Evaluating Field-Building Intermediaries

Challenges and emerging approaches for early-stage field catalysts

Early Years Catalyst

Clear Horizon

Acknowledgements

BHP | Foundation

Suggested Citation:

Barkley, E., 2023. Evaluating Field-Building Intermediaries: Challenges and emerging approaches for early-stage field catalysts prepared by Clear Horizon for the Early Years Catalyst.

Contents

Contents

Executive Summary

This concept paper discusses the challenges of and emerging approaches for evaluating field-building intermediaries, with a focus on field catalysts. It unpacks the need to go beyond conventional programmatic approaches to understand and demonstrate their influence on systems and longer-term social transformation, presenting leading approaches from international and Australian practice.

A field catalyst is a type of field-building intermediary that aims to unite and leverage the work of individual initiatives whilst linking these to the efforts of influential systems players to influence systems transformation. Field catalysts are a distinct type of intermediary; however, they often include elements from other intermediary models such as capability specialists, place-based backbones and evidence-action labs. One initial issue in evaluating field catalysts is the limited understanding and lack of existing literature about the model itself, as the identification of this type of intermediary is relatively recent. We need to better understand how field catalysts work to effectively wrap evaluation around them.

Based on what we are learning so far, our hunch is that the nature and form of the field catalyst model itself have the potential to turbocharge the challenges of systems change and evaluation, even compared to other types of intermediaries. Compared to other field-building intermediaries, field catalysts often have looser boundaries and are less likely to implement tangible activities. As such, they are more opportunistic and emergent, with longer 'on-ramps' to achieving tangible results. They can also have wider geographic boundaries than place-based models, a wider set of roles than capacity-building intermediaries, and their work is usually more indirect than that of action labs.

This paper is structured around three key evaluation challenges and suggested evaluation responses:

CHALLENGE 1

Measuring progress and managing expectations when there are long timeframes, non-linearity, and emergence.

In response, we discuss evaluation approaches that:

a) Rethink how 'results' are framed and evaluated to include process, enabling conditions, systemic results, and the mission and/or population level.

b) Broaden notions of accountability to include being accountable to learning and to wider cohorts of stakeholders.

c) Use a mixed method, systems-aware toolkit.

CHALLENGE 2

Conventional evaluation planning doesn't hold up when the "thing" being evaluated is dynamic.

To address this challenge, we recommended:

a) A phased evaluation strategy that balances planning and rigour with flexibility.

b) Developing agreed principles to guide evaluation over time.

c) Learning-focused evaluation that fosters sharing and learning at different scales.

CHALLENGE 3

Determining contribution and managing the politics of impact claims.

To tackle this, we focus on:

a) The importance of framing and assessing contribution, not attribution.

b) Getting clear on the contribution of the 'whole' as well as unpacking contribution from different parts of intermediary models.

The paper is intended as a snapshot of current practice and is not a user guide or extensive literature review. It is an output of discovery work commissioned by the Early Years Catalyst and conducted by Clear Horizon. The findings are informed by key informant interviews and a light literature and practice scan of evaluation of field-building intermediaries, field catalysts, and systems change initiatives more broadly.

About this document

This concept paper shares findings from discovery work undertaken by Clear Horizon for the Early Years Catalyst (EYC). The EYC is a national collaboration working to improve early childhood development outcomes for children experiencing disadvantage and vulnerability in Australia. The EYC was established in 2021 with the ultimate goal of becoming a field catalyst for early childhood development in Australia.

This discovery work aimed to investigate approaches for **evaluating field catalysts**, particularly in the **early stage** of set-up. This paper brings together the findings to inform the evaluation design for the EYC. It is being shared more broadly to support other field catalysts to better understand and demonstrate their influence on systems and social transformation, given there is so little in the public domain on this topic.

The findings are based on a light literature scan and interviews with thirteen key informants with expertise in evaluating systems change initiatives and early childhood development. The paper represents **a snapshot of selected perspectives and practice** and is not intended as an extensive literature review or a user guide. The lessons and themes from the key informants are woven throughout and were important in shaping and nuancing the findings. We reference key literature sources that may be of interest for further reading. We also draw on Clear Horizon's substantial experience as an evaluation partner of over 15 field-building intermediaries and the many systems transformation initiatives we work with, including Our Town, The Front Project, Indigenous Eye Health, and numerous place-based backbone organisations across Australia.

How the paper is structured

From our understanding of evaluating intermediaries and systems-change initiatives, **we focus on three key challenges** and offer evaluation approaches that address these.
The paper is set out as follows:

Introduction

Provides background on field catalysts and the challenges for evaluation. It includes definitions for the concepts of evaluation, systems change, and complexity.

SECTION 1
Challenge 1

Measuring progress and managing expectations when there are long timeframes, non-linearity, and emergence.

The approaches are:

1a) re-thinking how 'results' are framed and evaluated

1b) broadening notions of accountability

1c) using a systems-aware toolkit

SECTION 2
Challenge 2

Conventional evaluation planning doesn't hold up when the "thing" being evaluated is dynamic.

Evaluation approaches to respond to this challenge include:

2a) creating a phased evaluation strategy that balances planning and rigour with flexibility

2b) developing principles to guide evaluation over time

2c) implementing learning-focused evaluation

SECTION 3
Challenge 3

Determining contribution and managing the politics of impact claims.

For this challenge we explore:

3a) focusing on contribution not attributiony

3b) getting clear on the contribution of the 'whole' and unpacking contribution from different parts of the field catalyst

Final Thoughts

A few concluding remarks and looking forward.

The **Annex** provides further detail about complexity-aware toolkits for evaluators.

Field-building intermediaries and field catalysts

Intermediary:

An organisation that fills capability gaps that individual, community-based change efforts cannot fill on their own

Field catalyst:

A **type of field-building intermediary that deploys different capabilities to amplify and orchestrate the efforts of multiple actors** who are focused on a **common goal**, quietly influencing and augmenting the field's efforts to **achieve population-level change**

'Field':

A set of individuals and organisations (actors) working to address a common social issue or problem, often developing and using a common knowledge base

'Field-building':

The activities or investments that unlock a field's progress toward greater impact at scale

The number of **field-building intermediaries** is growing in Australia and they play an important role in systems transformation. Field-building intermediaries aim to support and develop the capacity of a given field to better achieve impact at scale around an identified social issue or problem. They do this by engaging and coordinating stakeholders across a field or system (e.g. through partnerships or sharing leading practices)[2]. It often takes the form of a dedicated catalysing organisation (also known as a facilitating partner or backbone) that convenes stakeholders addressing a social issue from different angles and parts of the system, creating an 'intentional ecosystem' for social ventures to do their work[3].

There are **four main types of field-building intermediaries** covered in the literature: capability specialists, place-based backbones, evidence-action labs, and field catalysts[4]. The identification and codification of the different types of field-building intermediaries is a relatively recent phenomenon. Papers contributing to our understanding of field-building intermediaries include: Cheuy, Cabaj, and Weaver's 'How Field Catalysts Accelerate Collective Impact' (2022); Social Ventures Australia's report, *Insights on Australian field-building intermediaries and their funding journeys towards sustainable impact* (2022) and case studies, Cabaj's paper *Evaluating the Results of Intermediary Organisations: A Paper for Intermediaries in Australia* (2021); The Bridgespan Group's article *How Philanthropy Can Support System-Change Leaders* (2021) and paper *Field Building for Population-Level Change: How Funders and Practitioners Can Increase the Odds of Success* (2020); and Hussein, Plummer and Breen's article 'How Field Catalysts Galvanise Social Change' (2018).

A **field catalyst** is a type of intermediary that aims to unite and leverage the work of individual initiatives whilst linking these to the efforts of influential systems players (actors)[5]. Field catalysts deploy different capabilities to nudge powerful systems players and augment efforts across a field to achieve population-level change, as no one organisation is in a position to drive and organise such a distributed system. Field catalysts influence the actions of others rather than acting directly. They need to be adaptive, recognising and seizing opportunities to advance systems change – identifying promising strategies as they emerge and translating them into practical actions that can be disseminated[6]. They are distinctive from other kinds of intermediaries because they weave together a diversity of skillsets in order to facilitate integrated bottom-up and top-down approaches to change. While they can include elements of capability building, regional and place-focus, and evidence-action learning, field catalysts are not confined or defined by specific niche strategies or scope. (See *Catalysing Change at Scale: Features and enablers of effective field catalysts and field-building intermediaries*, 2023 for further description.)

The **models and structures for field catalyst intermediaries** vary[7]. They can include (but are not limited to) a dedicated facilitating partner or backbone team, a leadership table made up of strategic system actors from different organisations and communities who play leadership and governance roles; the broader collaboration, which includes all the organisations and partners involved in implementation; and the funders sponsoring parts of the effort and/or the facilitating partner (usually government and philanthropic organisations). In some models, the facilitating partner is deliberately neutral and does not set the advocacy or strategic agenda. As with field-building intermediaries more broadly, they will go through distinctive phases such as catalysing, growing, sustaining, and renewal or wind-down phases[8].

While we are still very much learning about field catalysts, our hunch is that the nature and form of the field catalyst model itself have the potential to turbocharge the challenges of systems change, even compared to other types of intermediaries. Based on what we know so far, compared to other field-building intermediaries, field catalysts tend to have looser boundaries, be less likely to implement tangible activities, and usually play in highly political spaces. As such, they are more opportunistic and emergent, with longer 'on-ramps' to achieving tangible results. Field catalysts have wider geographic boundaries than place-based intermediaries, which are intended to centre the needs of a defined local community. They have a wider set of roles than capacity-building intermediaries, which have a more clearly defined focus on skill-building. Finally, their work is more facilitatory and indirect than that of action labs, which often prototype solutions with key cohorts or micro-places.

Challenges for evaluation

As Michael Quinn Patton notes, evaluation is parasitic to the thing being evaluated. Our first big challenge in evaluating field catalysts is that the lack of existing literature about the model already places us somewhat in the dark. We need to understand more about how field catalysts work in order to effectively wrap evaluation around them, and understand their results and whether they are on track.

Nonetheless, our starting assumption is that **field catalysts experience similar challenges as other types of intermediaries,** but with greater intensity. In particular, a field catalyst's long on-ramps, emergent nature of results, distributed roles and indirect action mean that we need to take a radically different approach to measuring progress compared to conventional evaluation. Well-worn approaches that have their foundations in program evaluation – including conventional understandings of evaluation planning, measurement, learning and causality – are not suited to this context. Furthermore, a field catalyst's challenges are felt even more keenly in the early design and start-up phases of systems change initiatives.

In this paper, we consider the three key challenges identified in the discovery research on evaluating intermediaries and reflect on how this plays out for field catalysts.

- **Challenge 1:**
  Measuring progress and managing expectations when there are long timeframes, non-linearity, and emergence.

- **Challenge 2:**
  Conventional evaluation planning doesn't hold up when the "thing" being evaluated is dynamic.

- **Challenge 3**
  Determining contribution and managing the politics of impact claims.

Evaluation practice that addresses these challenges is still developing and we hope that this paper contributes to this evolving thinking and adds to the field of evaluating field catalysts, which is still nascent.

Key concepts

EVALUATION

Before we explore the findings, there are a few key concepts to explain. The first is what we mean by **evaluation**. In this paper, we are taking a very broad view of evaluation to encompass measurement, evaluation, learning, developmental evaluation, formative evaluation and summative evaluation.

- **Measurement** is the ongoing collection of data (numbers and stories) to understand what is changing as a result of our work.

- **Evaluation** involves posing and answering key question important to us in the short and longer term. This includes having evaluative thinking as an essential part of how we work as we try, test, learn, and adapt as we go.

- **Learning** refers to using both measurement and evaluation data to answer key evaluation and learning questions that inform strategy, practice, delivery, and adaptation. It includes both formal and informal learning across different scales.

- **Developmental evaluation** uses evaluation and learning practices to inform the iterative development of an initiative. It involves collecting data, analysing it and feeding it back to social innovators so they can make evidence-based decisions for the design, development, and implementation of an initiative[9].

- **Formative evaluation** is a type of evaluation that is intended to help an initiative make improvements[10].

- **Summative evaluation** is a type of evaluation that is intended to inform decisions about whether an initiative should start, continue, expand, or stop[11].

In this paper, we use the term 'evaluation' as a simple way to refer to all of the above components.

SYSTEMS CHANGE

Another key concept is **systems change and transformation interventions**, in which field-building intermediaries play a part. Systems change and transformation interventions more generally can be defined as the intent to address the causes of a societal issue, rather than the symptoms, by taking a holistic view of the many policies, practices, power dynamics, social norms, and mindsets that contribute to this issue[12]. The systems thinking that underpins these approaches acknowledges that the context of systems change is dynamic and complex. As opposed to programmatic approaches – in which evaluation 'grew up' – systems change initiatives often involve multiple programs, are multi-disciplinary and have much wider boundaries than a singular program[13].

COMPLEXITY

The last term is **complexity**, used in the context of complex adaptive social systems. In such systems, multiple semi-independent agents continually interact with each other, adapting to each other and the environment as a whole. Complexity in this context refers to the fact such systems are characterised by multiple points of influence, dynamic patterns, and unresolved tensions[14]. Here, systems change conditions, patterns and results can be unpredictable and emergent. Usually, it is a series of multiple actions and agents that contribute to change rather than one strategy or single entity[15].

SECTION 1

Challenge 1:

Measuring progress and managing expectations when there are long timeframes, non-linearity, and emergence

1.1 ABOUT THE CHALLENGES

A key challenge for evaluators is how to **measure progress and manage stakeholder expectations** given the complex nature of systems transformation. Transformative change can involve many phases and is long-term (often intergenerational) and initiatives can work across multiple systems and locations, involving multiple interventions and moving parts. Activities do not always lead directly to intended results, and sometimes unintended results occur instead. Systems change is also characteristically messy and has long 'on-ramps' before tangible results emerge. Defining progress goals and what success looks like for 'nudging the system' can be problematic given the journey is unpredictable. Moreover, conventional notions of causality and attribution are unsuited to systems change initiatives (we tackle this separately in Challenge 3). In this context, evaluation is crucial for informing implementation, demonstrating progress and impact, and securing funding. However, it is challenging for numerous reasons.

For one, **data may be limited or not available** at the systems level or for tracking ripple impacts. A challenge is to resist focusing on easy-to-count metrics and data over what is needed for the work.

"We keep measuring the stuff that's easy to measure, but it's actually not the stuff that makes the difference. And the stuff that makes the difference therefore is not visible and doesn't get funded."

– Key Informant

In addition, the complexity of systems change can lead to **expectations failure** for stakeholders and sponsors. Funders, especially government agencies, can have impetus to 'demonstrate results' through output targets or a heavy focus on singular quantitative measures that are insufficient for measuring systemic changes in relationships, power, resources flows and purpose. This can be problematic, especially in the early phases of designing the model and governance structures. Managing expectations about realistic phase-relevant goals and the pace of results is particularly challenging where initiatives are sponsored over shorter-term funding periods. Expectations failure can short-circuit the work and soak up time and resources that provide little value[16].

It can also be **difficult to co-define shared measures** for a data-driven approach across stakeholders. This is compounded for national initiatives, which are trying to measure impact at scale in a way that is meaningful in place and locally. A key challenge is to design and implement measurement, evaluation and learning that holds true across diverse contexts and translates voices of lived experience into national results.

Of all the field-building intermediary models, it would seem that the **field catalyst is likely to experience these challenges most acutely**. They have longer on-ramps and perhaps the least control over the direction of results compared to other field-building intermediaries. By definition, field catalysts are not directing advocacy or running activities, but rather are mobilising a field. There are likely to be more actors to get on board, with longer timeframes to broker relationships and set up governance and leadership structures (compared to some place-based or more direct advocacy approaches). Furthermore, field catalysts often work across the state, national or even international scales.

1.2 EVALUATION APPROACHES FOR CHALLENGE 1

To address the challenge of measuring progress and managing expectations, we suggest three evaluation approaches:

    a) Rethink how 'results' are framed and evaluated

    b) Broaden notions of accountability

    c) Use a systems-aware toolkit

1A.

Rethink how 'results' are framed and evaluated

When framing and measuring progress for systems transformation, we need to think about results differently to programmatic approaches. Results for field catalysts are likely to be dispersed and are more about catalysing, nudging, or boosting a change effort than directly driving outcomes. Below, we frame the results relevant to the field catalyst across three levels: process and enablers, systemic changes, and mission-level outcomes that reflect transformative change (Figure 1). Based on what we heard from key informants and in the literature, we offer tips for evaluating results for each level.
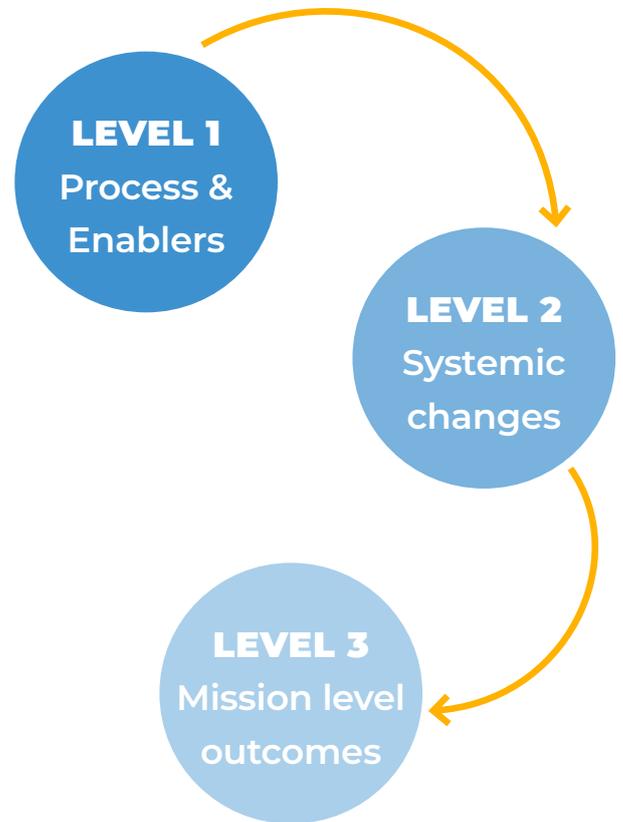
LEVEL 1
Process & Enablers

LEVEL 2
Systemic changes

LEVEL 3
Mission level outcomes

Figure 1: Three levels of results on the evaluation horizon

Level 1: Process and enablers

When initiatives are in their discovery and design phase, it is most appropriate to focus on process and enablers as results. **Developmental evaluation** is a particularly suitable evaluation methodology for this. Developmental evaluation can assist social innovators and change makers to develop initiatives by generating insights that help the design of the model itself[17]. It is well suited to uncertain environments and can help by framing concepts, testing quick iterations, and surfacing assumptions.

During this early phase, **learning can be considered a result** in and of itself[18]. At a time when few other tangible results will be evident, demonstrating learning and adaptation can be an important way to hold the intermediary accountable to its aims, as well as being essential to support the developmental work. Understanding the performance of the intermediary and the extent to which it learned from and adapted their work will be critical to ensuring that the intermediary can contribute to strategic results and impact[19].

It can be useful to focus on evaluating how well initiatives are upholding their **principles** – which again can be considered a result in and of itself. One key informant said there is a need for more resources that support system stakeholders to enact principles. They said that there can be a tendency for evaluation to focus on the technical side of outcomes measurement as opposed to "dig[ging] down to what [ways of working] are enabling an organisation to do really good work." Principles-focused evaluation can also facilitate accountability as the initiative can stop to reflect if principles of social justice and equity are being followed.

As the field catalyst model solidifies with time, evaluation can support an understanding of whether the right **conditions** are in place for the intermediary (such as trust, relational practice, governance, having the right people at the table, etc.), as well as the intermediary's early progress

in strengthening or disrupting the conditions for change. One key informant talked about the importance of having explicit conversations about values, power and institutional racism, highlighting that creating the conditions for change requires acknowledging the relationships and context in which that happens. Milligan, Zerda and Kania similarly emphasise the importance of relational practice:

"Making meaningful progress on the complex challenges of our time requires totally different ways of working together that prioritise relational practices."[20]

Given the intent, role, and context of field catalysts, tracking early results will need metrics about process (including principles and learning), relational elements, and enabling conditions. One key informant emphasised that given the need to learn and pivot as findings emerge, setting too many or unsuitable key performance indicators (KPIs) early on can be counterproductive. It is often necessary to work with boards and governance bodies to make sure they are aware of the appropriate type of milestones and progress measures to set to monitor progress. Evaluating progress may be more about understanding and mapping where progress has or hasn't been made, rather than using simplistic measures.
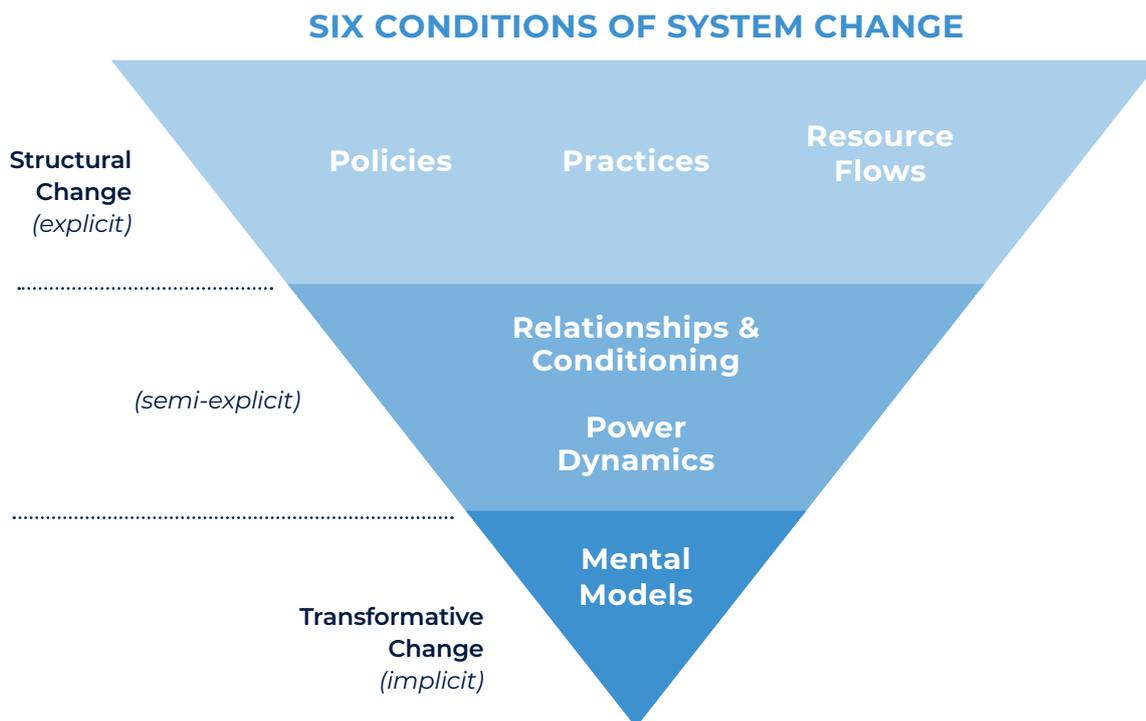
Level 2: Systemic changes

A key part of evaluating intermediaries will be **tracking their influence on systemic change**. The results most causally proximate to an intermediary's activities are more likely to be related to progress in nudging the conditions or the direction of systems transformation, rather than outcomes for populations. How and where we track results will depend on the field catalyst's agreed role in systems transformation, which may include diagnosing and assessing systems contexts and ecosystems, convening and activating system actors, and galvanising and amplifying efforts. For field catalysts, results might include research and evaluation, building public awareness, systems evaluation, influencing shifts in capability, mindsets, practice, policy, and sharing data and learnings. We need to look for intended results and unintended results to understand what is happening as a consequence of the intermediary's efforts.

There are various schemas that help identify the conditions of social systems that interventions might set out to shift to support sustainable transformation. These can be useful for strategy design as well as for evaluators. A commonly used framework is *The Water of Systems Change* framework created by FSG[21], which identifies six conditions for systems change (Figure 2): policies, practices, and resource flows (explicit); relationships/connections and power dynamics (semi-explicit); mental models (implicit). Mark Cabaj places disruption to systems components within 'strategic outcomes' in his typology of results in *Evaluating the Results of Intermediary Organisations*[22]. Another framework is the 'four keys' for unlocking systems innovation, identified as changing the purpose, power, relationships, and resource flows of a system[23].

Figure 2:

SIX CONDITIONS OF SYSTEM CHANGE

Structural Change *(explicit)*

Policies   Practices   Resource Flows

(semi-explicit)

Relationships & Conditioning

Power Dynamics

Transformative Change *(implicit)*

Mental Models

Figure 2:
The Water of Systems Change's six conditions of systems change

Of the different systems components, **power and relationships** were most frequently cited as important to keep in focus. This requires a high degree of self-awareness for systems change practitioners, and an ongoing process of examining biases, assumptions, and privilege. An interplay between this individual learning and broader process of collective change creates the potential for transformation of relationships and power dynamics[24]. Several key informants said that the **strengthening of relationships and collaboration** is both an enabler as well as an interim result for field catalysts (given their role to convene stakeholders and learning). Tracking relationships and collaborative innovation, for example, can be important as systems change initiatives usually require new connections to be made.

Evaluation of changes to the systems, therefore, requires nuance and close attention to **context, power, and positionality**. It will involve tracking, reflecting on, and reporting on what some practitioners describe as 'a mix of tangible and intangible outcomes' (e.g. shifts in language, narratives, relations, power sharing, and agency).

Level 3: Mission-level impact and results at scale
(including population-level outcomes)

Systems change interventions often look to influence impact at scale – either at the population level aligning with the intermediary's mission to break the cycle of disadvantage, or around the reimagined purpose of the field or system. Intermediaries and partners need to be forward-looking and define their longer-term results. Measuring mission-level impact and results at scale typically has an emphasis on **outcomes and impact evaluation, including contribution**. Evaluation also informs decisions about scalability of interventions and innovations.

We have learned that population results need to be considered at **different scales**. Mark Cabaj splits what he calls 'mission impact' into three categories: intervention level, or the direct effects of a key intervention on a group; targeted level, or the cumulative effects of multiple key initiatives on a subset of a target population; and population level, or the cumulative effect of multiple key initiatives on a population of people[25]. *The Place-based Evaluation Framework* takes a similar angle to this consideration, separating instances of impact (for family, cohort or micro-communities), and sustainable population-level impacts[26].

While results at scale take a long time to achieve, one key informant emphasised the need to **keep a sense of urgency regarding achieving the mission-level results** given what is at stake and to not be 'forgiving' about slow progress and results because it is systems change.

1B.

Broaden notions of accountability

A way to address the challenge of managing expectations is **to broaden the definition of accountability** (from the start). A strong theme from the interviews was that conventional definitions of accountability and inclusion are often too narrow. Accountability is commonly viewed as the evaluation reporting back the activities and outcomes of the initiative to a funder who wants to know if their funds were spent effectively so they can make decisions about whether funding is renewed or not.

For field catalyst efforts driven by diverse and cross-sector cohorts (including community members and people with lived experience), accountability needs to be much broader. Rather than only being accountable to funders, intermediaries should be accountable to the cohorts most affected by the initiative's work – whether that be organisations, communities, or both. For this reason, it is critical to involve multiple perspectives (especially those with lived experience) early on, for the work and for evaluation.

Broadening accountability can also refer to broadening the scope of evaluation to include evaluating the enabling role of sponsors and partners, so they are accountable within the ecosystem as well. Field catalysts also need to be held accountable for learning and adaption.

1C.

Use a systems-aware toolkit

When measuring progress for intermediaries, we will need to **look beyond conventional program evaluation toolkits**, as some standard methods are either not possible (e.g. experimental research design and trials with control groups) or not appropriate without context-specific adaption. While there is much to draw on from program evaluation, particularly from participatory and equity-driven methodologies, a clear message from our interviews was that it is not enough to just retrofit conventional approaches – we need to shift our evaluation mindsets and toolkits.

To evaluate field catalysts, methodologies need to wrap around the model and utilise key strategies to generate learnings and findings that can be "rolled up" to evidence initiative- or organisation-level impact. While we are still exploring, innovating, and learning what methods work, we do know that a systems-aware toolkit will utilise mixed methods for data, measurement, and evaluation; give attention to process indicators; and facilitate feedback loops for rapid and continual learning across networks. Annex 1 outlines a few relevant tools being applied.

"You've got communities for whom this is actually their lives… for whom these are life and death questions. And then you've got other people who might do some work on this for a while and move on to a different kind of sector. And I think acknowledging the values that everyone brings and the perspective and positionality and having some reflective work around that is really critical."

– Key Informant

SECTION 2

Challenge 2:

Conventional evaluation planning doesn't hold up when the "thing" being evaluated is dynamic

2.1 ABOUT THE CHALLENGES

Across the initiative phases, a **challenge for evaluation is to stay adaptive and useful**. This is particularly tricky in the early stages of system change initiatives. Both the strategies and results of systems change initiatives are emergent and the nature of the "thing" we are evaluating is dynamic. The journey is unprecedented and unpredictable, and plans are iterated or dropped as more is learned. Boundaries of system change initiatives are harder to pin down and less stable, as the scope tends to change and pivot. As one key informant noted, given there is no definitive guide for how to do the work, people can be left with no choice but to reinvent the wheel and there is a lot of this happening across initiatives.

For evaluation, a key challenge is not moving into technical evaluation planning too quickly, and positioning evaluation so that it can be useful in 'real time' and not stifle innovation.

One strong theme we heard is that **conventional approaches to evaluation planning are not suited** to such dynamic initiatives. In this context, commonly used 'waterfall' approaches – where the evaluation is planned in detail and then implemented – can risk impeding progress or leading to budget overspend as evaluators constantly update frameworks. Conventional procurement can contribute to this challenge, as setting up pre-defined milestones can undermine the reflexivity needed to walk alongside systems change initiatives.

To support adaptive and emergent systems transformation, evaluation planning needs to incorporate learning and adaptation. However, there is a **lack of capacity, capability, and investment for learning-focused evaluation**, and our understanding of and readiness for learning

at scale is still developing. When investment in learning does happen, it is sometimes met with resistance as some players caution it takes time away from the 'doing'. It can be uncomfortable to hold the necessary conversations for learning that help move the work forward. Another challenge is balancing conventional expectations of subject matter experts (often seen as having the answers) with the adaptive learner stance needed for try, test, learn approaches.

Again, these challenges are accentuated in the field catalyst model. Field catalysts are likely to have little direct control over the course of the work and will experience changes in direction as the leadership group learn, adapt and seize opportunities and policy windows. Furthermore, for field catalysts, the challenge of planning for and facilitating learning is not just at the individual or group level, it is at the systems scale. Across scales and contexts, creating fast-paced insight generation and feedback loops for sharing is challenging. Therefore, to suit the dynamic and adaptive nature of field catalysts, evaluation planning and design require tailored approaches.

2.2 EVALUATION APPROACHES FOR CHALLENGE 2

To ensure evaluation is adaptive and useful, we suggest three evaluation approaches:

a) A phased evaluation strategy that balances planning and rigour with flexibility

b) Developing principles to guide evaluation over time

c) Learning-focused evaluation

2A.

A phased evaluation strategy that balances planning and rigour with flexibility

A key strategy to working with complexity is to **balance planning and rigour with flexibility**. Evaluation planning will need to be stepped out and iterated over phases. Planning needs to strike a balance between having enough of a framework to give direction to the evaluation activities, whilst also being able to account for significant shifts in the initiative's strategy.

Agile planning is crucial for driving the development and adaptation of not only the initiative but also the evaluation. As Mark Cabaj writes, evaluators should "resist the temptation to build a 'perfect' and 'fixed' evaluation design from the get-go, and instead put together the 'best possible' design to begin with, and then continue to develop, adapt and refine their indicators and methods over time."[27] This can help avoid the problem of doing a 'big design up front', which is time-consuming, expensive and can create rigidity. Evaluators can also write up concept notes in the early phases that foreshadow the shape and the approach to evaluation, whilst avoiding detailing specific indicators and methods until the design is more formed.

The complexity and emergence of the work also mean that **learning is key across all phases** – not only through developmental evaluation but also through continual learning that deepens contextual understanding, drives innovation, and supports practice, adaption, and improvement. To be effective in the development phase, evaluation needs to be more generative than judgemental. It should aim to support evaluative thinking, shared learning, and feedback loops, wrapping around key parts of the work.

"Short feedback loops are probably much more important in the short term… Because you don't know what's going to happen and you're taking your best guess as you start out. Having a mechanism where you can do that kind of rapid testing, trying and adapting is really important."

– Key Informant

Some specific evaluation approaches that can facilitate agile evaluation planning and practice include:

- **Principles-focused evaluation:**
As discussed in Challenge 1, when in the early messy stage, sometimes all that is possible is to evaluate against a set of pre-agreed principles[28]. This involves reflecting on the initiative's principles, whether they are meaningful and actionable, whether they are being followed, and whether this is leading to desired results.

- **Co-create a dynamic theory of change:**
A high-level theory of change that can be refined over time to reflect new learnings and changes in the initiative's strategy can help manage expectations about the work of field catalysts, expected results, and pace of transformation. It needs to be created in a way that acknowledges complexity and interconnection; as one key informant stated, there is still a tendency for people in the change-making or government sectors to want a silver bullet solution. While theory of change models have limitations in this context, they can be a powerful participatory tool to scaffold design and across phases. Developing theories of change for complex settings will need to be tailored and will differ from program-focused theory of change[29], often having a focus on different scales, levels of change, and actors[30]. The model can be layered to represent the different levels of change, such as enablers, systemic change, instances of change, and mission-level change. This is a key difference to programmatic models and there is no one-size-fits-all approach.

- **Use an 'umbrella' evaluation framework with patch evaluation strategies:**
An overarching evaluation framework that has enough structure but not too much detail is good to accommodate changes while setting out the high-level approach, objectives, and principles of evaluation across phases. For field catalysts, the framework will need to pay attention to the different roles, elements and processes of the structure. Evaluation strategies for sub-components can wrap around particular areas of the intervention as 'patch evaluations'[31], in which multiple evaluative processes (rather than one standard design) are planned as needed. The umbrella framework can include guidance across measurement, evaluation, and learning, including how to aggregate initiative-level results and insights.

- **Facilitate short learning cycles:** Facilitating short loops of learning and adaptation can support innovation and results. This is especially important in the early stages to set the initiative up for the long term.

- **Be prepared to iterate inquiry frameworks over time:**
The focus of key evaluation and learning questions is likely to shift over time. They may look different from conventional key evaluation criteria and evolve to reflect the current phase of the initiative. At different points, key questions might include: the model and success factors; context; rapid learnings; process and progress; strategy review for optimisation; evaluating key roles, governance, and funding models within the catalyst structure; scalability of interventions for impact; contribution to results; or how self-organising and self-sustaining the system conditions are.

2B.

Developing principles to guide evaluation over time

Agreeing on a set of principles for the evaluation approach can help us to make decisions and guide it as we go. Principles are highly relevant when there is no one-size-fits-all evaluation solution or detailed upfront plan. Some principles that were highlighted as relevant for field catalysts are below:

- **Be participatory, share power, and promote self-determination.** This includes co-designing and implementing evaluation in ways that genuinely reflect the principles of the initiative and that acknowledge evaluation as a catalytic intervention that also has the potential to reinforce and disrupt systems conditions. Key informants talked about sharing power to enable participation and leadership, whilst similarly Mark Cabaj has noted the importance of getting serious about 360 stakeholder judgement, particularly in interpretation of lessons and results[32].

- **Cultural appropriateness and inclusion** across evaluation design, data collection, sense-making and the communication of the findings. Several key informants emphasised the need to place First Nations peoples at the centre of both systems change work and the evaluation. For meaningful inclusion, a systemic approach to inclusion is needed[33]. Furthermore, consideration should be given to cultural inclusion and appropriateness in terms of multiculturally valid methods and processes oriented towards participant ownership[34].

- **Integrate an explicit equity lens** into evaluation planning and practice. This involves aiming to understand how historical and structural conditions have contributed to the conditions being addressed, the effect of strategy on different demographic groups and systemic drivers of inequity.

- **Strong evaluation governance** is important and can ensure evaluation maintains broad accountability and an equity and inclusion lens[35]. This is significant because the tensions in the system across different groups will play out in the evaluation too, and it is important to have intention about who is brought together and how, to avoid reproducing dynamics that hold the system in place.

- **A partnership approach to funding and evaluation**. We heard that funding relationships were particularly meaningful and effective when they were built in the spirit of trusting partnership, rather than retaining an arm's length as in more traditional funding relationships. Another theme was the importance of evaluators working in partnership with intermediaries, rather than coming in as external consultants to do an 'objective' assessment. We also heard it can be preferable to frame the evaluator as a co-learner[36], with one key informant emphasising that effective evaluation processes were deeply relational and involve evaluators being deeply connected into the team. Evaluators with 'one foot in, one foot out' afford a different perspective and the ability to 'get up onto the balcony and see the dance floor', which is hard when you are 'in the washing machine'[37].

- **Making systems change and evaluation more accessible**: Evaluation can play a role in creating a shared understanding and language about systems change, and as a means to engage practitioners, policymakers, and local partners. Integrating evaluation and communications – especially in communicating complex data and findings simply – was also mentioned as an important way for evaluation to stay accountable to stakeholders and enable them to shape and direct the work.

"It is really useful to have someone who's a kind of insider-outsider who brings that expertise through working across different projects in this way, but who is trusted and who understands the detail, so it's not our community collaborators having to endlessly explain and prove their knowledge or expertise."

– Key Informant

2c.

Learning-focused evaluation

As discussed in Challenge 1, learning-focused evaluation helps address the challenges of complexity and supports adaptive leadership. Strategies for learning will be multiple, overlapping, and will need to be structured for emergence. For field catalysts, it is necessary to plan and embed strategic learning and feedback loops at various scales: individual, group, and system level[38]. Below, we provide general insights into learning approaches relevant to field catalysts, which will need to be designed and planned across the various scales relevant to the model and learning objective.

Embedding strategic learning

Whilst this is common knowledge for many working in evaluating systems change, the interviews and document review reaffirmed the vital importance of **strategic learning** in systems change initiatives. Strategic learning is defined as using data from a variety of sources to inform decision-making about strategy[39]. Social problems and interventions "are not static but inherently reflexive"[40], and strategic learning helps drive adaption and innovation when working in complexity. This includes learning about the system and how we show up through trying, testing, and reflecting. It is an effective way to stay attuned to where the system is moving and what is needed at that time, rather than sticking to a plan that soon may no longer be appropriate and may even lead to recreating the system. Including strategic learning in your evaluation design and planning will be critical for success.

"Rigorous strategic learning is not a technical problem solved by simply having the right tool, the right template, or even the right data and findings at hand. It is a practice, a way of working and thinking, a set of habits — a capacity. As such, it must be cultivated over time and in a way that clearly connects to program staff needs."[41]

Embedding strategic learning will involve shifting mindsets, practices, power relations, and relationships as well as creating deliberate processes and structures that promote learning and utilisation[42]. Below are a few key considerations for embedding strategic learning in this context:

- In a **learning culture**, learning and adaptation is valued as inherent to the work, rather than seeing any deviation from a delivery plan or intended outcomes as failure. A learning culture is required for an initiative's sustainability, meaning that habits and ways of working need to be set up early on. Harnessing, documenting, and sharing learning at scale and cultivating the culture of learning takes commitment, resourcing, and deliberate structures from the start. One key informant said that if this learning culture is not built early on, or if funders enter the picture in future – which is highly likely in a field catalyst model – and they want to apply more traditional approaches to evaluation, it can be harder to push back on this. Another key informant noted that having a **learning mindset** is crucial to both addressing and evaluating complex, adaptive challenges (as well as having context-specific expertise for systems change).

- **Capability building** is important for learning, within the intermediary team and beyond. From what we heard, for field catalysts to convene or support learning sufficiently, we should expect that capacity and capability building for learning is a given for the intermediary and the network. For funders and boards, it is likely that a huge uplift in skills will be needed to step into different models of accountability for learning and adaptation. Mark Cabaj suggests that to build this capability for learning, smaller scale experiments should be introduced in order for participants to observe the value it has for the work. This can then be scaled and integrated into the rest of the work so that learning is experienced as integral than just something extra[43].

- **Learning can be uncomfortable**. Many collaborating organisations or players are used to working in competitive funding environments, where it is not safe or encouraged to share or learn from 'failure', or have been previously disempowered or not had a seat at the table.

- **Processes and structures for learning** will need to be tailored for the field catalyst, and may involve shared sense-making, routine reflection, communities of practice, regular rounds of Most Significant Learning, experimental action sprints, peer-to-peer supervision, and coaching. One key informant suggested that an important way to embed and maintain a learning culture is through having a rhythm of reflection every month so that it forms a habit. As well as driving innovation, convening shared learning builds trust and a better understanding of the context and evaluation across the network, and also means that intermediary staff build their evaluation capacity through doing.

Establishing learning networks and feedback loops at scale

Approaches that support **feedback loops for shared learning** across systems, or between local and national levels and diverse contexts are still in the early days of being explored, tested, and developed. Learning at a systems scale will need a long-term orientation. Processes and outcomes can take much longer than individual and group learning. The momentum and benefits of system scale learning are slower to happen and do not present equally across the system[44].

One process that can help achieve learning at scale is **learning networks**. Learning networks are interactive processes for engaging collaborators and cross-system players to continue to share learnings and experiences related to applied systems thinking across organisational boundaries, while drawing on diverse sources of knowledge, experience and capabilities[45]. One key informant spoke about the need for networks where people working in this space can share resources from different places and showcase examples. Learning networks serve the work and evaluation simultaneously. They support feedback loops for mutual learning across the collaborating systems players and for this to be utilised and applied in people's own context, organisations, networks etc.

If field catalysts have knowledge sharing, translation and learning as a key strategy, then investment in and evaluation of learning networks will be important. For evaluators, it can also mean walking alongside intermediaries with 'one foot in, one foot out' to support the learning network through developmental evaluation and documenting learning for sharing. We also need to track the effectiveness and impact of the learning network. For example, this may involve evaluating the process, progress, and outcomes of capacity building, knowledge sharing, and ongoing learning related to application of systems thinking.

"We need to give people doing systems change time to reflect and think… but also evaluators need time to share and think and integrate resources together. That's when we have 'aha' moments."

– Key Informant

SECTION 3

Challenge 3

Determining contribution and managing the politics of impact claims

3.1 ABOUT THE CHALLENGES

In highly complex systems-change contexts, a key challenge is how to evaluate the contribution that an intermediary has made to results, as conventional notions of causality and attribution don't hold up well here. The challenge of determining and managing claims of contribution is both technical and political.

Part of the challenge of assessing contribution comes from the **role of intermediaries** and issues of who should be claiming impact. Because field catalysts play a facilitator's role - supporting, connecting, and amplifying other groups and organisations – their influence on the system is **indirect**. There are also many players who may not be directly linked with the intermediary working towards similar changes. Mark Cabaj notes this makes conventional attribution analysis poorly suited to evaluating field-building intermediaries[46].

Methods to evaluate the contribution of intermediaries are especially underdeveloped, yet are needed to understand and evidence the deliberate influence and impact intermediaries are having in addressing complex issues. While some emerging approaches do exist, they are not yet tested for this context.

In addition, the layered structure and the way intermediaries work (with a facilitating partner, leadership and governance structure, wider systems players, etc.) compounds the challenge as we need to **'untangle' contribution to impact** across these layers. The field catalyst model

and facilitating partner (often the main funded party) is even more hidden, working backstage. Their work is not implementing activities but convening and supporting the leadership group and system players to build the field for systems transformation and to influence systems conditions. Unpacking the work and influence through the leadership group is also challenging, as this happens in their organisations and own networks.

For field catalysts, getting clear on whose impact it is and managing the sensitivities around the intermediary as a whole and its parts is difficult. A challenge will be making sure the very act of claiming impact doesn't go against the spirit of shared measurement and purpose. Cabaj cautions:

"the more that intermediaries try to "drive" and "organise" actors – or take credit for such progress – the more they are apt to alienate the many other actors whose participation is critical, and diminish their own credibility and usefulness in the change process."[47]

3.2 EVALUATION APPROACHES FOR CHALLENGE 3

To address the challenge of determining the contribution and impact of a field catalyst we suggest:

a) Focus on contribution not attribution

b) Get clear on the contribution of the 'whole' and unpacking contribution across the layers and parts of the field catalyst model.

3a.

Focus on contribution not attribution

Evaluating field catalysts entails understanding and tracing their **contribution to results**. Notions of attribution don't hold up given that there are multiple factors likely to be contributing to change. We are interested in an intermediary's contribution to both intended and unintended results, as well as wider ripple effects. It will be important to consider the contribution of single interventions towards understanding the combined effect of multiple players and interventions.

Identifying contribution better fits with an **intermediary's role of influencing and supporting change** from behind the scenes, working with many other field actors. Intermediaries contribute to, rather than drive mission-level changes. Mark Cabaj suggests that initiatives should embrace a 'contribution mindset' for two key reasons: 1) intermediaries are required to 'nudge' powerful systems actors, as they are not in a position to drive and organise such a distributed system, and 2) claiming credit can potentially alienate other stakeholders and is politically sensitive (see Section 3b)[48]. Cabaj's paper

explains how understanding the contribution of intermediaries involves tracking and reporting on results around their inputs, outputs, and the immediate influence on actors and organisations that they support, as well as the more difficult-to-track and evidence "downstream effects of their work".

Evidencing contribution is about **establishing a reasonable case** for how a field catalyst contributes to and/or accelerates results, or what would have happened without it. Tracing the role and contribution of the model and strategies to emergent results means looking at them as connected to yet distinct from the contribution of the whole systems effort. It will also involve identifying and assessing other contributing causes for the results. While we should aim for rigour, we need to be explicit about uncertainty in judgements and that at best we can use evidence to draw credible conclusions between systems influence and change[49].

Relevant approaches to assessing contribution include **methodologies** such as 'contribution analysis' and process tracing, which offer rigorous ways to explore causal links and competing explanations[50]. However, they have not been adequately tested and applied to field catalysts as yet. Other potential approaches to be explored include Scriven's General Elimination Methodology[51], which is a forensic approach that evaluates impact by eliminating alternative explanations until the most valid explanation remains[52], or lighter analysis tools such as the 'What Else Test'[53]. While there is still much debate about how to undertake and present contribution analysis effectively, there has been a particular focus on how to build a stronger sense of the counterfactual (what would have happened without the intervention). Of late there has been emerging work on building counterfactual scenarios based on historic data as well as synthetic counterfactuals.

3b.

Manage the political nature of impact claims

For field catalysts, **getting clear on whose impact it is** and managing the sensitivities around the intermediary as a whole and its parts will be paramount. When demonstrating the contribution of field catalysts, we need to watch out for overclaiming and get clear on how field catalysts affect change. This may involve clearly articulating the difference between the impact of the field catalyst and that of other models or wider efforts they are working to support.

In the Health Justice Australia example below, one of the field catalyst's key strategies was to support an existing systems-wide partnership model to drive systems change. The example highlights the challenge and importance for getting clear on the difference between the impact of the field catalyst and that of the models or wider efforts they support.

Health Justice Australia – Rethinking results and getting clear on contribution

Health Justice Australia (HJA) is a not-for-profit field catalyst established in 2016 that supports health justice partnerships across Australia. Its purpose is to influence systemic impact related to better health and justice outcomes for people experiencing intersecting health and legal need. Informed by the partnership focus, they run three key initiatives: knowledge and its translation; building capability for collaboration; and driving systems change by connecting cross-system players to drive reforms to policy setting, service design and funding.

One of the challenges encountered has been the conflation of HJA model and the health justice partnership model it supports. HJA found that funders were not asking about the evidence behind HJA or its work, but instead were wanting evidence to support the concept of health justice partnerships. This was happening at the funding, activity and outcome level. Setting expectations and keeping focus on HJA outcomes and accountability has been challenging. HJA does not wish to claim the outcomes of the health justice partnerships for itself or create expectation failure, given the outcomes they can influence are for improved capability and capacity of the sector.

Another challenge for HJA has been in articulating evidence and progress, which has meant securing funding has been difficult due to a relative lack of evidence, track record, or demonstrated impact (particularly in the beginning). In rethinking results, the initial funding from sponsor Paul Ramsay Foundation was not focused directly on the delivery of outcomes, but instead on 'making the case' for health justice partnership and HJA. This involved gathering evidence and creating a track record to support further funding. Even though an evidence base and track record has now been established, data and measurement continue to be challenging for the field catalyst due to the nature and complexity of the work.[54]

As well as understanding the contribution of the whole, we may need to explore the *contributing roles of the different layers across an intermediary* (i.e. unpacking the unique contribution story of the facilitating partner, leadership and governance structure, wider systems players, etc.). This will help generate a deeper understanding of the model and how they influence change. However, claiming impact is political. In some cases, it can potentially be counterproductive to the whole idea of working together as a collective towards shared impact. As well as the technical approaches needed to tackle evaluating contribution, it requires capacity building and learning mindsets to hold the necessary processes to reflect and unpack results and contribution in participatory, inclusive, and rigorous ways. The trust, relationships, culture, and processes that are set up for evaluation more broadly will be a key enabler when it comes to navigating the potential sensitivities of impact claims and contribution.

Final Thoughts

One key insight from undertaking the discovery work is that understanding of, and codification about, field catalysts is still in its early days. While we tried to reflect on field catalysts specifically, given there is so little in the public domain on this topic, the findings still land fairly generically around the challenges faced by field-building intermediaries and systems change initiatives more broadly. Ultimately, we need to better understand how field catalysts work to effectively wrap evaluation around them. Future work is needed to further explore with specificity the challenges and approaches for evaluating field catalysts.

Our starting assumption – that field catalysts experience similar challenges as other types of intermediaries, but with greater intensity – held up through the discovery work. However, we will continue to test this through our work with the Early Years Catalyst. Our hunch is that the challenges of systems change and evaluation are amplified for field catalysts given their wide scope, long 'on-ramps' to results, and indirect role in influencing systems change. This creates technical and political challenges for evaluation, such as for determining and managing contribution claims to impact. Further, a field catalyst's challenges are felt even more keenly in the early design and start-up phases of systems change initiatives.

A strong theme from key informants and in the literature is the call for a radically different approach to framing and measuring progress and results given the limitations of conventional evaluation to handle the complexity of field catalysts. The urgency of the call is heightened given there is such a strong commitment to being data- and evidence-driven in this work. There is immediate and significant work to be done to support the data needs and capacity of intermediaries and their networks, so nuanced systems-level results and lessons can be tracked and shared, and so measurement is kept meaningful, useful, and relevant to the collective effort, context, and diverse stakeholders involved. There were also calls for: reframing accountability and what constitutes results for field catalysts; the valuing and inclusion of lived experience; and that equity and learning focused evaluation are essential for this work.

KEY TAKEAWAYS

In closing, we package up some of the key takeaways for evaluation presented in the concept paper:

- **Stay flexible** and take a phased approach to evaluation.

- **Use developmental evaluation,** particularly in the early phases of a field catalyst's design and implementation.

- **Invest in co-defining what "good" looks like**, even though this will change as you go – for example, basing evaluation against an iterative theory of change that orientates to the future and what you expect to see.

- **Broaden notions of accountability** to include being accountable to learning and to wider cohorts of stakeholders.

- **Be participatory and inclusive**, for example by including funders and beneficiaries in evaluation design and implementation.

- **Include principles** and ways of working in evaluation, not just activity.

- **Rethink how 'results' are framed and evaluated** to include the process, enabling conditions, systemic results, and the mission and/or population level.

- Use a **mixed method, systems-aware toolkit**.

- **Lean into learning-focused evaluation** that embeds learning at different scales and across the field.

- **Maintain a contribution mindset**.

- **Get clear on the contribution of the 'whole'** as well as unpacking contribution from different parts of intermediary models.

- **Set up an evaluation governance group** – implementation needs to be held by a small group.

Annex 1

SYSTEMS-AWARE TOOLKIT

Below are a few relevant systems-aware methodologies and tools being applied. This is not intended as a complete toolkit, but rather is intended to highlight a few examples identified during the research.

Data and measurement

A clear message from the discovery research is that a data-driven approach and the ability to articulate results and demonstrate impact is essential for intermediaries and field catalysts. As field catalysts are set up and started, **data and the development of indicators** (measures) will come into focus. Data and measurement will need commitment and data literacy and is an area that needs further exploration and innovation for field catalysts, including data and development of indicators that allow for comparison of applied systems thinking approaches across contexts.[55]

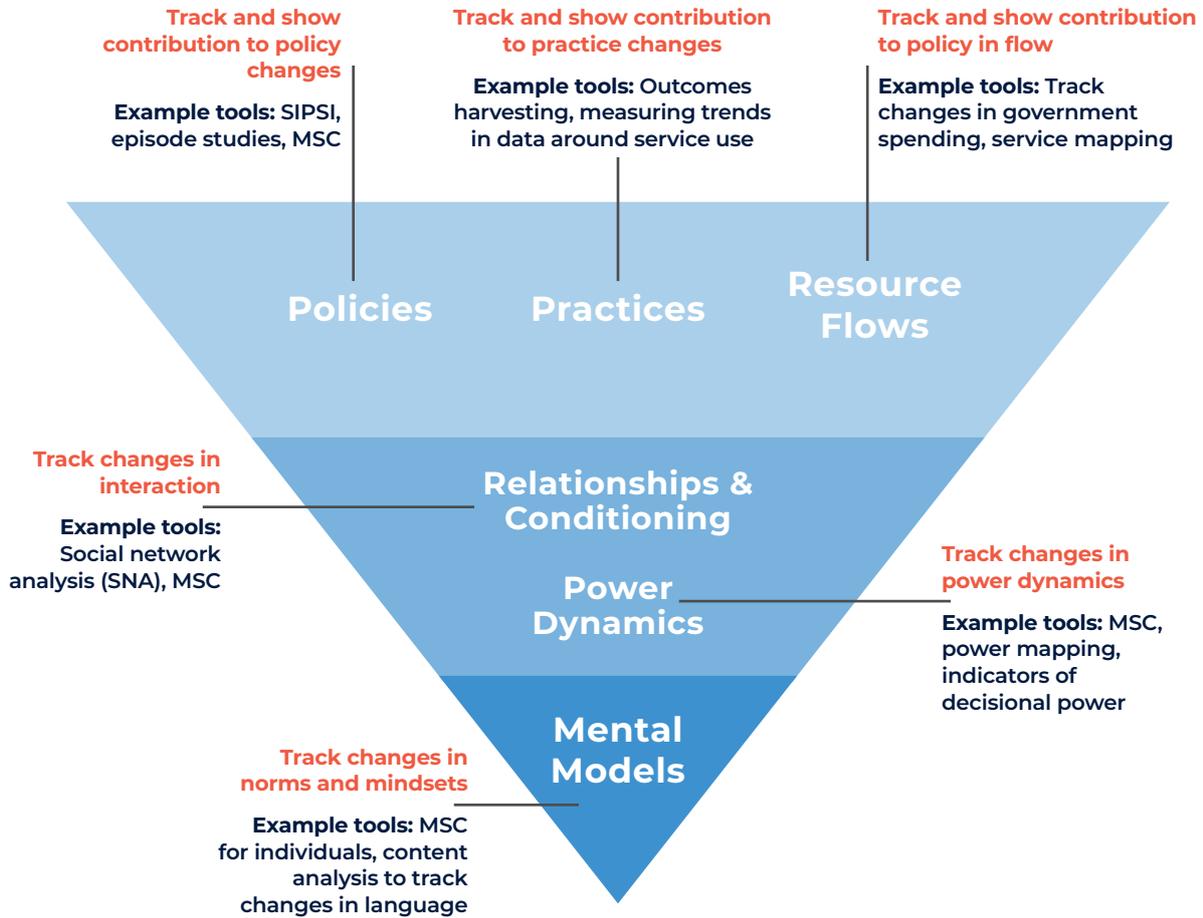Below are a few considerations to keep in mind:

- What constitutes **relevant, meaningful, and appropriate measures** will depend on how the catalyst field model intends to influence systems change, the context, and intended results. We heard it is best to start small and not develop too many too early. It can also be helpful to create agreed criteria for selecting measures (through a participatory and inclusive process), particularly when initiatives involve multiple and diverse players, communities and sites/places.

- Milner et al. recommend giving attention to **metrics to measure systems thinking process and impact**, which may include metrics for mechanistic approaches that aim

to improve system performance (through outputs and outcomes) and/or ecological approaches where learning networks are used as a key activation lever and the focus is more on process and participation rather than performance.[56]

- For some initiatives, **'shared measurement'** may be relevant, where an agreed set of measures are used to monitor progress, process, participation, and/or performance. Shared measures can enable a group or collective to determine what success looks like and scaffold a data-driven approach to learning and measurement[57]. However, a note of caution – shared measurement is not always necessary and determining them in participatory and inclusive ways is challenging and takes time. Sometimes groups can get 'stuck' or end up overinvesting in measurement[58]. While important, shared measurement is not sufficient as a sole approach, and effort and time should also be invested in evaluation.[59]

For measurement and evaluation, **data collection methods that cast an open net to identify both expected and unexpected results** are relevant. For example, inductive evidence-gathering methodologies such as Outcomes Harvesting and Significant Instances of Policy and Systems Influence (SIPSI) are useful for tracking change and showing the contribution of field catalyst influence (e.g. new/expanded partnerships, policy, or the reflexive practice of key players). Story and community voice collection methods – such as community conversations, Most Significant Change technique[60], and impact yarns – can be good for identifying a diverse range of systems changes. Systems evaluation and analysis techniques, such as social network analysis or the use of bellwether indicators, can evidence pre and post systems conditions.[61]

Below are a few examples of data collection methods laid over the Water of Systems Change model.[62]

Track and show contribution to policy changes

Example tools: SIPSI, episode studies, MSC

Track and show contribution to practice changes

Example tools: Outcomes harvesting, measuring trends in data around service use

Track and show contribution to policy in flow

Example tools: Track changes in government spending, service mapping

Policies

Practices

Resource Flows

Track changes in interaction

Example tools: Social network analysis (SNA), MSC

Relationships & Conditioning

Power Dynamics

Track changes in power dynamics

Example tools: MSC, power mapping, indicators of decisional power

Track changes in norms and mindsets

Example tools: MSC for individuals, content analysis to track changes in language

Mental Models

Figure 3:

Examples of data collection methods against The Water of Systems Change's six conditions for change.

Logs are useful for tracking both process and changes and can support building an evidence base. They are useful for keeping track of engagement, participation, pivots, or impact. A log is a shared document or record-keeping system where entries can be entered and tracked over time (e.g. Excel, Microsoft Forms). Impact logs, which record emerging changes, are beneficial when multiple stakeholders contribute to recording data, as the field catalyst can harness diverse 'eyes and ears' across the system. One key informant relayed how during coaching sessions they would ask the intermediary team to tell them updates about what they had seen shifting, which not only helped to capture change but also document the increasing literacy of the intermediary.

Evaluative tools and reporting

- **Partnership assessment tools** help collect data and analyse how stakeholders are experiencing the partnership and collaboration. A bespoke and simple version could be developed based on existing comprehensive tools such as the VicHealth partnerships analysis tool[63] or Collaboration Health Assessment Tool.

- **Rubrics:** Rubrics can be used for learning, progress mapping, or performance measurement. They can help by defining what success or progress looks like over phases of the field catalyst, across key pillars, strategies or principles. Developing rubrics can be participatory and enable us to articulate practical demonstrations of desired systems changes or influence (such as related to relationships, power, trust, uptake of ideas, capacity, or resource flows). They can also be used to assess scalability of innovations or to help be transparent about the strength of evidence and confidence in conclusions (such as for contribution ratings).

- **Digital dashboards** can display real-time process and results metrics that support co-design and innovation, or outcomes measurement.

Learning and shared sense-making

- **Gathering and documenting learning** can occur by identifying key reflections/insights and critical incidents that highlight that the organisation/system players are aware of their assumptions and are applying insights to develop and adapt its strategies[64]. Systematic processes for harvesting learnings include tools such as Most Significant Learning technique[65] to collect stories of learning, or a learning log to capture moments of key learning and adaptation. Particularly in the early phases, short learning summaries, documentation of applied systems thinking, and journey mapping can help drive innovation and start building the evidence-base.

- **Triple loop learning framework:** see Tamarack Institute's three loops tool[66]. Triple loop learning includes reflection on how we are showing up, as well as what is happening in the work and the system.

- **Shared learning and sense-making** can be promoted through reflection workshops and learning circles that bring together different stakeholders[67]. One key informant recommended that when conducting reflection and learning sessions, it can be more productive to cover one or two questions rather than overloading the agenda, saying "deeper is better than broader". This is because it allows people to make sense of things and get to the things that sit below the waterline, like beliefs and mindsets that are not as obvious in the day-to-day work. These sessions give people permission to slow down and do the adaptive work that is so important for complex systems change problems, according to a key informant.

References

[1] JESSICA DART *"Place-based Evaluation Framework: A national guide for evaluation of place-based approaches"* commissioned by the Queensland Government Department of Communities, Disability Services and Seniors (DCDSS) and the Australian Government Department of Social Services (DSS), 2018.

[2] SOCIAL VENTURES AUSTRALIA. *"Insights on Australian field-building intermediaries and their funding journeys towards sustainable impact"* 2022, https://www.socialventures.com.au/assets/Insights-on-field-building-intermediaries.pdf.

[3] STEPHANIE DODSON CORNELL *"How Can Intermediaries Accelerate Social Change?"* Social Innovation Forum, published on 15 February, 2018, https://www.socialinnovationforum.org/blog/how-can-intermediaries-accelerate-social-change.

[4] TAZ HUSSEIN, MATT PLUMMER AND BILL BREEN. "How Field Catalysts Galvanise Social Change", *Stanford Social Innovation Review,* Winter (2018): 48-54, https://ssir.org/articles/entry/field_catalysts#:~:text=Surmounting%20daunting%20social%20challenges%20such,amplifies%20the%20efforts%20of%20others.

[5] SILVIA CHEUY, MARK CABAJ, AND LIZ WEAVER. "How Field Catalysts Accelerate Collective Impact" *Stanford Social Innovation Review,* published 4 January, 2022, https://ssir.org/articles/entry/how_field_catalysts_accelerate_collective_impact#.

[6] SEE: FARNHAM, LIJA ET. AL., *"How Philanthropy Can Support System-Change Leaders"* The Bridgespan Group, 2021, https://www.bridgespan.org/bridgespan/Images/articles/how-philanthropy-can-support-systems-change-leader/how-philanthropy-can-support-systems-change-leaders.pdf; SILVIA CHEUY, MARK CABAJ AND LIZ WEAVER, "How Field Catalysts Accelerate Collective Impact"; AND TAZ HUSSEIN, MATT PLUMMER AND BILL BREEN, "How Field Catalysts Galvanise Social Change".

[7] For examples, see SOCIAL VENTURES AUSTRALIA, *Case studies of Australian field-building intermediaries.* 2022, https://www.socialventures.com.au/assets/Case-studies-of-Australian-field-building-intermediaries.pdf=.

[8] SOCIAL VENTURES AUSTRALIA. *Insights on Australian field-building intermediaries.*

[9] MARCIE PARKHURST, HALLIE PRESKILL, JEWLYA LYNN, AND MARAH MOORE. *"The Case for Developmental Evaluation,"* FSG, published 1 March, 2016, https://www.fsg.org/blog/case-developmental-evaluation/.

[10] BETTER EVALUATION, *"What is Evaluation?"* Better Evaluation, 2022, https://www.betterevaluation.org/en/what-evaluation#:~:text=Formative%20evaluation%20refers%20to%20the,how%20it%20is%20being%20implemented).

[11] BETTER EVALUATION, *"What is Evaluation?"*

[12] CATALYST 2030, *"About Systems Change"*, published 2022, https://catalyst2030.net/what-is-systems-change/

[13] It is important to clarify that we are not talking about 'systems evaluation', which involves diagnosing and tracking the conditions of a particular system, and rather are focused on how to evaluate systems change initiatives.

[14] Snowden's Cynefin framework can be useful in diagnosing complex problems. See the following video for details: *"The Cynefin Framework"* YouTube, 12 July, 2010. https://www.youtube.com/watch?v=N7oz366X0-8.

[15] BEVERLY PARSONS, *"Using Complexity Science Concepts When Designing System Interventions and Evaluations,"* InSites (2012), https://insites.org/resource/using-complexity-science-concepts-when-designing-system-interventions-and-evaluations/.

[16] TAZ HUSSEIN, MATT PLUMMER AND BILL BREEN, *"How Field Catalysts Galvanise Social Change."*

[17] MICHAEL QUINN PATTON, *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, Guilford Press, New York, 2010.

[18] This was raised by multiple key informants and more can be read about this from the literature, such as in Mark Cabaj, Evaluating the Results of Intermediary Organisations, Paul Ramsay Foundation (2021).

[19] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[20] KATHERINE MILLIGAN, JUANITA ZERDA, AND JOHN KANIA, *"The Relational Work of Systems Change,"* Stanford Social Innovation Review, 2022, https://doi.org/10.48558/MDBH-DA38.

[21] JOHN KANIA, MARK KRAMER, AND PETER SENGE, *The Water of Systems Change*, FSG (2018), https://www.fsg.org/resource/water_of_systems_change/.

[22] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[23] CHARLES LEADBEATER AND JENNIE WINHALL, *The Patterns of Possibility: How to Recast Relationships to Create Healthier Systems and Better Outcomes*, System Innovation Initiative, 2022, https://static1.squarespace.com/static/6166ed8d6518934cae8a0e0a/t/621f82b11b697f6a6530a6ac/1646232278572/SII_ThePatternsOfPossibility_2022.pdf.

[24] KATHERINE MILLIGAN, JUANITA ZERDA, AND JOHN KANIA, *"The Relational Work of Systems Change."*

[25] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[26] JESSICA DART, *Place-based Evaluation Framework*

[27] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*

[28] MICHAEL QUINN PATTON, *Principles-focused Evaluation: The Guide,* Guilford Press: New York, 2017.

[29] Examples approaches include: *Linda Doyle, Change & Complexity: Vector Theory of Change*, The Cynefin Centre, (2022), https://cdn.cognitive-edge.com/wp-content/uploads/sites/12/2022/02/02160119/VTOC-paper-2022.pdf; and Transformative Innovation Policy Consortium (TIPC) and Utrecht University Centre for Global Challenges, Motion Handbook: Developing a Transformative Theory of Change, EIT Climate-KIC, 2022, https://transitionshub.climate-kic.org/publications/motion-handbook-developing-a-transformative-theory-of-change/.

[30] JESSICA DART AND ZAZIE TOLMER*, "Agile theory of change for system change endeavours,"* Clear Horizon, n.d., https://clearhorizon.circle.so/c/place-based-approaches/toolkit-for-evaluating-place-based-delivery-approaches.

[31] MARK CABAJ IN MICHAEL QUINN PATTON, KATE MCKEGG, AND NAN WEHIPEIHANA, *Developmental Evaluation Exemplars: Principles in Practice,* Guilford Press, New York, 2015.

[32] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[33] CLEAR HORIZON, *EYC Field Catalyst Review: Prepared for EYC Field Catalyst*, Clear Horizon, 2022.

[34] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[35] CLEAR HORIZON, *Literature Scan for Place-based Evaluation Framework*, Clear Horizon, 2018.

[36] NATASHA JOSHI, *"Plotting Impact Beyond Simple Metrics,"* Stanford Social Innovation Review, published 17 August, 2022, https://doi.org/10.48558/NMK0-PQ67.

[37] From several key informants and as similarly described in Charles Leadbeater, and Jennie Winhall. *"The Patterns of Possibility: How to Recast Relationships to Create Healthier Systems and Better Outcomes."* System Innovation Initiative, 2022, https://static1.squarespace.com/static/6166ed8d6518934cae8a0e0a/t/621f82b11b697f6a6530a6ac/1646232278572/SII_ThePatternsOfPossibility_2022.pdf

[38] FIONA MCKENZIE, *Building a culture of learning at scale: learning networks for systems change*, Orange Compass, commissioned by the Paul Ramsay Foundation, 2021, https://www.orangecompass.com.au/images/Scoping_Paper_Culture_of_Learning.pdf.

[39] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[40] NATASHA JOSHI, *"Plotting Impact Beyond Simple Metrics."*

[41] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[42] FIONA MCKENZIE, *Building a culture of learning at scale*.

[43] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[44] FIONA MCKENZIE, *Building a culture of learning at scale*.

[45] FIONA MCKENZIE, *Building a culture of learning at scale*.

[46] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*.

[47] MARK CABAJ, *Evaluating the Results of Intermediary Organisations*, p. 10.

[48] MARK CABAJ, *"Evaluating Collective Impact: Five Simple Rules"*, The Philanthropist: Vol 26: Issue 1, (2014).

[49] JIM COE AND RHONDA SCHLANGEN, *No Royal Road: Finding and following the natural pathways in advocacy evaluation*, Centre for Evaluation Innovation, (2019).

[50] See John Mayne, *Contribution Analysis: An approach to exploring cause and effect*, Institutional Learning and Change (ILAC) Initiative, ILAC Briefs, 2008, https://www.researchgate.net/publication/46472564_Contribution_analysis_An_approach_to_exploring_cause_and_effect ; and Jewlya Lynn et al, When Collective Impact Has an Impact. A Cross-site Study of 25 Collective Impact Initiatives, ORS Impact and Spark Policy Institute, Seattle, WA, 2018. For an example of a hybrid contribution analysis from the Australian context, see Clear Horizon, Logan's Community Maternal and Child Health Hubs: The Contribution Case for Collective Impact Practice, Clear Horizon, n.d. https://clearhorizon.circle.so/c/place-based-approaches/logan-together-contribution-analysis.

[51] BETTER EVALUATION, *"General Elimination Methodology,"* Better Evaluation, 2022, https://www.betterevaluation.org/en/evaluation-options/list_possible_causes_general_elimination_methodology.

[52] For an example application, see Michael Quinn Patton, *"Advocacy Impact Evaluation," Journal of MultiDisciplinary Evaluation 5*, (2008): 1-10, https://journals.sfu.ca/jmde/index.php/jmde_1/article/download/159/181/.

[53] CLEAR HORIZON, *What Else Test: A Basic Tool for Strengthening Contribution Claims*, Clear Horizon, n.d. https://clearhorizon.circle.so/c/mel/what-else-test.

[54] SOCIAL VENTURES AUSTRALIA, *Case studies of Australian field-building intermediaries*.

[55] MILNER, KATE ET AL., *Mixed-method evidence review of the potential role of systems thinking in accelerating and scaling promotion of early child development*, Melbourne, Australia: Murdoch Children's Research Institute, 2022. https://ecdan.org/wp-content/uploads/2022/08/ECD_Systems_Evidence_Executive20Summary_31.5.22_FINAL_JD.pdf.

[56] KATE MILNER ET AL, *Mixed-method evidence review*.

[57] Collaboration for Impact (CFI). (2018). Collaborative Approaches – The Collective Impact Approach. Sourced from: http://www.collaborationforimpact.com/collective-impact/

[58] MARK CABAJ, *"Evaluating Collective Impact: Five Simple Rules."*

[59] JESSICA DART, *Place-based Evaluation Framework*.

[60] MSC is participatory and involves stakeholders collecting stories about significant change and can be used to help understand the impact on individuals in an authentic and powerful manner. See the following resource: Jessica Dart and Davies, Rick, The 'Most Significant Change' (MSC) Technique: A Guide to Its Use, 2005.

[61] See the Clear Horizon resource page for links and guides for many of these examples: https://clearhorizon.circle.so/c/mel/

[62] Content from Clear Horizon Academy online course *'Evaluating Systems Change and Place-based Approaches'*

[63] VICHEALTH,*"The Partnerships Analysis Tool,"* VicHealth, published 5 October, 2011, https://www.health.vic.gov.au/primary-care/strengthening-partnerships-tools-and-resources.

[64] MARK CABAJ, *Evaluating the Results of Intermediary Organisations.*

[65] JESSICA DART AND RICK DAVIES, *The 'Most Significant Change'* (MSC) Technique: A Guide to Its Use, (2005), https://clearhorizon.circle.so/c/msc/most-significant-change-user-guide-2005.

[66] TAMARACK INSTITUTE, *"Single, Double and Triple Loop Learning,"* Tamarack Institute, n.d. https://www.tamarackcommunity.ca/hubfs/Events/Multi-Day%20Events/Community%20Change%20Institute%20-%20CCI/2017%20CCI%20Vancouver/Resources/Tool%20-%20Single%20Double%20Triple%20Loop%20Learning.pdf.

[67] CLEAR HORIZON, *Reflection Workshops or Evaluation Parties: Sense-making as a Collaborative Effort*, Clear Horizon, n.d. https://clearhorizon.circle.so/c/mel/reflection-workshop-tool.