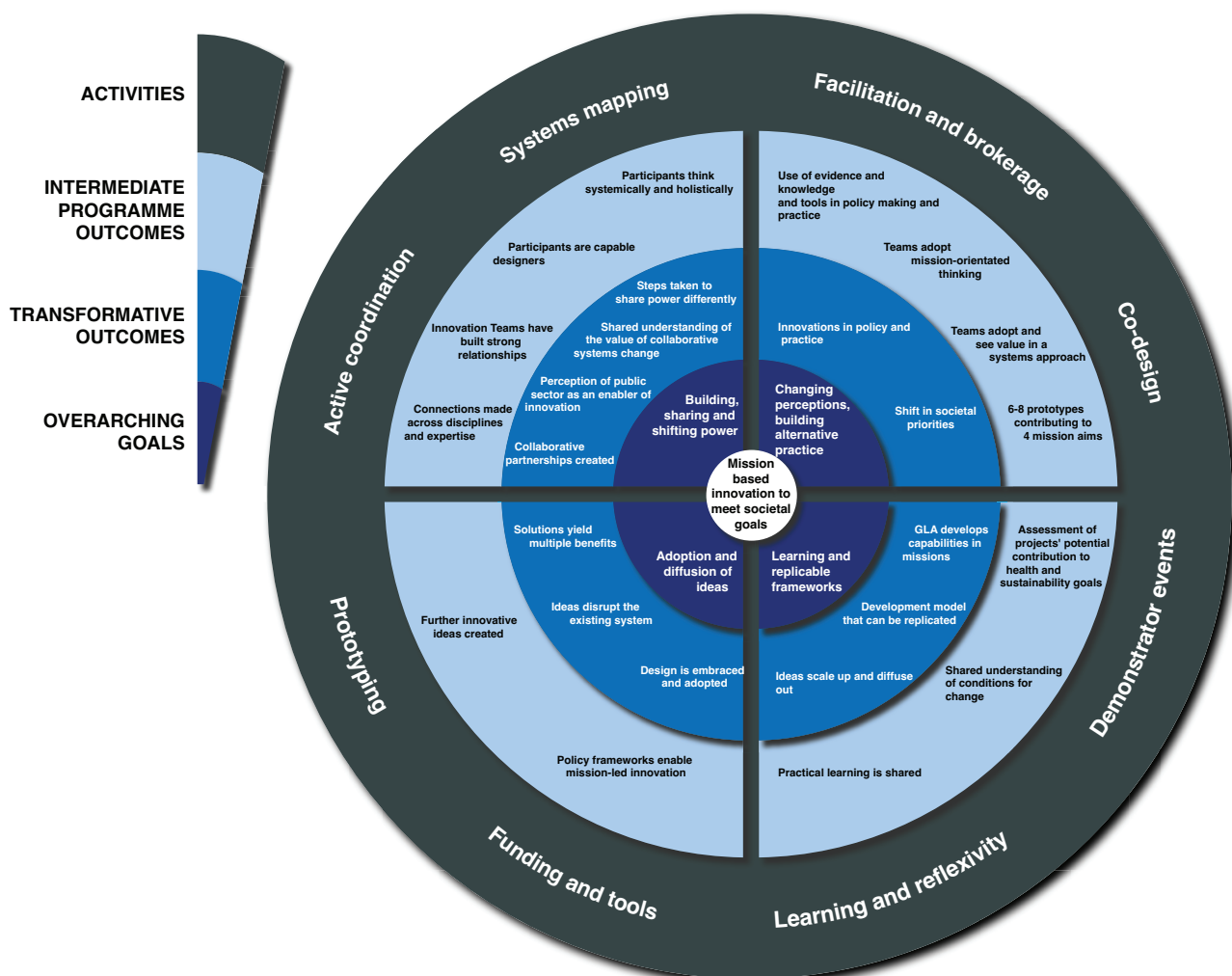
Table 2. (continued)

Goals	Building capacity, sharing and shifting power	Changing perceptions and building alternative practices	Adoption and adaptation	Learning
Intermediate outcomes	<p>Flexible processes enabled connections and coordination across different disciplines and expertise, to come up with effective and resilient ideas.</p> <p>Innovation teams have built strong relationships with one another and with wider stakeholders across each of the Mission areas; they have extensive networks.</p> <p>Programme participants have built their design skills and capabilities; they are capable designers.</p> <p>Participants feel confident navigating and creating around their mission: they are thinking systemically and holistically.</p>	<p>6–8 physical prototypes of what these new ideas are, which bring them to life for people to imagine what a new way of learning skills/ gaining jobs, living in communities and living regeneratively feels like.</p> <p>Innovation teams have pivoted their ideas so that they clearly lead to mission outcomes, and also have connections and dependencies to each other.</p> <p>Evaluation, behavioural and systems thinking approaches are included in the applications and projects. Innovation teams adopt, and see the value in using, a systems approach across their activities.</p> <p>The incorporation of scientific evidence, knowledge, and tools in decision and policy making, and innovation practice.</p>	<p>Further innovative ideas have been created from the initial 20 (i.e. further innovation).</p> <p>Policy frameworks enable decision making and investment to continue innovation- and mission-led programmes.</p>	<p>Projects have established how they will create change and will have shared their learnings so they can be scaled and diffused.</p> <p>The delivery team have a better understanding of what is needed to create fertile ground for these innovations to 'stick and grow' (i.e. there is a shared understanding of the conditions for change).</p> <p>Assessment of the degree to which projects have the potential to contribute meaningfully to health and sustainability goals.</p>

The aims of the sessions were to bring together programme partners at regular intervals to connect, capture and share learning. Alongside these aims, we agreed on some values to work towards within the sessions, which were: to create a space for sharing; to ensure that everyone could contribute equally; and that curiosity and criticality were welcomed. Sixteen participants were invited to sessions; the invited attendees included those involved in the delivery of the programme, but also wider connections within the GLA and Design Council. For instance, this included the programme sponsors (from the GLA and Design Council), GLA mission leads and Design Council appointed Design Experts.

Roles within these sessions were made explicit prior to the first session taking place, so it was clear why everyone was invited and what their expected contribution would be. What worked well about these sessions was that they provided a space for the different partners within the programme to share learning and to reflect on the overall programme delivery. The design experts provided emerging insights about the innovation teams, sharing connections across teams or emerging themes across the mission areas and raised any support needs identified. The mission leads provided updates on relevant mission-level activities from the GLA's wider work at a mission level. The evaluation and learning leads (including the two authors) captured learning, insights and any connections made, with the intention to integrate this learning into the ToC and programme evaluation.

Figure 3. Revised theory of change for DLRP (Source: designed by David Osrin)



Box 1. Researcher's notes from the second system session (Source: authors)

Making sense of the delivery and evaluation of Mission approaches – some points discussed helping us think through how this comes together:

- *The DLRP seeks to generate and incubate diverse projects addressing a mission from multiple angles in a specific context/place. The purpose being to build portfolios of distributed projects that collectively are addressing the mission, whether aimed at dismantling the current system and/or contributing to the building of new systems. Also, each of those individual projects involve multiple actors across different sectors working together on a specific challenge or task.*
- *Observation that the governance of these approaches happens at multiple levels – project, mission, overarching programme and wider system – and those levels interact.*
- *From an evaluation view – how do we know if they're successful? The approach we're using is one of needs to be flexible in response to the open-ended nature of possible transformative outcomes and also incorporating learning and reflection to enhance those outcomes. It does make sense to use this tool or another to indicate the 'full stack' approach to possible impact areas.*

What did not work as well was finding the space and sufficient time in these sessions to fully apply the ToC as a framework for reflection, problem solving and double-loop learning. Conducting these sessions online and intermittently proved to be a limitation, as discussion tended to drift towards status updates rather than structured reflection. In the first system session, the ToC was presented, receiving comments and clarification questions. Rather than reflecting on the programme, the ToC reminded participants that there was a need to support the innovation teams themselves to articulate their vision and impact for their projects. Following this, the authors ran a session at one of the workshops with the project teams on evaluating impacts from innovation.

In the second system session, the authors ran an exercise for the group to reflect on achieved outcomes against the ToC; this worked well to provoke a wider discussion of the layered, or nestled, nature of the programme and changes that may or may not result at project, programme and systems levels. The research notes from this session are set out in [Box 1](#). The exercise, applying the ToC, enabled the group to unpack the complexity of the programme, seeing the interconnected layers – between innovation projects, missions, the overarching programme and the wider system – more clearly. However, this articulation of the programme also raised a reflection point for the authors about whether the ToC tool (and connected learning and evaluation plan) was adequate to capture and represent the dynamic, multi-layered nature of the programme and its potential to drive systemic change.

Finally, within the third system session the ToC was used to present some key learning points from the programme. The aim was for the ToC to provoke a deeper reflection on the learning from Phase 1 of DLRP. Within the session the participants broke into four groups and each group were given a section of the ToC (e.g. a transformative outcome). Participants were asked to re-look at the outcomes and consider if they were still relevant, alongside if they felt the outcomes were being achieved and what evidence (if any) supported their claims. Some participants, who were not part of Stage 1 developing the ToC, commented on the language used, noting it 'felt technical and too academic', there was a request to 'tone down emphasis on "scientific" knowledge vis a vis other forms of knowledge'. The outcome 'contribution to health and sustainability goals' was noted to be 'too vague and also too specific for the programme', with some participants asking 'which goals?', 'why these and not others?'. The conception of the layered nature of the programme (which was cemented in the second system session) was returned to in the discussion: DLRP was a distributed set of projects catalysing change towards multiple missions, hence the programme-level ToC focused on means rather than the end. This approach for some was too vague. Following the third systems session the team revised the ToC drawing upon the reflections from sessions. For example, the language within the ToC was made more accessible (to address the feeling of it being too technical and academic). A tension lay within the ToC co-production process: although it was beneficial to co-produce the ToC to reach a shared agreement on the programme, co-production became a challenge

during the application stage of the ToC due to new ideas and people involved. As learning partners, the authors were to remind the team of past decisions, provide an open space for new perspectives to come in, and continuously translate the revisioning of the programme into an evolving ToC.

Within Phase 2 of DLRP the system sessions were not continued due to personnel changes within the Design Council, GLA and UCL. The lack of structured sessions reduced the opportunity to apply the ToC as a tool for collective reflection during Phase 2. Furthermore, the changes in personnel brought in 'new' expectations on innovation and how transformation happens, which were sometimes at odds with the collective aspiration built at the outset. It was at this point a challenge to maintain and invest in ongoing shared learning and understanding about the programme.

Stage 3: Reflect – reflection and the theory of change

Towards the end of the programme the ToC was used again to reflect on learning via interviews with policy-practice-research partnership. We conducted semi-structured interviews with 13 participants who were involved or linked to the programme, the majority of these participants were involved in Stage 1 (co-developed) and/or Stage 2 (application). These interviews were undertaken online, and we used purposive sampling to recruit participants. Fourteen potential participants were invited to take part, and 13 accepted, a response rate of 93 per cent. Within the sampling process we repeatedly assessed the balance of participants over the different organisations, that is CUSSH, GLA and Design Council.

The aim of the interviews during this stage were to:

- understand achievements, challenges and opportunities in delivering the programme from individuals involved;
- capture the outcomes and impacts resulting from the programme; and
- reflect on the theory of change based on learning from the programme.

These interviews were part of a wider evaluation and learning process for the programme. In the interviews we asked multiple questions to understand achievements, challenges and opportunities in delivering the programme. The visual of the ToC was used to encourage reflection on achievements of the programme, providing a systematic way to evaluate the perceived outcomes of the programme. Overall, participants used the opportunity to assess, from their perspectives and knowledge, the achievements of the intermediate programme and transformative outcomes. However, some participants (who had not been involved in Stage 1 of the ToC) questioned why the transformative outcomes and goals represented on ToC were about ways of doing (rather than the missions) and needed to be reminded of the development and focus of the ToC. Again, highlighting the challenge of maintaining a shared understanding about the programme and its goals.

When discussing the transformative outcomes some participants questioned the extent of their achievement but spoke of the value of having the transformative outcomes as a collective goal, and the role of the ToC in that process. As illustrated by one interviewee:

I'm going to murder the quote but there's an Abraham Lincoln quote which is about the abolition of slavery. He's trying to say that you can't take a direct line to the North Star, you'll end up in a swamp soon enough. You can't be absolutely dead set and follow it to the letter, you have to compromise on the way and adapt and adjust but it's your North Star.

For me, that's what mission-driven work is. It provides a North Star. There're lots of ways that you can get there that are really established, so developing a theory of change can help to find those North Stars and say, 'What's the future, where are we now, how might we get there?' Then you try to... well you start to action that plan and, sure enough, it meanders and it changes and it adapts and adjusts. Theories of change, you know, are things that you revisit, they're not just... a final outcome.

Furthermore, in discussing the transformative outcomes there was a focus on specific partners (e.g. the GLA) and whether the programme had influenced their ways of working and delivering policy, as noted by the quote below from an interviewee:

Respondent: I just don't think it... it didn't shift the ordinary way that things are adopted and the way the GLA adapts. I don't think it shifted, actually, this mega beast.

Interviewer: Is that the GLA, the mega beast?

Respondent: Yes, yes. Just the general way that it might impact the way policy is done, I don't think it did that.

Due to the iterative, formative nature of the evaluation, there were points for individual reflection and co-learning throughout the delivery of the programme. The learning stayed within the boundaries of the programme, rather than leading to specific actions or shifted practice in policies for organisation (e.g. the GLA). Some interviewees spoke about the active use of their learning for their own practice and their own projects, as outlined below:

I do have a sense that now that we have this as a pilot, I don't know what else will be built upon it but, it would be really useful... I refer to this project all the time and say, 'If we're designing this project, here's what worked in this other project. Here is what didn't work as well.' Actually that gets conveyed quite a lot.

In terms of the intermediate programme outcomes, all of the participants felt that they were relevant and had, mostly, been achieved.

Discussion

This paper has reported a case where a developmental, democratic evaluation approach, involving the development and application of a theory of change, was used to embed learning in a policy programme, being delivered by a policy-practice-research partnership.

There were strengths and limitations to our approach. Our 'learning partner' approach enabled the researchers to access and contribute to the delivery of a policy programme. We feel that, on balance, this was a strength of our approach; however, we acknowledge that it involved tensions around researcher roles and practical considerations. We were operating a responsive process which involved a range of engagement activities, including reviewing applicants to the programme, contributing to workshops with innovation teams, and supporting the innovation teams directly. In the programme we built up trusting relationships that afforded openness about what was working well and less well, but our connections with those who were interviewed could be seen as a limitation in terms of being 'objective'. The theory of change was developed through a participatory process to ensure the input of a broad range of perspectives and create a shared understanding of the goals of the programme among stakeholders from key organisations. Again, we consider this a strength, but there was staff turnover in the delivery team which resulted in those involved in the co-creation of the ToC were not present or active in the programme by the end of it. The changes in personnel resulted in the ambition of what we were doing to catalyse change – shifting or dwindling from something long-term and radical to more of a short-term and standard policy programme approach. Another limitation was that participation was only at a programme level, we could have gone further and included members from the innovation teams in the process.

TIPs and programmes, being delivered by a range of partners, require different approaches and ways of doing, but three points seem particularly pertinent from our experience. First of all, co-developing a ToC helped to unpack and explain how the TIP programme would 'work' to achieve its desired effects. The resulting ToC provided a *compass*, to help guide those delivering the programme. Our compass articulated what DLRP was contributing towards and how, for example a contributory approach to

systems change. The process of creating the ToC was valuable at the beginning of the programme as a way of bringing partners together, building consensus and co-producing shared goals. In this way it provided directionality and guided what we wanted to achieve. It was an incredibly helpful tool for setting ambitious goals beyond the scope of the immediate programme, setting the ambition to achieve systems change, as found in [Molas-Gallart et al. \(2021\)](#) and [Palavicino et al. \(2023\)](#). We were mindful that the DLRP programme was nestled within a larger programme – the London Recovery Programme – and the missions were being managed at that broader programme level and therefore DLRP was open to shift or pull. Hence, acting as a navigational tool, there needed to be opportunities for collective reflection to make sure the coordinates remained valid and true to those participating.

Second, we aimed for learning to move from a passive to active role in the evaluation of TIP based on collective learning in the governance of the programme. The presence of the ToC encouraged the sharing of knowledge and information, dialogue and reflection around goals and transformative outcomes. However, the distribution of power and expertise among the group meant that ‘ownership’ of the evaluation wasn’t fully distributed or democratised, instead it rested with one element of the partnership: the researchers. The ToC provided a reflective tool for capturing learning and informing policy implementation, but the process of learning rested with members of the delivery team. We would argue, as with [Ghosh et al. \(2021\)](#), that learning should be identified as a normative, shared goal within the TIP, governance of the programme and the wider context of system change. Capturing and sharing learning should be a key strategy of collaborative development, as is agreeing upon the aim of learning. We found that TIPs embrace different forms of knowledge and diverse routes of learning, and tools like a ToC can aid this, but not on their own. Within this case study there was not transformative learning – other things could have been done. For instance, others, including the innovation teams, could have been involved in the reflection (as in an ‘extended peer community’ or ‘co-operative inquiry’) and we could have pushed for more reflective practice from those involved in delivery. [Baldwin \(2001\)](#) notes the benefits of using co-operative inquiry to explore (and evaluate) issues of mutual interest: ‘the way in which co-operative inquiry ensures ownership of learning within the direct meaning and experience of participating individuals provides a very high likelihood of successful outcome’ (p. 235). We, like Baldwin, would recommend a distributed learning model to enable learning as well as organisational change. A distributed learning model would create structures and processes where learning is not confined to specific people (i.e. researchers) or teams but is shared across all participants, or partners. Formal activities (e.g. workshops, seminars and case studies) can help share learning, alongside more informal activities (e.g. mentoring, digital platforms).

Finally, there was a tension between standard programme delivery and learning-oriented approaches, echoing [Rohracher et al. \(2023\)](#). In reality, within DLRP, after an initial bold and ambitious start to the programme around innovation, systems change, and learning, priority reverted to ‘competition over collaboration’ and on allocative efficiency and the possibility of achieving (and evidencing) future economic value over creating the conditions and capacity for systemic change. As noted by one interviewee:

The way of thinking about which projects get awarded funding, they were being judged on an old GLA model rather than what we tried to promote through systems thinking and saying... Even if we have slightly less funding, but more abilities for the projects to collaborate, then that would be better. Instead, a lot of funding just went to the XX project which was much more, kind of, product innovation?

This drift from a learning-orientated programme to a more standard programme meant that stakeholders focused more on tangible outcomes delivered, than learning about creating conditions and capacities for systems change. Some interviewees reflected upon those conditions, but spoke of how that learning wasn’t shared: ‘Maybe, for me, it has probably shown real silos within [local] government where they’ve done a bit of building that trust but it’s not rippled through the GLA.’

The ToC wasn't directly used to inform decision making about the implementation of the programme; other factors were prioritised. It is therefore important to have routine ways to take a step back and return to the goals of the programme and what it's hoping to achieve. Scholars including [Molas-Gallart et al. \(2021\)](#) promote working principles for innovation programmes that are set at the start and revisited; reflexivity and sharing learning could be such a principle. However, we have found that in reality the principle alone isn't enough but needs to be accompanied by plans to fully integrate into programme decision making. For the ToC to 'work', it needs to be part of a wider learning culture within organisations (e.g. GLA) with the freedom to experiment and innovate. Learning can be embedded as a core value of an organisation and actively supported by leadership and structures, which make it clear and visible how lessons from one programme influence changes in wider practice or policy. In our role of learning partner, this was outside our scope; instead we focused on building relationships, creating trust and the delivery of one programme.

Our participatory action research has demonstrated that ToCs have the potential to be a useful compass in transformative change programmes. However, being effective in that role requires shared investment and ownership of the learning processes, alongside dedicated and consistent spaces for collective reflection.

Conclusions

The paradigm shift towards transformative innovation delivered by complex partnerships cannot unfold unless it is followed by a shift in the practices of policy formation, programme implementation and evaluation. Our work, focused on embedding learning within DLRP, raises questions on how knowledges come together and how learning happens in practice. Within our case study, the ToC provided a tool for convening and building consensus around the direction of the operational delivery of the programme – but less so on the overall ambition of the systemic change, which sat within the wider recovery mission programme. Our study has found that ToCs are useful tools, but not enough on their own. There needs to be a building of capabilities and strengthening of capacity for more transformative ways of working, across diverse partnerships, and this is an ongoing process.

Acknowledgements

Funding was provided by the Wellcome Trust under the CUSSH programme (grant number 209387/Z/17/Z). We would like to thank those involved in Designing London's Recovery Programme, including colleagues from CUSSH (notably Dr Guiseppe Salvia, who led on the delivery of Phase 1) and the participating innovation teams. We are extremely grateful to Professor David Osrin, UCL, who designed the theory of change.

Declarations and conflicts of interest

Research ethics statement

The CUSSH project was approved through the UCL low-risk ethics process, as part of the collaboration a memorandum of understanding was drafted between the GLA and CUSSH for their work on CUSSH. Verbal consent was given during workshops, and for the interviews participants received an information sheet and consent form and gave signed consent prior to participating.

Consent for publication statement

The author declares that research participants' informed consent to publication of findings – including photos, videos and any personal or identifiable information – was secured prior to publication.

Conflicts of interest statement

The authors declare that there is no conflict of interest. All authors were learning partners embedded in the DLRP programme. GM designed the participatory workshops to create the ToC; JM led the interviews; all authors analysed the data and prepared the manuscript. All authors read and approved the final manuscript.

References

- Baldwin, M. (2001). Co-operative inquiry as a tool for professional development. *Systemic Practice and Action Research*, 15(3), 223–235. <https://doi.org/10.1023/A:1016392325258>
- BEIS. (2020). *Alternative policy evaluation frameworks* (BEIS Research Paper Number 2020/044). Department for Business, Energy & Industrial Strategy. https://www.ucl.ac.uk/bartlett/public-purpose/sites/public-purpose/files/iipp-beis-alternative_policy_evaluation_frameworks_and_tools_oct_2020_final.pdf
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Breuer, E., Lee, L., De Silva, M., & Lund, C. (2015). Using theory of change to design and evaluate public health interventions: A systematic review. *Implementation Science*, 11, Article 63. <https://doi.org/10.1186/s13012-016-0422-6>
- Coffman, J., Beer, T., Stid, D., & Armstrong, K. (2023). Using developmental evaluation to support adaptive strategies: An application from a social change initiative. In A. Rangarajan (Ed.), *The Oxford handbook of program design and implementation evaluation* (pp. 151–171). Oxford Academic. <https://doi.org/10.1093/oxfordhb/9780190059668.013.10>
- Cook, S. N., & Wagenaar, H. (2012). Navigating the eternally unfolding present: Toward an epistemology of practice. *The American Review of Public Administration*, 42(1), 3–38. <https://doi.org/10.1177/0275074011407404>
- Estrella, M., & Gaventa, J. (1998). Who counts reality? Participatory monitoring and evaluation: A literature review (Working Paper No. 70) Institute for Development Studies, University of Sussex.
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5(1), 80–92. <https://doi.org/10.1177/160940690600500107>
- Floc'hlay, B., & Plottu, E. (1998). Democratic evaluation: From empowerment evaluation to public decision-making. *Evaluation*, 4(3), 261–277.
- Ghosh, B., Kivimaa, P., Ramirez, M., Schot, J., & Torrens, J. (2021). Transformative outcomes: Assessing and reorienting experimentation with transformative innovation policy. *Science and Public Policy*, 48(5), 739–756. <https://doi.org/10.1093/scipol/scab045>
- Hanberger, A. (2006). Evaluation of and for democracy. *Evaluation*, 12(1), 17–37. <https://doi.org/10.1177/1356389006064194>
- Hargreaves, M. B. (2010). *Evaluating system change: A planning guide*. Mathematica Policy Research, Inc. <https://www.tamarackcommunity.ca/hubfs/Resources/Evaluating%20Systems%20Change%20a%20Planning%20Guide.pdf>
- Hildén, M., Jordan, A., & Rayner, T. (2014). Climate policy innovation: Developing an evaluation perspective. *Environmental Politics*, 23(5), 884–905. <https://doi.org/10.1080/09644016.2014.924205>
- Hölscher, K., Wittmayer, J. M., & Loorbach, D. (2018). Transition versus transformation: What's the difference? *Environmental Innovation and Societal Transitions*, 27, 1–3. <https://doi.org/10.1016/j.eist.2017.10.007>
- Köhler, J., Geels, F. W., Kern, F., Markard, J., Onsongo, E., Wiecek, A., Alkemade, F., Avelino, F., Bergek, A., Boons, F., Fünfschilling, L., Hess, D., Holtz, G., Hyysalo, S., Jenkins, K., Kivimaa, P., Martiskainen, M., McMeekin, A., Mühlemeier, M. S., Nykvist, B., Pel, B., Raven, R., Rohrer, H., Sandén, B., Schot, J., Sovacool, B., Turnheim, B., Welch, D., & Wells, P. (2019). An agenda for sustainability transitions research: State of the art and future directions. *Environmental Innovation and Societal Transitions*, 31, 1–32. <https://doi.org/10.1016/j.eist.2019.01.004>
- Linnér, B., & Wibeck, V. (2020). Conceptualising variations in societal transformations towards sustainability. *Environmental Science and Policy*, 106, 221–227. <https://doi.org/10.1016/j.envsci.2020.01.007>
- Molas-Gallart, J. (2015). Research evaluation and the assessment of public value. *Arts and Humanities in Higher Education*, 14(1), 111–126. <https://doi.org/10.1177/1474022214534381>
- Molas-Gallart, J., Boni, A., Giachi, S., & Schot, J. (2021). A formative approach to the evaluation of transformative innovation policies. *Research Evaluation*, 30(4), 431–442. <https://doi.org/10.1093/reseval/rvab016>
- Moore, G., Michie, S., Anderson, J., Belesova, K., Crane, M., Deloly, C., Dimitroulopoulou, S., Gitau, H., Hale, J., Lloyd, S. J., Mberu, B., Muindi, K., Niu, Y., Pineo, H., Pluchinotta, I., Prasad, A., Roue-Le Gall, A., Shrubsole, C., Turcu, C., Tsoulou, I., Wilkinson, P., Zhou, K., Zimmermann, N., Davies, M., & Osrin, D. (2021). Developing a programme theory for a transdisciplinary research collaboration: Complex urban systems for sustainability and health. *Wellcome Open Research*, 6, 35. <https://doi.org/10.12688/wellcomeopenres.16542.1>
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1), 1–13. <https://doi.org/10.1177/1609406917733847>

- Palavicino, C. A., Matti, C., & Brodnik, C. (2023). Co-creation for transformative innovation policy: An implementation case for projects structured as portfolio of knowledge services. *Evidence and Policy*, 19(2), 323–339. <https://doi.org/10.1332/174426421X16711051078462>
- Pineo, H., Turnbull, E. R., Davies, M., Rowson, M., Hayward, A. C., Hart, G., Johnson, A. M., & Aldridge, R. W. (2021). A new transdisciplinary research model to investigate and improve the health of the public. *Health Promotion International*, 36(2), 481–492. <https://doi.org/10.1093/heapro/daaa125>
- Plottu, B., & Plottu, E. (2011). Participatory evaluation: The virtues for public governance, the constraints on implementation. *Group Decision and Negotiation*, 20, 805–824. <https://doi.org/10.1007/s10726-010-9212-8>
- Pontikakis, D., González Vázquez, I., Bianchi, G., Ranga, M., Marques Santos, A., Reimeris, R., Mifsud, S., Morgan, K., Madrid, C., & Stierna, J. (2022). *Partnerships for regional innovation – Playbook* (EUR 31064 EN), Publications Office of the European Union. <https://doi.org/10.2760/775610>
- Reason, P., & Bradbury, H. (Eds). (2008). *Handbook of action research: Concise paperback edition*. Sage.
- Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14, 29–48. <https://doi.org/10.1177/1356389007084674>
- Rohracher, H., Coenen, L., & Kordas, O. (2023). Mission incomplete: Layered practices of monitoring and evaluation in Swedish transformative innovation policy. *Science and Public Policy*, 50(2), 336–349. <https://doi.org/10.1093/scipol/scac071>
- Sabel, C. F., & Zeitlin, J. (2012). *Experimentalist governance*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199560530.013.0012>
- Schot, J., & Steinmueller, W. E. (2018). Three frames for innovation policy: RandD, systems of innovation and transformative change. *Research Policy*, 47(9), 1554–1567. <https://doi.org/10.1016/j.respol.2018.08.011>
- Völker, T., Kovacic, Z., & Strand, R. (2020). Indicator development as a site of collective imagination? The case of European Commission policies on the circular economy. *Culture and Organization*, 26(2), 103–120. <https://doi.org/10.1080/14759551.2019.1699092>
- Wanzenböck, I., & Frenken, K. (2020). The subsidiarity principle in innovation policy for societal challenges. *Global Transitions*, 2, 51–59. <https://doi.org/10.1016/j.glt.2020.02.002>
- Weber, K. M., & Rohracher, H. (2012). Legitimizing research, technology and innovation policies for transformative change. *Research Policy*, 41(6), 1037–1047. <https://doi.org/10.1016/j.respol.2011.10.015>