Assessing Whether Mission-Driven Innovation Makes a Difference: Mission Impossible? Developing a Guiding Framework for the Evaluation of Five Mission Driven Environments for Health in Sweden

Anna Essén

Anna Krohwinkel

Karl Wennberg

Summary

Background. Mission-driven innovation (MDI) policies are founded on governmental attempts to address fundamental but complex societal challenges. The rationale behind such attempts is typically to influence the directionality of innovation towards addressing the perceived challenge.

This report focuses on a particular instance of MDI policy executed by Sweden's innovation agency, Vinnova: the funding of five so-called "mission-driven environments" (MDEs) in 2019. The policy in question is called 'Vision-Driven Health' and was initiated in 2019 to support the establishment of inter-organizational and cross-disciplinary coalitions that work towards a common vision and a long-term systemic transformation within the Swedish health care and life science sector.

Aim. The report aims to provide a framework for evaluating five MDEs funded by Vinnova. Vinnova asked us to consider, in particular, the role of eight "Work Principles" (WPs) they recommended the MDEs implement.

This report is the result of the first (of two) possible steps in evaluating the five MDEs. The first step is about developing a framework for evaluating MDEs. We hereafter refer to it as a *pre-study*. A second step would involve actually evaluating the five MDEs based on the framework in this report.

Methods. The report is based on selective reviews of relevant literature providing insights about best practices for setting up and governing MDE-like initiatives and possible approaches and challenges to evaluating such initiatives. We also collected empirical data about how the five Swedish MDEs operationalized the principles. We surveyed members of the participating MDEs, asking them what a meaningful evaluation could imply from their perspectives. Finally, we consulted a group of external experts on three occasions.

Findings. At an overall level, the Vinnova-recommended WPs partly align with practices recommended in the relevant literatures. However, the WPs are formulated abstractly and implemented heterogeneously by the five MDEs. We argue that this heterogeneous implementation is necessary for the MDEs to progress towards their visions but complicates a uniform set of evaluation principles. The MDEs also prioritize the WPs differently, and we observed an additional set of informal WPs.

The literature consists primarily of normative studies defining MDI and its relevance and studies that discuss sets of challenges tied to evaluating MDI policies and initiatives. Empirical studies and evaluations remain scarce.

Suggestions. Drawing on insights from the literature, we outline a *framework for formative and summative evaluation* that could be used to evaluate *the MDEs* and *the WPs with which they are set to work*. We specifically argue for combining contribution and attribution approaches to evaluation, which could include the following steps:

Formative Evaluation Steps

(A) If and to what extent the *MDE is justified* due to a "failure" of the system, market, or current development direction;

(B) If and how the *MDE's governance arrangements are purposeful, consistent, and coherent* (processes and structures; i.e., ways of working and formalized routines, standards, decisions, and rules);

(C) If and how there is a *"match" between the MDE's interventions and identified barriers* (weaknesses, bottlenecks, impeding regulations, social norms, etc.).

Formative and Summative Evaluation Step

(D) If and how the targeted overarching sociotechnical system/field demonstrates *improved performance*, such as capabilities (system functions and interactions like knowledge sharing), transition processes, and outcomes.

Summative Evaluation Steps

(E) If and how the targeted overarching sociotechnical system/field exhibits *structural changes*, such as a change in the types of innovations, new forms of cross-sectorial collaborations, or new networks constellations in the system, because of the MDE;

(F) If and to what extent there is *measurable impact* on the societal level in terms of mitigating the failure addressed and reaching the MDE's "vision" or "mission."

For evaluating specific MDEs, we conclude that the formative Steps B and C (and after the MDEs have been in operation for some time, Steps D and E, which also are discussed in the report) are of utmost relevance. Step A is a policy-mix decision, and Step F is an evaluation of the overall policy). For Steps B through E, we detail how an evaluation could be done and the type of data needed and exemplify useful methods for each evaluation step.

Continuous Evaluations

For Step B (governance arrangements), we suggest that evaluations focus on: Are the WP formulated necessary and sufficient for MDEs? Are some WPs more important than others to achieve the expected process outcomes? How do MDEs develop routines and decision rules to operationalize the WPs, and what are the results of their progress?

For Step C, we suggest that each MDE evaluate the "match" between the interventions and initiatives they initiate and the barriers to reaching the vision they identified. This involves assessing whether an MDE seems to contribute to eliminating or diminishing the power of bottlenecks in a sociotechnical system. Ideally, this should focus on the most crucial bottlenecks. This step is a necessary precursor to evaluating whether the MDE spurs the emergence of new, needed functions in the sociotechnical system (Steps D and E).

This type of evaluation must be (a) conducted on an ongoing basis and (b) handled or coordinated by the MDEs because identifying barriers to their goals and launching initiatives to address such barriers are, in fact, their *raisons d'être.*

3

Ex Post Evaluations

Summative and attribution-oriented evaluation steps aim to *assess outcomes* and *the degree to which an MDE reached its goals*. This implies a "working backwards" approach, where observable changes are reviewed, followed by an analysis of *whether they can be linked causally to an MDE intervention/activity*. Here we suggest evaluating whether and how the targeted sociotechnical system(s) demonstrates improved performance (formative/summative evaluation Step D) and whether the system exhibits any structural changes that facilitate reaching the vision (summative evaluation Steps E and F).

Ideally, such evaluations should be conducted ex post the current MDE initiatives because systematic change often takes years to accrue. As such, these types of evaluations instead should be conducted by the policy actor or external evaluators working on their behalf, not the MDEs.

Considerations. The MDEs in focus are similar in having received funding (relatively small relative to other MDI initiatives globally) from Vinnova and being instructed to implement eight WP. However, the MDEs also were given agency in determining what challenges to focus on, how to design their vision, and how to implement the WP. We show that the MDEs exhibit great differences in these regards, which has logical consequences for designing an evaluation approach that is useful for all five.

Thus, we caution against assessing the MDEs uniformly on *all* WPs or mere "vision attainment." Instead, we argue that an evaluation of the MDEs also needs to *assess the WPs;* that is, it should evaluate the policy design of the overall MDE program.

Finally, a prerequisite for addressing multiple and diverse stakeholders' needs is to gain their trust. Stakeholders who are more engaged with and understand the evaluation's wider purposes are less inclined to feel "threatened" and will impart more useful and meaningful information. Thus, we argue for actively involving the MDEs in the evaluation steps (especially Step C, which is a tool to actively help them prioritize, document, and evaluate the actions and initiatives they take) and, whenever needed, organize external expert panels to assist them in this work.

Table of Contents

List of Tables and Figures

Introduction

What Is Mission-Driven Innovation Policy?

Since Arrow's (1962) pathbreaking work in the 1960s, the core aim of public innovation policy has been geared towards addressing alleged "market failures," notably by investing in public research universities and stimulating private firms' investment in research and development (R&D). From the 1990s onwards, a second generation of innovation policy also strived to address institutional constraints in countries and regions by creating *national innovation systems* frameworks (Edquist, 2010; Nelson, 1993). A third stream of thought emerging in the 21st century is *transformative innovation policy* (TIP; Schot & Steinmueller, 2018; Weber & Rohracher, 2012). Transformative innovation policy provides a broad perspective or rationale on innovation policy (linking innovation policy to addressing societal challenges) rather than outlining concrete policy strategies. One prominent concept within this stream of thought is mission-driven innovation (MDI) policy,[1] in which the focus is on formulating long-term missions of societal transformation and progress and gearing public and private actors' innovation efforts jointly towards accomplishing these missions (Mazzucato, 2018).

Mission-driven innovation policies are founded on governmental attempts to address fundamental but complex societal challenges (sometimes, but not always, depicted as "wicked problems").[2] Corresponding missions, such as affordable health care or a carbon-neutral economy, legitimizing government intervention, aim at influencing the directionality of innovation towards addressing the perceived societal problem; that is, working towards the mission (Wanzenböck et al., 2020; Wesseling & Edquist, 2018).

By positing policymakers as initiators and mission coordinators, MDI policies share important features with existing frameworks such as *demand-based innovation policy* (Thirtle & Ruttan, 1987), *policy-induced innovation* (Lindman & Söderholm, 2016), *challenge-led innovation* (Kuhlmann & Rip, 2018; Raven & Walrave, 2020), and *dedicated innovation systems*, which focus on social inclusion (Pyka, 2017). Mission-driven innovation policy is hence more systemic and holistic than, for instance, the Manhattan or Apollo programs, which were directed towards specific technical goals (Organisation for Economic Cooperation and Development [OECD], 2021). The OECD (2021, p. 15) defines MDI policy as:

> a co-ordinated package of policy and regulatory measures tailored specifically to mobilise science, technology and innovation in order to address well-defined objectives related to a societal challenge, in a defined timeframe. These measures possibly span different stages of the innovation cycle from research to demonstration and market deployment, mix supply-push and demand-pull instruments, and cut across various policy fields, sectors and disciplines.

---

[1] These policies are sometimes also referred to as *mission-oriented* or *mission-led*. We use mission-driven here because this is the term Vinnova uses.

[2] "Wicked problems" are societal problems characterized by conflict, complexity, and uncertainty in policy-making where vision-driven innovation has been emphasized as a possible approach (Wanzenböck et al., 2020).

The OECD (2021) outlines two sources of variation in MDEs: (1) the wickedness and (2) the scope of the missions they seek to address. The expert group on the Economic and Societal Impact of Research (ESIR, 2017) distinguished two broad categories of challenges:

- Challenges that are potentially solvable and can thus be relatively easily reduced to discrete or verifiable goals (e.g., technological challenges or the development of a new vaccine).
- Challenges where the problems escape simple definition, and the solutions are unknown. Wider societal problems, such as sustainability or migration, fall into this category.[3]

These two categories of challenges echo Nelson's (1977) famous dichotomy between "the moon and the ghetto." They highlight that under the umbrella of "mission policy,"' two inherently different types of missions are included: missions aiming for faster scientific and technological advancement, and missions targeting societal challenges with implications for transformational changes (Kuittinen et al., 2018; OECD, 2021, p. 34).

The scope of the mission challenge puts specific challenges on policy design and policy trade-offs. As the OECD (2021) highlights:

> Any challenge can be represented as a problem tree and governments can set their own objectives at any position in the tree. Positioning the challenge at a lower position in the problem tree narrows the scope of potential options for solving the broader challenge but makes the policy more concrete and feasible. . . . When selecting the challenge to be addressed, governments thus face a trade-off: The challenge must be broad enough to engage a broad set of actors across policy fields and sectors without "picking winners" (i.e., be overly prescriptive in terms of potential solutions), but sufficiently concrete and well-defined so that it provides strong orientation and is "actionable." (p. 35)

This challenge and potential policy trade-off are also relevant for the current report, which deals with how to evaluate five ecosystem-based mission projects, so-called mission-driven environments (MDEs). According to the OECD (2021), such projects are particularly subject to "mission capture" because, to be effective at developing consensual strategic agendas, they must rely on established communities. Often, these communities relate to incumbents in key sectors that tend to avoid transformational agendas involving reshuffling established economic positions. Among others, Mazzucato (2021) discussed the risk of "mission capture."

Overall, MDI policy implies a holistic approach in which the policymakers seek to encourage and stimulate a broad set of actors in society to address a societal problem. However, what exact actions are to be performed is not specified ex ante. Hence, MDI policy implies a shift from what is sometimes referred to as the "complexity paradox" of public policy. In this paradox, the more complex the issues are, the more compartmentalized policy-making becomes, fragmenting problems and solutions into different and competing initiatives and departments (Molas-Gallart et al., 2021, p. 75). Instead, MDI policy implies encouraging and

---

[3] See also Wanzenböck et al.'s (2020) discussion of how "problem wickedness" varies with the problem's level of contestation, complexity, and uncertainty.

supporting innovation processes that involve multiple private and public stakeholders in society working together towards a mission in a way that is expected to be experimental, risk-taking, failure-tolerant, and adaptive to new information and circumstances.

In theory, MDI policy should support the identification of major hurdles to moving towards the vision and negotiations about ways to reduce or eliminate such hurdles. Here again, suboptimal solutions benefiting only certain parties should be avoided. Instead, actors should be encouraged to see not only their part but also the whole and thus strive towards solutions that contribute positively to both. Expectedly, this will be challenging because addressing societal challenges often requires actors to build new capabilities and reconfigure external relationships to embrace collaboration and spillovers across sectors (European Commission, 2018; Mazzucato, 2018, p. 60).

Our Focus: Vinnova's Mission-Driven Innovation Environments within Health

This report focuses on a particular instance of MDI policy executed by Sweden's innovation agency, Vinnova: the funding of five MDEs in 2019. Vinnova's initiative can be categorized as an ecosystem-based mission program, which is slightly different from other types of mission-led policies, such as mission-oriented strategic frameworks or thematic mission-oriented programs (OECD, 2021). The *ecosystem-based mission program* can be described as focusing on

> directionality and legitimacy by delegating responsibilities related to strategic orientation to relevant community (or ecosystems) of stakeholders in priority or emerging areas. One of the main added values of this type of [MDE] . . . is to engage wider participation and significant investment from a variety of partners in initiatives that build on the strategic agenda they have collectively designed. (p. 31)

The MDI program in question is called Vision-Driven Health.[4] It was initiated in 2019 to support the establishment of interorganizational and cross-disciplinary coalitions (MDEs) that work towards a common vision and long-term systemic transformation within the Swedish health care and life science sector.[5] A key feature of this initiative was its promotion of self-organization: Applicants propose the objective they will pursue in addressing a societal challenge, the means to realize this objective, and the governance of the MDE. Prior to the call, Vinnova performed an extensive consultation with 50 to 80 organizations (firms, government

---

[4] In the MDI literature, *mission-driven* is the theoretical logic for the constellation of partners and activities to achieve transformation and social and economic development, whereas "vision" is seen as a working principle (WP). In the Swedish program translation, "mission" has been translated to "vision."

[5] This vision-driven approach can be seen as a successor to the so-called strategic innovation programs (SIPs) that have been a main vehicle for innovation funding by the Swedish government in the past decade. A recent meta-evaluation of the SIPs concludes, "If the SIP instrument is to be used for more radical, system-changing purposes and to meet societal challenges, it needs to be modified to involve and prioritize the needs of societal stakeholders to a greater extent. In future programs, participants need to agree on clearer and more specific guiding visions and to focus on making a small number of big changes, rather than trying to support many interests from industry and the public sector through R&I agendas that aim to be very inclusive and therefore tend to become fragmented" (Åström et al., 2021, p. 8). The new MDI policy could thus be seen in light of Sweden moving towards a "third way" in innovation policy—from the traditional focus on improving competitiveness towards system-changing purposes seeking to meet societal challenges (Tillväxtanalys, 2020).

agencies, civil society, and patient organizations) over 1.5 years. Figure 1 was influential in Vinnova's design of this policy.

Figure 1
Design Process for National Missions
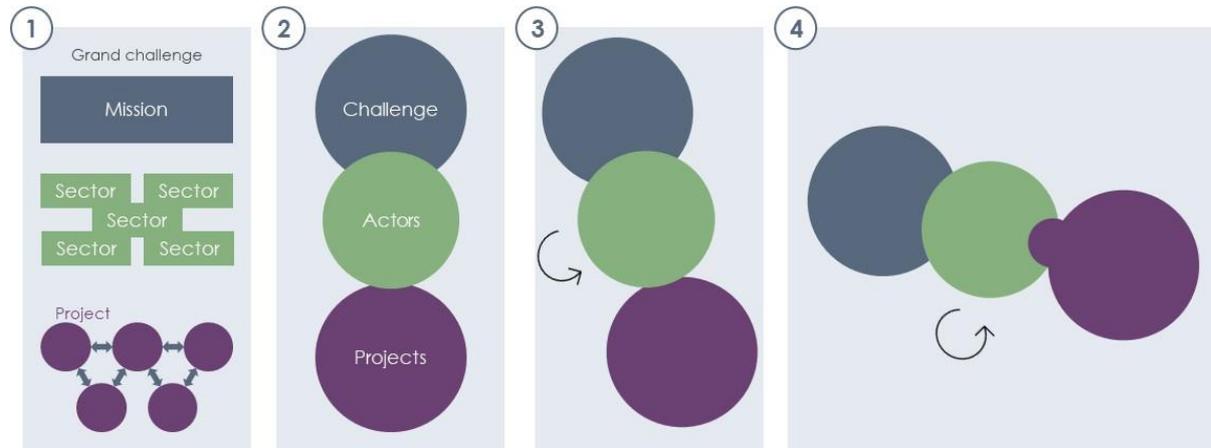
Source: Vinnova (2020).

As Vinnova (2020) described, the process prioritizes participation and views policy explication as a subsequent step. The idea is to gradually develop a more precise understanding of what is possible and what capabilities exist, create broad engagement, and mobilize diverse forms of knowledge to achieve tangible outcomes.

> This approach to missions looks to co-design the agenda, build a network for change, produce action on the ground, and then create a set of stories based on insights from these ongoing activities. Having built up that body of work, then we can approach national politics in a more meaningful discussion about the many ways that a national mission could add value to this work. (Vinnova, 2020)

Five coalitions have been selected so far to build their own MDEs. Each MDE received SEK 5 million (approximately 500,000 EUR) of yearly funding for an initial 2-year-period. A conditional second phase offers an additional 3 years with the same funding level. Table 1 summarizes the basic MDE properties.

Table 1

Basic Properties of the Five Vinnova-Funded Mission-Driven Environments within Health

Topic area	Advanced therapeutics	Cancer prevention	Antibiotics usage	Older adult nutrition	Health informatics
Vision	Sweden is a world leader by 2030 in development and availability of advanced therapies that can cure patients of serious illness	Vision Zero Cancer: turning cancer from a deadly into a curable or chronic disease	Antibiotic-Smart Sweden	By 2030, no older adult person should be at risk of malnourishment	Sweden offers information-driven health care supported by artificial intelligence applications
Coordinator	Independent research institute	University	Government agency + independent research institute	Nongovernmental agency	National technology center
Core-team partners	7 organizations	6 organizations	12 organizations	12 organizations	4 organizations

This report does not focus on how to evaluate Vinnova's funding strategy or the MDI policy logic per se.[6] Our focus is on *how the five MDEs Vinnova funded could be evaluated*. Vinnova commissioned the report in 2021. At the time of writing the report, the MDEs had been operating for 2 years. In late 2021, all MDEs were granted funding for an additional 3 years of operations.

Aim and Methods

This report aims to provide a framework for evaluating five MDEs funded by Vinnova. Vinnova asked us to consider particularly the role of the eight work principles (WPs) Vinnova recommended the MDEs implement. These WPs were based on what the agency considered "best practice" for MDEs (principles outlined in Part 2).

This report results from the first (of two possible) steps in evaluating the five MDEs. The first step is about developing principles for evaluating MDEs. A second step would involve evaluating the MDEs based on the principles outlined in this report.

Methods and Report Outline

This report is based on the following activities and materials:

1) A review of relevant literature documenting *"best practice" WPs for MDEs*. Understanding the MDEs' WP is essential to designing evaluation approaches for them. The literature review is summarized in Part 1.
2) To provide meaningful evaluation principles suited to the MDEs in question, we collected empirical data about how the five MDEs operationalized the eight WPs Vinnova recommended. The result of this empirical study is presented in Part 2. We also sent a survey to members of the participating MDEs, asking what a meaningful evaluation could imply from their perspectives. The (limited) survey results are incorporated in Part 2.
3) A review of recent literature on *how MDEs can be evaluated*. The result of the literature review is presented in Part 3.
4) Based on the joint analysis of these sources of material, we developed a framework for evaluating the five MDEs. The result of this phase is presented in Part 4.

We consulted a group of senior advisors on three occasions. The senior advisors had expertise in areas such as health care ethics (Charlotte Hall), evaluating MDI policy (Matthijs Janssen), innovation policy design, evaluation, and impact assessments (Jelena Angelis), leadership in health care (Barbro Fridén), and health care microsystems, patient perspectives, and patient data (Andreas Hager). The senior advisors were selected to provide critical input based on their insights into the inner workings of the Swedish health care system and broader issues on policy evaluation from different perspectives. A PhD student involved in one MDE (John-Erik Bergqvist) also commented on the report.

---

[6] For the latter type of comparative evaluations of MDE policies vs. other innovation policies, see, for example, Brown (2021), Janssen, Torrens et al. (2021), and Mateos-Garcia (2019).

Literature Review

Several research streams contributed theoretically infused prescriptive models outlining *what mission-led innovation should involve* and *how it should be organized.* We reviewed the literature to identify how MDE/MDE-like environments, programs, or projects *could be* and *have been* organized and evaluated. Given our project's limited scope, our review was based on a snowballing procedure. First, we performed a selective search based on the criterion of empirical papers citing any groundbreaking paper by Nelson (1977, 2011) or Mazzucato (2018) on "innovation missions" or conceptual papers including the word "evaluation" in the abstract or keywords. We also conducted a Boolean string search with various combinations of the terms "Mission*", "Evaluat*", and "Innovat*" in the SCOPUS, ABI-INFORM, and Google Scholar databases. This yielded several hundred papers, of which only a few mentioned empirical data or evaluations of MDIs or MDEs in their abstracts.

Second, we asked the project's advisory board to suggest recent policy and academic material on how MDE/MDE-like environments or policies *could be* and *have been* evaluated. They provided a few dozen additional examples. Within this corpus of papers and reports, we skimmed all abstracts and selected only papers that included empirical data or explicitly mentioned MDE evaluations, leading to a sample of 57 papers. Upon closer reading, most of these were conceptual/programmatic papers and reports outlining how MDE policies could/should be *designed*. They did not discuss how they are or could be *evaluated*. Although there was agreement across these papers and reports that MDEs require new forms of evaluation beyond cost-benefit analyses, the current literature detailed very few evaluation papers or reports. Janssen (2019a) and colleagues (e.g., 2021) constituted notable exceptions, which is why we drew primarily on their work, including white papers/reports. Hence, our literature review is selective and should not be read as an exhaustive systematic review of the field of MDEs/MDE-like initiatives or MDE policy.

Empirical Study Methods

The empirical material is based mainly on interviews with key people ("coordinators") in the MDEs. In autumn 2021, semistructured interviews (Kvale et al., 2014) of approximately 1.5 hours each were conducted with seven coordinators (one of the five MDEs had two project coordinators; in one MDE where the project coordinator was newly appointed, her precursor also participated). Four women and three men were interviewed. The interviews consisted of open-ended questions about thematic areas related to the studied WPs (vision, core team, forms of collaboration, innovation types, etc.; cf. Table 2). The interview purpose was to determine if and how the MDEs operationalized and implemented the principles relative to de facto work done in each WP, rather than determine whether they *intended* (or articulated that they "would soon") to work according to the respective principle. In other words, the interview focus was *realized* practice rather than plans and ambitions.

To complement the interviews, we sent a web-based survey (Appendix A) to all core-team members of the five MDEs. E-mail addresses were collected from coordinators. In total, 40 respondents received the survey, but only nine responded. Survey questions were mainly open-ended, asking for the informants' views on central outcome measures and their evaluability. We also asked for input on how a future evaluation can be *useful* for the MDE itself. Given the rather weak response rate of 22.5%, survey results are not presented separately for validity and integrity reasons. Instead, the material is fed into the qualitative descriptions. To obtain a fuller

picture, we supplemented the interviews and survey with document studies of application texts, interim reports, and, when applicable, other MDE outputs (e.g., reports, seminars, and communication materials). In agreement with the respondents, all observations are presented in a de-identified way.

Analysis of Empirical Data

As Vinnova requested, our analysis focused on how the MDEs implemented the WPs outlined in the Vinnova grant. However, we simultaneously were open to themes suggesting additional, informal WPs (beyond those Vinnova recommended). In the first analysis round, we started by breaking down and "deductively" (Braun & Clarke, 2006) applying the eight WPs—scanning through interview transcripts to identify if and how the MDEs applied the WP attributes. We particularly sought to identify variability in operationalizations of each WP and how these could be linked to the MDEs' different approaches to addressing their respective visions. We also inductively reviewed the material to identify additional ways of working that, according to the respondents, were significant for the MDEs' work and outcomes. These attributes are presented in Part 2 as "additional work practices." In a third round, we conducted a summarizing analysis of the entire material to identify "archetypal change mechanisms" the MDEs apply to influence society in the direction of their vision (see further Table 3). As we explain in the last section of this empirical chapter, early visualization of these different logics can be seen as a tool for identifying what aspects a future evaluation needs to include and what scope such an evaluation needs to have.

Delimitations

This pre-study primarily draws on oral and written representations the MDEs produced. We did not take part in Vinnova's evaluation of the MDEs or examine other actors' perceptions of them.

Selectiveness and Interpretation

The pre-study was conducted during the same period as Vinnova's midterm evaluation of the MDEs, which includes oral interviews and written reporting and planning for a possible next phase. In our interviews, we clearly emphasized that they did not form part of Vinnova's evaluation and that, in that phase, we did not aim to assess the individual MDEs' results normatively. Nevertheless, the evaluation going on at the same time probably, to some extent, colored respondents' presentation of the work within their MDEs.

This pre-study was carried out by researchers at a higher education institution that is also the coordinating party for one of the MDEs. The interview preconditions thus differed between the cases studied. We believe this circumstance had no decisive significance on the pre-study results. However, the potential impact of vested interests must be considered in view of a future full-scale evaluation.

Part 1: Principles of Mission-Driven Innovation

Here we summarize key findings from the literature reviews, highlighting how MDI initiatives should be organized. Appendix B summarizes a selected set of core studies outlining *what MDEs should involve* and *how MDEs should be organized.*

Core Principles for Organizing MDEs

We focus the review on the most influential works in the literature, such as Mazzucato (2018, 2021), as well as derivate work by Janssen, Torrens et al. (2021) and Morgan and Marques (2019) and related work by Sabel and Zeitlin (2012) on "experimentalist governance" and how such can be evaluated (Hellquist & Birksjö, 2021; Hertting & Vedung, 2012). We identify four major principles highlighted as central in the MDIs/MDE literature: *vision and leadership, organization, stakeholders and interaction, and outcomes.*[7]

Vision and Leadership

The *vision* of an MDE should be *clear and understandable* (Mazzucato, 2018, 2021). Visions can be operationalized in broad framework goals, and metrics provisionally established by central and local units. The MDE *leadership* should be *visible* to all actors involved and *responsible* in terms of providing direction for joint work towards the vision, considering equity and inclusion among those involved in the mission work (Mazzucato, 2021). The leadership also needs to encourage *democratic participation*, in which relevant societal groups are represented in the mission formulation. The process of selecting missions must be operationalized into *practicable mechanisms* of how to be open and inclusive (Roth, 2012).[8]

Organization

Organizing of MDEs should *incorporate risk-taking and innovation* among the actors (Mazzucato, 2018, 2021). This involves pilot testing new initiatives, continuous monitoring and evaluating "what seems to work" and what does not, and connecting risks and rewards across actors to avoid shortsightedness and misaligned incentives. The MDEs need to *avoid the compartmentalization* common in public bodies and large private actors by enabling actors to "see the whole" and their own parts, necessitating clear communication between functions and actors. Over time, these actors should seek to *build new capabilities* to integrate, develop, and reconfigure the competencies needed to realize the vision (Mazzucato, 2021). This also involves *information retrieval* in the form of mechanisms for identifying system-specific problems and solutions (Janssen, Torrens et al., 2019).

Stakeholders

Stakeholders and interaction within MDEs should encourage *collaboration across sectors* and involve central actors from domains in society relevant to the mission (private, public, and

---

[7] To some extent, these principles overlap with the more generic form of TIP (see Grillitsch et al., 2019, for an evaluation of the SIPs from a transformative perspective).

[8] Our expert panel noted that there might be tensions between these elements. A clear and selective vision can be hard to achieve when also respecting principles of openness, inclusiveness, and democratic participation.

nonprofit actors). Especially important is that collaboration between public and private actors should be in the form of "dynamic partnerships" rather than private-sector actors used as consultants or on fixed procurement contracts by public actors. The MDEs also should *create competition* rather than collusion. Importantly, however, the risks and rewards of value creation are shared so that competition intensifies experimentation among all types of actors where unexpected benefits are embraced. *Resource mobilization* is necessary for such emerging development paths to function and set new directionality in the developmental paths of technologies and broader sociotechnical systems. Public agencies can co-fund or initiate, but private actors must have incentives to invest and realign investments (Mazzucato, 2021).

Goals, metrics, and decision-making procedures need to be revised to track progress and ensure the vision is being realized. This can occur by *widening the circle of actors*, meaning there cannot be solely one or a few actors governing what is to be accomplished and measured (Sabel & Zeitlin, 2012). Actors should be involved as legitimate participants in decisions on based on (a) being directly affected by a decision (the production/consumption of an innovation), (b) having a formal mandate to represent others (being elected), (c) having expert knowledge, and (d) having a legitimate right to represent others (Hellquist & Birksjö, 2021; Hertting & Vedung, 2012).

Outcomes

Because MDEs aim for large-scale and hard-to-reach visions, long-term financial horizons must be in place (Mazzucato, 2021), with budgeting focused on the outcomes to be achieved and the *actors involved jointly accountable* (Hertting & Vedung, 2012). This is best achieved by a governance structure and culture of transparency. Outcomes also need to be inclusive and aim for structural change, not prioritizing profits or benefits solely among a few or a specific type of actor. To create legitimacy for necessary change, MDEs may need to identify bottlenecks and suggest remedies, such as tax regimes or putting new technology on the political agenda (Schot & Steinmueller, 2018).

Does Vinnova Adhere to the Principles Outlined in the Literature?

Table 2 lists Vinnova's outlined WPs for projects applying for funding within the health MDE framework. The first row lists the WPs as formulated in the call for applications (in 2019); the second row shows the slightly revised WPs (in 2021) as formulated in the evaluation instructions for the overall MDE program. The process of updating and changing plans and WPs during a mission is consistent with the overall MDE philosophy of needing iterations and continuous refinement (Mazzucato, 2018). To compare with the literature review findings (summarized earlier), we include a third row comparing the WPs with the literature's core principles for MDE organizations.

Table 2 *Work Principles for Vinnova's Mission-Driven Innovation Environments*

Work principle	Grant announcement (2019)	Revised principle (2021)	Compared to MDI literature
Vision	A long-term measurable, simple vision with associated goals that are engaging	WP1: A bold and inspiring vision that contributes to gathering forces and clarifies a common direction	Important to articulate a concrete and measurable vision that can be met
Leadership	Headed by a gender-equal and committed core team	WP2: Coordination by a dedicated core team	Leadership should be visible, provide direction towards the vision, and emphasize openness and inclusiveness.
Organization	Attract actors cross-sectorially and interdisciplinarily; explore new, dynamic ways of organizing	WP3: Establish a flexible form of collaboration and multidisciplinarily mobilize actors	Important to engage a wide set of stakeholders with complementary knowledge and work practices that allow adjusting activities to emergent results
Stakeholders and interaction	Relate to and attract ongoing initiatives + drive innovative projects	WP4: Initiate new and connect to ongoing innovation efforts	Private, public, and nonprofit actors engaged
	A plan for national anchoring and international connections	WP5: Communication work with the goal of becoming nationally renowned and internationally known	Encourage experimentation and competition (not collusion) where risks and rewards are shared.
	Capacity to continuously conduct external intelligence assessment[a]	WP8: Receive support from Vinnova with innovation management, external intelligence assessment, and communication	Goals, metrics, and decision-making revised by widening the circle of actors
Outcomes	A plan for anchoring with relevant decision-makers and financiers	WP6: Create conditions for long-term sustainability in the environment after support ceases	Long-term financial horizons with budgeting focused on outcomes (not costs) and actors involved jointly accountable.
		WP7: Work towards equality within the MDE and in the system that the MDE tries to influence	Governance structure and culture of transparency
			Outcomes inclusive and aim for structural change

Note. [a]*External intelligence analysis* or *business intelligence* (Swe: *omvärldsbevakning*) is defined generally as systematic efforts to identify patterns in the relevant environment, understand how these create threats or opportunities for the project or organization, and ensure this knowledge is considered in decision-making (Hamrefors, 1995).

Overall, Vinnova's principles seem to align with the principles in the literature: The main dimensions "recommended" or considered "important" for MDE/MDE-like interventions are, to some extent, operationalized in Vinnova's eight WPs. However, the terms used are abstract and allow flexibility in interpretation. Hence, some nuances emphasized in the literature run the risk of being backgrounded (e.g., the MDEs' competitive aspects and the need to include a wider circle of affected actors in defining and redefining an MDEs' goals, WPs, and evaluation metrics).[9] On the other hand, abstract guiding principles provide interpretive flexibility, which is what Vinnova intended.

As noted in Table 2, Vinnova reformulated some WPs over time. For instance, the requirement to "have a plan" was replaced by requirements about what the WP should achieve. In terms of *stakeholders and interaction*, each MDE was originally required to present "a plan for national anchoring and international connections." This WP was reformulated to "active and long-term communication and anchoring work" with the objective to "become nationally recognized and internationally known." In terms of *outcomes* envisioned, the original WP of "a plan for anchoring with relevant decision-makers and financiers" was reformulated to "creating conditions for long-term sustainability in the environment after support ceases." These reformulated WPs can be understood from the perspectives of (a) MDEs needing to formulate a plan upon application and commencement of how they plan, and (b) plans needing to be reformulated to more concrete goals or envisioned outcomes as the MDEs progress in their work.

Other WPs also have been reformulated. For example, the original WPs noted that *visions* should be articulated as a "long-term measurable, simple vision with associated goals that are engaging." In the revised version, for which we suggest evaluation procedures, visions are described as "bold and inspiring vision, which contributes to gathering strength and clarifies a common direction." A core principle highlighted in the MDE literature is that visions must be concrete, bold, and measurable. Thus, relative to the original version, the revised vision principles seem to stress more the boldness of visions but less need to be long-term measurable. In terms of *organzation*, the revised WP emphasizes a "flexible form of collaboration" rather than "new dynamic ways of organizing," as in the original WPs. When it comes to principles for *leadership*, the original WP emphasized a "gender-equal and committed core team," whereas the revision emphasizes a "dedicated core team," with a new long-term WP being that the MDEs "actively work for increased gender equality." We can only speculate whether some feedback from the MDEs or reflective change among Vinnova desk officers underpinned the revised WP.

---

[9] Vinnova also influenced and guided the MDEs by asking them to report annual progress and resource use according to certain templates. These templates served as governing mechanisms, as well, but are beyond the scope of this report.

Part 2: Empirical Observations from the MDEs

This section consists of empirical observations from the five MDEs in health funded by Vinnova in 2019. The material describes how the MDEs interpreted and *applied* the WP Vinnova outlined during their 2-year establishment phase.

Empirical Observations: The MDEs' Application of the Work Principles

We organized our description of the operationalizations of the WPs by first outlining how the MDEs operationalized the central attributes of each WP as Vinnova suggested (key words in the phrasing of the principles, version of 2021, see Table 2). We then comment on the MDEs' additional WPs relative to these attributes.

As described earlier, the WPs' exact wording has varied somewhat between Vinnova's call for applications and assignment to us. Because our interview questions were framed thematically (cf. methods presented in the Introduction), we do not believe that changes in the formulation have any decisive significance for the observations in this feasibility study.

WP 1: A bold and inspiring vision that contributes to gathering forces and clarifies a common direction

(A) Vision attributes suggested by Vinnova:

- *Bold and inspiring*. Several MDEs actively work to include words reflecting "values" in their vision formulations. For example, one MDE uses wording that turns a negatively charged problem (antibiotics usage) into a positive possibility for action (being clever about antibiotics). A few MDEs deliberately choose an unrealistic vision that attracts attention (eradicating cancer or malnutrition), and several use concepts ("vision zero," "smart," "leading") established for innovation and advocacy work in other sectors.
- *Clarifies a common direction*. The MDEs display large variability in the vision's content and specificity. Some have concise visions (a few words) also used as the MDEs' names. Others use longer formulations that more elaborately describe what the MDE wants to achieve and for whom.

(B) Additional attributes observed:

- *Vision adjustment*. Some MDEs altered their vision formulations during the establishment phase, including abbreviations to a more concise vision and extensions to incorporate added perspectives.
- *Vision function*. The vision fulfills various functions for the MDEs. Some use it as a brand name in their external communication, whereas others use it primarily as internal support for strategic work within the core team. In the latter case, the exact wording of the vision appears less important. In one MDE, the vision fulfills a meta-function in that it enables the MDE to connect to other vision- and mission-driven initiatives internationally.

- *Timeline and measurability of vision*. Several MDEs include the year 2030 in their vision formulations. This dating appears to underline the general future orientation of the MDEs' work rather than represent an actual target for when the vision should be reached. Some MDEs state that the attainment of their vision is measurable. Others, however, state that it would be hard to determine whether they reached their vision because the phrasing includes a measure of judgment (e.g., how is being a "leading country" defined?). Our survey results indicate there might be differing views within the same MDE regarding whether the vision is, in fact, attainable.
- *Development towards the vision*. The extent to which the MDEs follow up or plan to follow up on whether there is societal development towards the vision varies. In some cases, the MDEs identify existing indicators as relevant to follow-up. In many other cases, however, such indicators are lacking. One MDE conducted its own baseline study to better assess the status quo relative to the vision.
- *Attribution*. All MDEs saw challenges in assessing their own contribution to societal development towards the vision because of the great complexity of actors and activities in each field.[10]

WP 2: Coordination by a dedicated core team

(A) Core team attributes suggested by Vinnova:

- *Set up a core team*. All MDEs have a core team. However, the working methods within the core teams vary between MDEs. Some gather the entire core team for joint work with high frequency. In other MDEs, work is conducted mainly in smaller projects that run in parallel. There the core-team meetings have a more informing function.
- *Assume a coordinating role*. All MDEs have a project coordinator responsible for gathering the core team. Apart from this, the project coordinator's role differs among the MDEs, for example, regarding how operationally he/she works. Time dedicated to the coordinating function varies from one part-time person to two full-time positions. In one MDE, project coordination is divided between two organizations, perceived as something positive.
- *Ensure dedication*. It is difficult to infer if or the extent to which the MDEs have "dedicated" members. All claim that core-team members are dedicated, although they varied among MDEs, and commitment in terms of working hours could vary greatly among members of the same core team. Some MDEs describe involvement in the core team as a personal issue; others link it more broadly to participating organizations. In many cases, the MDEs' financing of working time and dedicated job positions is described as a necessary condition for each person's or organization's commitment.

---

[10] Part 2 discusses the challenging tasks of attributing actual societal impacts (including economic, scientific, and policy impacts) because societal impact is a type of system-level output. Any change in this output can depend on any myriad combinations of system inputs, processes, or conditions. Often, such impact also takes time to accrue. Part 4 recommends that such tasks not be put on any specific MDE but potentially evaluated post hoc.

(B) Additional attributes observed:

- ***Departing from familiar ties***. Common to the MDEs is that their core teams emerged from pre-established personal or organizational networks around the coordinating party.
- ***Departing from pre-existing projects***. In several cases, these networks received financial support from Vinnova even before the MDE was formally formed. It was common that the vision and collaboration forms developed within these so-called planning projects. Several MDEs also have close links to other state-supported forms of collaboration (centers, testbeds, etc.) within their area of activity.
- ***Enabling new constellations***. In connection with the vision formulation and work on the application, the initial actor constellations often drew attention to "other key players" who were then recruited to the core team. Although the core teams are based largely on established relationships, all MDEs also state that new relationships and collaborations have been fostered. For example, one MDE describes that being part of the same core team enables joint work between a capital-based research institute and a provincial region, a collaboration unlikely to occur otherwise.
- ***Budgeting for broad actor involvement***. The delineation between the core team and the MDEs' broader stakeholder network corresponds in most cases to the funding structure (if there are resources allocated to a project partner in the MDE's budget, it is part of the core team). However, a couple of MDEs state they also have core team members participating on a nonprofit basis.
- ***Achieving control over dedication***. A challenge is that the MDE itself has limited influence over crucial aspects of the core-team partners' contributions. For example, decisions on how much employee time to set aside for MDE activities and how much proprietary work could be shared within the MDE have to be taken within each participating organization. Several MDEs struggle to achieve the needed committed dedication from partners. This is a common issue in publicly funded innovation efforts, especially those that involve actors such as public authorities and private firms for whom such cross-sectorial collaborative projects are not "core business" (Ramböll, 2021).

WP 3: Establish a flexible form of collaboration and multidisciplinarily mobilize actors

(A) Collaboration attributes suggested by Vinnova:

- ***Flexible forms of collaboration***. The collaboration flexibility that the MDEs are supposed to display is interpreted in different ways: being able to accommodate actors with different expectations and opportunities for commitment and concrete efforts within the MDE, allowing actors' workload to vary over time, and allowing for adjustments in the constellation of actors.
- ***Multidisciplinarity of actors***. The MDEs unanimously state that a multidisciplinary core team facilitates the mobilization of different types of actors in the sector. Some MDEs also work actively with external reference or steering groups with a multidisciplinary composition. However, they have different perspectives on which actors outside the MDE need to be mobilized for the vision to become a reality. Many MDEs describe a trade-off between broad (multidisciplinary) outreach and more focused efforts.

(B) Additional attributes observed:

- ***Budgets allowing for flexibility***. The extent to which the constellation of actors is perceived as adjustable during the establishment largely depends on whether the MDEs have unspecified funds in their budgets. Some that have unspecified funds used these to recruit new core-team members. In other cases, independent initiatives outside the core team were financed.
- ***Flexibility in the transition between phases***. Flexibility in incorporating new ideas and subprojects is described as something the MDEs work on continuously. However, most MDEs concentrated on the projects outlined in their original applications during the establishment phase. The transition to the subsequent financing cycle is seen as an opportunity to adjust plans and core-team composition formally.
- ***Flexibility by necessity.*** All MDEs experienced unplanned but inevitable changes in their core team due to staff rotation, changed operating conditions within a partner organization (bankruptcy, acquisitions), and changes in the surrounding environment. For example, a change in the environment that affected all MDEs towards more flexible working methods is the coronavirus pandemic. The impact varies in scope and consequence. One MDE describes that the work became more centralized due to the reduced availability of partners operatively working in health care. In other MDEs, the consequences are largely limited to the transition from physical to digital meetings.
- ***Multidisciplinarity of issues***. One MDE describes that its placement under a "sector-spanning" center enables cross-sectoral collaborations on technology issues that are not unique to health care.

WP 4: Initiate new and connect to ongoing innovation efforts

(A) Innovation attributes suggested by Vinnova:

- ***Initiation of new innovation efforts***. The MDEs cover a wide spectrum balanced between running their own innovative projects and relating to other ongoing innovation initiatives. During the establishment phase, a couple of MDEs worked almost exclusively with their own pilot projects, developing and spreading specific new working methods. An idea that occurred in several MDEs is that making lead innovators and pilot projects visible could inspire mimicking innovation in other parts of the surrounding environment.
- ***Connecting to ongoing efforts***. The balance between working broadly and finding a niche relates to how the MDEs frame their respective problem areas, visions, and views on additionality (opportunities to influence societal development towards the vision, in addition to what established actors already do). Seeing a need for a unifying force with a holistic view within the sector, one MDE focuses strongly on capturing and channeling what is occurring in the surrounding environment within the scope of the vision. A few MDEs describe that they want to increase understanding of what is needed, inspire with good examples, and engage in and support change rather than drive innovation. Some MDEs see themselves primarily as contact arenas where different actors with common development interests can come together.

(B) Additional attributes observed:

- ***Addressing barriers to innovation***. Many MDEs emphasize the need to reduce system barriers to innovation and suggest increasing knowledge about system dependencies and focusing on the overall needs of users and patients. Some MDEs also work actively with legislators and authorities on these issues.

WP 5: Communication work with the goal of becoming nationally renowned and internationally known

(A) External communication attributes suggested by Vinnova:

- ***Efforts to become nationally renowned***. The extent to which the MDEs prioritize external communication varies greatly. One MDE employs a full-time communicator. In the other MDEs, communication work is among the project coordinator's tasks. Several MDEs describe that their focus was on internal work during the establishment phase, and they expect the importance of external communication to increase in the future.
- ***Efforts to become internationally known***. Most MDEs see international networking as a next step. Most have sporadic international participation at their events and sporadically participate in events abroad. One MDE that aims specifically at profiling Sweden as a leader in the innovation field has a greater international commitment.

(B) Additional attributes observed:

- ***Varying degree of formality in communication***. How formalized the communication strategy and tools are differ between MDEs. Some have a logo and work strategically with their presentation via their own websites. Others have only brief, summary information on the web. Some MDEs present themselves alongside other projects and activities conducted by the coordinating partner.
- ***Varying channels for communication***. Apart from information on their homepages, most MDEs use open seminars as a channel to present their work in public. In several cases, MDE representatives also participate in seminars and conferences arranged by others. Some MDEs launch stand-alone products, such as reports and manuals. The spread of these is part of the communication around the MDE. However, how prominent the MDE itself is in the material varies.
- ***Varying frequency of communication***. Day-to-day communication about the MDEs' work and activities in social media also varies from nonexistent to very frequent.
- ***Varying scope of communication***. In addition to public communication, most MDEs work more intensively with outreach towards certain target audiences seen as crucial for developing the vision (e.g., municipalities, public administration, potential financiers).
- ***Using personal ties for communication***. Several MDEs state they benefit from existing personal networks at the partner organizations for recognition and trust. In some cases, communication about the MDE is strongly associated with a certain front figure (leader or pioneer). This makes communication clear but also more vulnerable in the long run.

WP 6: Create conditions for long-term sustainability in the MDE after support ceases

(A) Sustainability attributes suggested by Vinnova:

- ***Long-term (self-)sustainability****.* Several MDEs have ideas for developing organizational forms and financing models to support continued operation after Vinnova's support ends. Examples include partner-funded collaboration models and commercialization of specific solutions developed within the MDE. Several MDEs also are investigating opportunities to spin off smaller research or innovation projects that could qualify for their own funding.

(B) Additional attributes observed:

- ***Sustainability of results****.* Some MDEs reflect on whether they should seek long-term sustainability of the MDE itself or longevity of the results achieved. Some aim to contribute to legislative changes or common infrastructure investments that would simplify future innovation. One MDE aims to contribute to behavioral change that eventually would become accepted practice.
- ***Sustainability of operations with more stable financing****.* Several MDEs hope for continued state or regional base funding. Working towards a more stable "Center" establishment in public–private collaboration is a common vision.

WP 7: Work towards equality within the MDE and in the system the MDE tries to influence

(A) Equality attributes suggested by Vinnova:

- ***Equality within the MDE.*** All MDEs express that they strive for equal distribution between men and women in core teams, reference groups, projects, and seminars. However, several MDEs express that they are not completely satisfied with their current distribution. There are distortions in terms of both fewer women and fewer men. The difficulty of achieving similar group compositions in the MDEs reflects the composition of professionals in the sector in which one operates. One challenge is that the MDEs can only partially influence which employees the participating organizations assign. One MDE with female project coordinators describes this as something positive because most of Vinnova's funds go to projects with male leadership.
- ***Equality in the surrounding system****.* Several MDEs draw attention to the fact that the social problem that their vision intends to tackle is unevenly distributed or takes different expressions between women and men. Some state they actively include gender equality aspects in their subprojects. This may, for example, be about ensuring data is collected in a gender-neutral manner. One MDE has a core-team partner with expert competence in gender-equality issues and the specific role to include these issues in pilot projects.

(B) Additional attributes observed:

- ***Patient participation.*** Several MDEs discuss the need to include patient representatives in their day-to-day work and develop specific solutions. However, only one MDE has a systematic process for patient participation in place.
- ***Broader ethical aspects within the MDE****.* A few MDEs mention the trade-off between working through networks based on familiar ties and ensuring projects have broad

benefits for the surrounding system. Ensuring that MDE members participate for the greater good rather than their own visibility is a challenge the MDEs need to handle.

- **Broader ethical aspects in the surrounding system**. Several MDEs state they look at the equality issue from a broader perspective as part of Sweden's ambition of equal care (swe: jämlik vård). In addition to gender, this includes care on equal terms regardless of age, ethnicity, geographical residence, and so forth. Within one MDE, the development of costly innovative treatments that benefit only a few patients is discussed as an ethical issue upon which the MDE needs to take a stand.

WP 8: Receive support from Vinnova with innovation management, external intelligence assessment, and communication

(A) Support attributes suggested by Vinnova:

- **Training in innovation management**. Several MDEs mention participating in an innovation leadership training offered by Vinnova but delivered by an external player.
- **External intelligence assessment support**. The MDEs have had some interaction with each other, mainly through meetings chaired by Vinnova. Three MDEs subsequently agreed to joint follow-up on a specific issue (patient researchers).
- **Communication support**. Some MDEs interacted more closely with Vinnova during the establishment phase. For example, Vinnova representatives were invited as commentators at MDE seminars. One MDE also collaborated with Vinnova on a possible future evaluation.

In summary, the five MDEs seem to have implemented the eight WPs Vinnova outlined at the overall level. However, operationalization of "the vision" varied greatly among the studied innovation MDEs. The MDEs also implemented the remaining WPs with different emphasis and prioritization. We elaborate on these differences next.

Different Approaches to Vision-Driven Work

One theme that emerges as salient in our analysis of the data collected (interviews, observations, and archival material)—and which the preceding description does not fully capture—is that the MDEs seem to *focus* their resources (time, attention, money) differently. That is, although all the MDEs make efforts to align with all of Vinnova's eight WPs in one way or another, the intensity and frequency of activities targeting different WPs differed across the MDEs. This insight is inferred from the interviews and documentation available from the MDEs, which emphasize some activities over others.

We analyzed the different priorities inductively (Braun & Clarke, 2006) as they appeared from our qualitative data. From that analysis, we identify five archetypical core-change mechanisms the MDEs apply to influence the sociotechnical system in focus to the desired direction. Because the change mechanisms include assumed bottlenecks and activities targeting them, they can be understood as relevant parts of abstracted "impact pathways" (Griniece et al., 2020).

Table 3 summarizes the archetypes. This is not a full-fledged archetype analysis (Eisenack et al., 2019), and we did not perform a quantitative analysis measuring the number of activities or participants attending to them. Hence, the archetypes outlined should be viewed as tentative and potential tools for visualizing what aspects a future evaluation must include and what scope

such an evaluation must have. The categories in Table 3 also are not to be interpreted as a description of each MDE. In practice, many MDEs mix different logics and activities, and many plan to adjust their working methods and focus over time.

Table 3

Archetypal Change Mechanisms of the Studied Mission-Driven Environments

Impact-creating activity	"Pilot testing"	"Investment hunting"	"Best practice showcasing"	"Network marketing"	"Social informing"
Core practices	Carries out own small, local pilot trials and analyzes effects and systemic obstacles; Feeds knowledge and analyses to public inquiries and policymaking	Produces proposals for future government initiatives; Branches off and seeks funding for parallel projects; Contact node for private investments in the sector's companies	Focuses on highlighting one innovative regional solution; Tries to spread insights and experiences from development, implementation, and operations	Aims to gather as many actors and projects as possible under its brand; Networking and external communication are main activities.	Seeks to inform various actors at grassroot level about the problem area and its consequences; Develops standards for desirable action
Typical output	Research/case report	Investment proposal, application for funding	Manual	Communication materials	Guidelines, certification
Underlying assumption about innovation system	System-level problems need to be inferred from hands-on implementations	Resources are lacking	Good examples and existing practical knowledge need to be disseminated	Attention and meeting places are lacking	Knowledge about the problem and guidance is lacking in community
Receivers in focus	Public administration and politics	Financiers	Regional decision-makers	All interested players in the sector, nationally and internationally	Individuals, small businesses, municipalities

Part 3: How Can Mission-Driven Environments Be Evaluated?

Given our aim to develop suggestions for how MDEs could be evaluated, we depart from the literature outlining ways to evaluate MDEs, "experimental policy," and TIPs[11] because these by and large overlap. These literatures use and extend previous research on policy science and economic development (Foray et al., 2012), industrial policy (Rodrik, 2008), and technological innovation systems (Edquist & Zabala-Iturriagagoitia, 2012).[12] We draw primarily on Janssen (2019a, 2019b) and colleagues (e.g., 2021) because their work aims to develop principles for evaluating environments similar to the five Vinnova-funded MDEs in focus here. Next, we briefly discuss (a) the challenges tied to evaluating interventions such as MDEs, (b) possible approaches to evaluating such interventions, and (c) how we chose to pursue the task of developing a framework for evaluating the Vinnova-funded MDEs in focus.

Challenges Tied to Evaluating MDEs

Overall, research on evaluating MDE-like innovation initiatives (often studies aiming to evaluate *the policies supporting MDEs or similar interventions*) emphasizes its many and partly overlapping challenges, as outlined next.

Scope gap and time lag

Authors note the inevitable "attribution" challenge faced in attempts to evaluate the impact of policies that are implemented and directly affect a limited set of activities that are distant (several steps "upstream") from the intended final objectives (Smutylo, 2005). The MDE-like initiatives often are characterized by a narrow, local, geographical, and activity scope, typically with the ultimate objective to achieve macro-level, systemic, and far-reaching change. Further, MDE-like initiatives commonly have a narrow temporal scope with intended outcomes expected to materialize in the long-term perspective, which makes evaluation difficult (Molas-Gallart et al., 2021). There is thus a "time lag" challenge in the temporal gap between initiation and support of MDEs and their potential impact. As by Janssen (2019b, p. 82) noted:

> An impact assessment serves to determine economic progress in terms of productivity growth, employment, added value, exports, etcetera. Because such final outcomes can

---

[11] Transformative innovation policies are based on the notion that key societal challenges require profound changes in current sociotechnical systems (Schot & Steinmueller, 2018). They refer to *coordinated attempts to make a leap towards selective economic activities*, characterized by three properties: *selectiveness* (prioritization and pursuit of specific development paths), *multi-instrumentality* (a wide set of policy actions to support the pursuit of development paths), and *process-orientation* (policy actions continuously adapted to current bottlenecks) (Janssen, 2019b).

[12] Mission-driven environments are seen as different from other ground-breaking "projects" or "initiatives" discussed in the evaluation literature. The MDEs address different challenges than those well-known projects (e.g., Manhattan and Apollo) address, which aim to develop a particular technological capability and its achievement at the end of the program (Foray et al., 2012). Policy supporting MDEs further differs from "generic innovation policy" because MDE policies aim to help create the conditions that allow specific technologies to be created and diffused (Janssen, 2019b; Molas-Gallart et al., 2021).

only be measured after a substantial number of years, appraisals on the shorter term typically resort to econometric analyses that might reveal an elevated level of experimentation. Even if policy is just trying to enhance the general amount of entrepreneurial experimentation in a country, there are many pitfalls to consider when determining the total societal benefits. For instance, indicators on R&D investments do not tell how efficiently these are translated into outputs, or how novel these outputs would be. Uncertainties of these kinds prevail throughout the entire framework for policy intervention, stretching from inputs to outputs to outcomes.

Of course, this time lag also applies to impact in areas such as scientific or societal progress.

Multi-instrumentality

The literature also suggests that the kind of innovation MDE typically aim for involves unpredictable interactions between processes of knowledge creation, technology development, innovation, and growth—processes that may move in multiple parallel, iterative, and nonlinear, rather than *single* or *linear*, directions. As an example, new technological development can foster basic research. Whether new knowledge and technologies will lead to innovation and growth depends on parallel innovation trajectories and a set of structural aspects as well as the connections between different stakeholders in the system in focus (Foray et al., 2012; Molas-Gallart et al., 2021).

The intermediate and ultimate "outcomes" of MDEs and the processes and functions through which desired outcomes can be attained thus often are assumed to be multiple and parallel (multi-instrumental). This makes MDEs particularly challenging to evaluate because they constitute a "bundle" of parallel functions and processes endogenously developed over time by the actors involved (Foray et al., 2012). For example, because MDEs often intend to "align" the direction of entrepreneurial experimentation around specific technological pathways (directionality) without steering exactly how these experiments are conducted or by whom, outcome assessments cannot be ex ante defined for a specific group of actors or activities. Extant policy evaluations of single policy interventions aiming for single goals among a pre-identified set of actors thus are not quite applicable in the MDE context (Grilli et al., 2018; Magro & Wilson, 2013; Mazzucato & Robinson, 2018). Evaluations of policy combinations sometimes acknowledge multiple and interacting outcomes but often tend to focus rather narrowly on predefined outcomes attributed to specific policy implementations (Nauwelaers et al., 2009). Such summative "effect-size" evaluations often aim to evaluate the extent to which the policy reached its intended goal (given the counterfactual scenario that no policy was initiated; Czarnitzki & Fier, 2002; Howell, 2017). For MDE initiatives intending to "align" the entrepreneurial experimentation around specific technological pathways, ex post evaluations instead may seek to use "impact pathways" logic to chart possible activity-output–outcome–impact pathways (Aspuru-Guzik & Persson, 2018; Douthwaite et al., 2007).

Partly Unknown/Emergent Goals

Partly unknown/emergent goals are other key features of MDEs, given that multiple policies are enacted simultaneously to enhance directionality and progress towards the mission but without specific ex ante milestones formulated. These partly unknown/emergent goals make traditional policy evaluation tools more difficult to apply (Janssen, 2019b). Several authors ask for more work developing approaches and frameworks applicable to evaluating the emergent

macro-level impact of policy related to complex innovation initiatives such as MDEs (Coenen et al., 2017; Magro & Wilson, 2013; Warwick & Nolan, 2014).

Lack of Relevant Counterfactuals

Because MDEs often are intended to transform entire sociotechnical systems and create new prerequisites for multiple different actors in a field, actors not "directly involved" in the MDE initiative also ideally should be affected. That is, it is hard to identify nonbeneficiaries relative to MDEs. As a result, it is difficult to identify a relevant counterfactual: What would have happened if the MDE had *not* been launched? (Bloom et al., 2019). As Janssen (2019b, p. 82) noted:

> Were we to measure developments taking place in beneficiaries and compare them with those in nonbeneficiaries, we would falsely calculate the policy impact as the observed difference between the two. Thus, rather than subtracting positive effects from one another, we should be adding them.

Directionality

The aim of MDEs is typically to provide "directionality"; that is, innovation that makes progress relative to a certain challenge. Again, this implies that the aim is not to fuel particular technologies. Instead, it is to foster fruitful social and regulatory conditions that facilitate the development of various technologies contributing to the challenge in focus while hindering (reducing incentives and facilitators for) undesirable development directions (Molas-Gallart et al., 2021). Hence, evaluating whether MDEs led to "any" innovation or "overall innovation capacity" would miss the point (Coenen et al., 2017). A more meaningful question is whether the capability of innovating in a way that addresses the identified societal challenge has been achieved at the aggregate level (Janssen, 2019b).

Accounting for Complementarity and Substitution Effects

Activities initiated or triggered by MDE initiatives may interact in multiple ways (complementing, attenuating, and substituting each other or some pre-existing policy or condition). These further complicate evaluating whether the MDEs contribute to an alignment around a certain direction. Such interactions are challenging to trace but important to acknowledge because they suggest that simply measuring the isolated impact of MDEs on specific activities is insufficient (Grilli et al., 2018; Janssen, 2019b; Magro & Wilson, 2013).

In sum, the literature identifies several difficulties related to evaluating MDEs. Next, we discuss possible paths forward given these challenging circumstances.

Possible Approaches to Evaluating MDEs

Formative versus Summative Approaches

Evaluations can be formative. They can aim to generate knowledge about *how* interventions are implemented (or not) and conducted in collaboration with stakeholders for the main purpose of improving the definition and ongoing implementation of the interventions being evaluated. Evaluations also can be summative. They can aim to attribute observed outcomes to intervention effects to establish if and to what extent the intervention reached its desired end goals (Molas-Gallart et al., 2021)

Formative evaluations allow for assessing efficiency: what works more or less for the intended goals. This also can feed into more refined summative evaluations because they shed light on what (intermediate) outcomes to measure. Summative evaluations allow for gauging effectiveness: whether an initiative reached its intended goal or outcomes to a higher or lower extent (Furubo et al., 2002). Formative and summative approaches are not mutually exclusive. It seems meaningful to use formative evaluations at regular intervals during an intervention, in combination with summative evaluations, to enable ongoing learning and relevant adjustment of policy. At a more general level, they support both accountability and reflexivity (e.g., Magro & Wilson, 2013).

Janssen, Wesseling et al. (2021) provide a model designed for evaluating transformative-oriented policies (similar to MDI policies), combining these two general approaches (reproduced in Figure 2). The figure highlights six types of evaluations (evaluation foci) spanning from formative to summative evaluation (A–F). The authors highlight these categories as complementary, all with the potential to affect policy impact.

Figure 2

Different Approaches to Evaluating Mission-Driven Environments
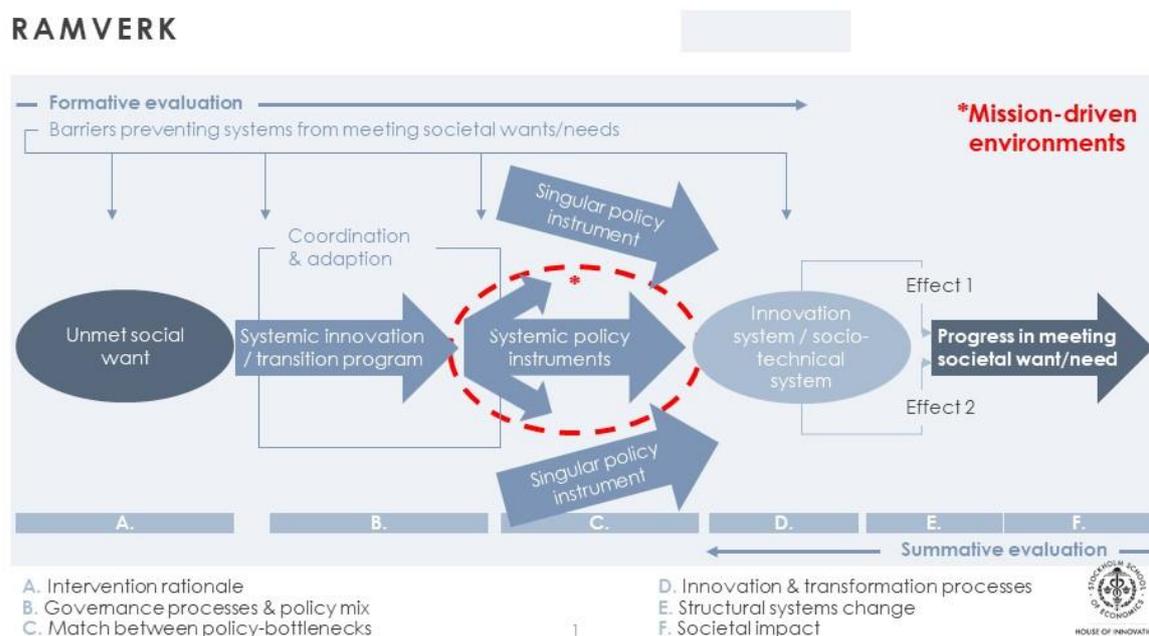
Source: Janssen, Bergerk et al. (2021).

Formative Evaluation Steps

(A) If and to what extent (and from whose perspective, authors note) the *intervention is justified* in terms of the existence of a "failure" of the system, market, or current development direction

(B) If and how (and from whose perspective, authors note) the *governance arrangements are purposeful, consistent, and coherent* (processes and structures; i.e., ways of working and formalized routines/standards/decisions/rules)

(C) If and how (and from whose perspective, authors note) there is a *"match" between the interventions and barriers* identified (weaknesses, bottlenecks, impeding regulations, social norms, etc.)

Formative and Summative[13] Evaluation Step

(D) If and how (and from whose perspective, authors note) the targeted, overarching sociotechnical system/field demonstrates *improved performance*,[14] for instance in terms of capabilities (system functions and interactions such as nowledge sharing), transition processes, and outcomes

Summative Evaluation Steps

(E) If and how (and from whose perspective, authors note) the targeted, overarching sociotechnical system/field exhibits *structural changes*, such as change in types of innovations, new forms of cross-sectorial collaborations, or new networks constellations in the system

(F) If and how (and from whose perspective, authors note) there is *measurable impact* on the societal level in terms of mitigating the failure addressed and reaching the intervention's "vision" or "mission." This type of evaluation is helped by the fact that some interventions, policy initiatives, and interventions like MDEs have measurable end goals that make possible a comparison of policy impact (perhaps not again the null of "no policy"). As previously discussed, attributing actual societal (including economic, scientific, and policy) impacts to system transformations and the MDE in focus is fraught with challenges. This is because societal impact is a type of system-level output, and any change in this output can depend on myriad combinations of system input, processes, or conditions (Levin et al., 2013).

Given these challenges, Janssen, Bergerk et al. (2021) suggest focusing evaluations with a "summative" orientations on intermediate goals such as changes in Steps D and E. Long-term "societal impact" (Step F) can be analyzed but interpreted with caution. In essence, such long-term impacts represent the vision articulated. By definition, MDIs should be guided by a concrete vision that is easy to articulate, understand, and understand when it is fulfilled or not (Mazzucato, 2021).

Contribution versus Attribution Approaches

Janssen (2019a) outlines two evaluation purposes linked to the formative-versus-summative approach distinction. In the context of MDEs, these are oriented towards:

(a) Learning how the MDEs prioritize opportunities and threats and respond by "working forward" (a formative contribution-oriented type of evaluation), and

---

[13] Step D can be defined as *formative* if involving the explorative identification of new capabilities that seem to emerge as the MDE operates, and insights from these are used to development the MDE further. Step D can be defined as *summative* if involving efforts to establish measurable effects between the MDE and identified changes in the sociotechnical system capabilities.

[14] *Performance* here is usually defined as the degree of attainment of program objectives. We further separate output (activities performed, services delivered), outcomes (effects of the outputs on the beneficiaries), and impact (the degree to which the outcomes are attributable to the intervention (e.g., an MDE).

(b) Determining whether MDEs achieve impact (a summative/impact-oriented type of evaluation).

In the "working forward" approach, observable structures and actions are reviewed. Analysts ask *where and how they are likely to create an innovation-evoking system*. In the "working backwards" approach, observable performance changes are reviewed, followed by an analysis of *whether they can be causally linked to the policy/MDE*. These modes of MDE evaluation processes are depicted in Figure 3.

Figure 3

Contribution versus Attribution Approaches to Evaluation
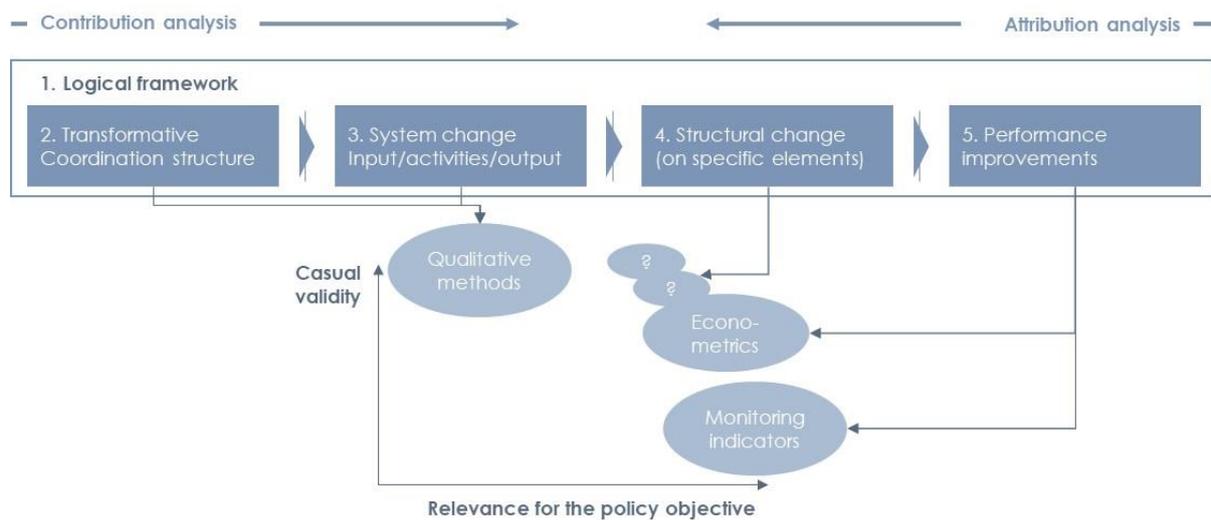
Source: Janssen (2019a).

Next, we outline some suggestions regarding how to implement the contribution approach, followed by the attribution approach.

Contribution Approach

The objective of MDEs is typically to overcome the inertia and system bottlenecks preventing directionality change in the sociotechnical system. Such barriers may differ greatly among MDEs, depending on what societal challenge is addressed. However, they often center around legal issues or a lack of complementary production factors. A contribution-oriented evaluation of MDEs would aim to *assess whether the MDE seems to contribute to eliminating/diminishing the power of the most critical bottlenecks* in the focal system.

Contribution-oriented evaluations also can study whether and how interventions contribute to conditions that stimulate creating and applying new technologies (Janssen, 2019a). Rather than focusing on the number of actual new technologies developed or new innovation actors in the field, this perspective suggests focusing on whether the overarching sociotechnical system seems to adjust and become more prone to stimulate innovation. For instance, it might shift focus from whether the system shifted from or abandoned impeding processes and functions to *developing underlying structures that provide the necessary conditions for technologies and*

34

new actors to emerge and sustain change over time (i.e., key processes and functions; Bergek et al., 2008; Hekkert et al., 2007). Janssen et al. (2019a) suggest assessing various system "functions" proven important in previous literature on technological innovation systems. These can be assessed on three dimensions:

(a) How important is system strengthening or transformation (what were the problems?)

(b) What was done to strengthen it (which inputs; i.e., investments, activities)

(c) What can we say about the realized outputs on each function?

As an example of these dimensions, if the problem assessment suggests that knowledge development and exchange are especially weak, evaluations could focus on *whether the intervention (MDE) invested in addressing these and if those investments seem to reduce bottlenecks in these areas*. The MDE should focus less on functions considered strong at baseline.

Studying what pathways are evolving out of MDEs. Given that the approach can be oriented towards one or multiple domains, a starting point for assessment is determining *which pathways* (Molas-Gallart et al., 2021) or sociotechnical trajectories are evolving from an MDE (Janssen, 2019a). Such pathways or trajectories can be multiple or parallel and occasionally intertwine.

Contribution mode analyses thus aim to create knowledge about whether ways of working and designing arrangements in the intervention (in our case, the MDE working and governance principles and policies supporting them) are *relevant and efficient*, similar to the formative evaluation style (Furubo et al., 2002).

The downside of such mode analyses is that

> they only shed light on the likelihood that performance-inducing conditions emerge; in this case the strengthening or transformation of systems in which innovation and entrepreneurship can flourish. This still provides little conclusive evidence regarding the extent ultimate policy objectives are truly met. (Janssen, 2019a, p. 12)

Attribution-oriented evaluations—a form of summative evaluation—aim to answer the latter type of questions, which concern effectiveness relative to the policy objective.

Attribution Approach/Impact Assessment

Like summative evaluations, the attribution-oriented evaluation approach aims to assess an intervention's outcomes and establish whether and to what degree the intervention reached its end goals (Molas-Gallart et al., 2021). This implies a "working backwards" approach, where observable performance changes (i.e., goal attainment) are reviewed, followed by an analysis of *whether they can be linked causally to the policy/MDE*; Janssen, 2019a). Such analyses might provide a basis for accountability. The literature provided a few approaches and ideas.

Comparative Impact Assessment

One type of impact assessment is assessing *whether some participants in the system in focus* perform better than others. Actors participating in the intervention exhibiting improved performance could indicate improved underlying conditions. However, that also might signal

that improvements were produced only in the direct "treatment" group while the overall system operates as usual (Janssen et al., 2019a). The system may even deteriorate "when in particular incumbents manage to obtain support for clinging on to traditional activities rather than giving space to the emergence of new ones" (p. 22).

System-Level Changes in the Direction of Innovation

A second type of impact assessment examines system-level changes in the *direction of innovation* (i.e., among all organizations, not only those involved directly in the MDE). Such an assessment could seek to gauge more generally *whether organizations are increasing their activity in existing, adjacent, or new markets or sectors in a desired way.* Such an assessment would seek to gauge changes in entrepreneurial experimentation on the system level (Lindholm-Dahlstrand et al., 2019). A system-level analysis also could seek to assess whether organizations are *shifting their R&D, innovation, and entrepreneurship activities towards topics that are prioritized relative to the MDEs' visions.* This type of assessment would answer whether resources are mobilized in the desired and intended direction (i.e., a desirable shift in knowledge development and search guidance). However, an increased firm-level focus on such topics cannot necessarily be attributed to the MDEs. Certain topics could be "hot" and attract attention because of other developments. Evaluators would need to exclude such alternative causation sources triggering system-level changes in the innovation direction. To do so, they would need either to identify firms likely not to have been influenced by the MDEs' environments as a plausible "control group" or to find another control group consisting of, for example, a similar but institutionally and geographically distinct innovation system (Jansen et al., 2019b).

Collaboration and Cross-sectorial Innovations

A third type of impact assessment is examining system-level changes in types of interorganizational innovation. Such an assessment would seek to gauge whether novel R&D and innovation collaboration patterns are emerging. This type of assessment emphasizes desirable changes in knowledge development and exchange across sectors and organizations. Here it is relevant to analyze whether recently formed partnerships structurally differ from those formed before the MDE initiative. Such assessments are common in innovation studies and innovation economics, using network analysis to track the co-development of technologies or products, patents, and R&D projects across organizations or sectors (Björk & Magnusson, 2009; Powell et al., 1996). Figure 4 illustrates such a network map consisting of nodes (circles) and links (collaboration projects between two nodes) among Australian firms (Janssen, 2019a).

Figure 4

Network Map of Joint Research and Development Projects among Australian Firms
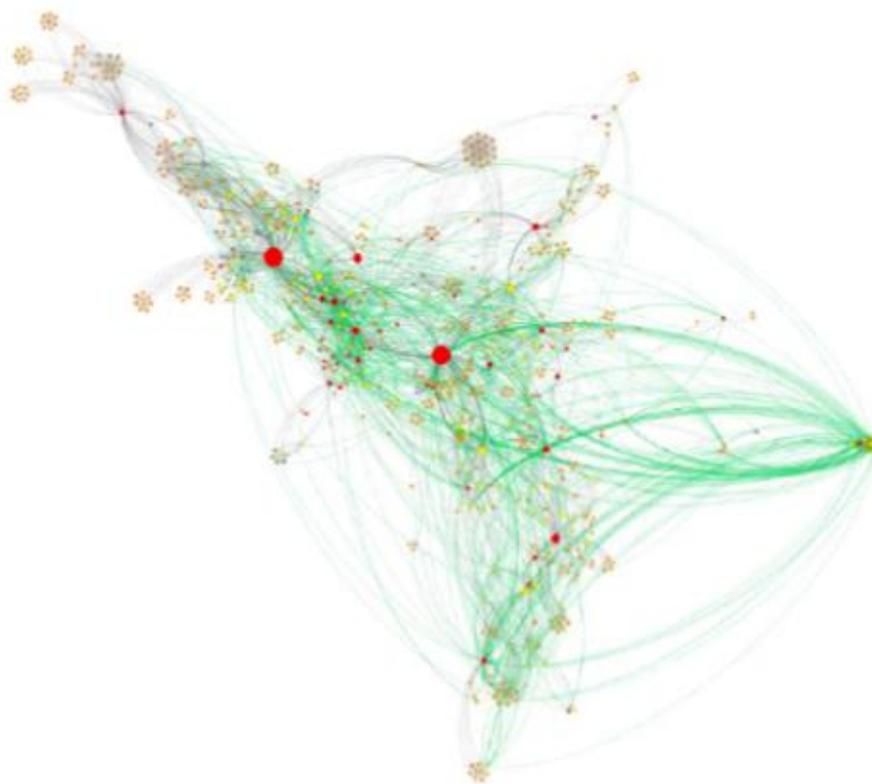
Source: Janssen (2019a).

Macro-level Indicators

Janssen (2019a) suggests an extensive list of indicators that concern actual increases in innovation, growth-related actions, and results thereof. Such macro-level indicators might shed light on economic developments closely linked to the MDE's ultimate objectives. Some monitoring indicators available in program data or aggregated business statistics concern sectorial dynamics, such as changes in turnover, employment, exports, R&D investments, and patenting in particular sectors across newer and older firms. Such data are readily available in Sweden and have been used in earlier innovation program evaluations (Aghion et al., 2021; Engberg et al., 2021; Grillitsch et al., 2019; Humphries, 2021). As noted, progress on such indicators alone does not necessarily tell whether the policy is working unless a comparison with a "no program" counterfactual can be achieved by comparative data or counterfactual simulations (Humphries, 2021). Causal effects are more plausible if the preceding parts of the assessment framework also point in this direction (Furubo et al., 2002). Reversely, no progress on short-term performance indicators might raise questions about the potential of system changes that have been enforced.

In summary, like the summative mode of evaluation, attribution-oriented evaluations typically use meso- and macro-level monitoring indicators that reflect the intended progress towards achieving actual policy objectives. However, such indicators in themselves do not answer

whether this change occurred because of the intervention or policy. In the MDE context, this could mean that investments, exports, and growth in the sociotechnical system in focus might increase but not necessarily through successful MDE actions.

As Janssen (2019a) suggested, using micro-level data and comparing organizations "treated" or "subject" to the influence of MDEs in the field in focus (including ones receiving spillovers) with suitable control groups before and after establishing the intervention (MDE in our case) is an interesting option. Using econometric techniques like propensity-score matching and difference-in-difference models to identify control groups would help shift indicator analysis more towards a valid causal analysis (Figure 3, right-upper corner). However, as already above, finding control groups is difficult; MDEs typically intend to trigger a set of activities that change the entire field (and thus all or, at least, many actors operating in it, including those not participating directly in the intervention). Janssen and colleagues conclude that a feasible way forward could be to triangulate methods and combine a set of approaches.

Ensuring Usefulness of Evaluations

Based on the literature and data collected from the MDEs, we find it relevant for evaluations of these to use a *combined-methods approach that includes elements of both formative and contribution-oriented analysis* on the one hand and *summative and attribution-oriented analyses* on the other. As Janssen (2019a) noted, analyzing process data on an ongoing basis in combination with multiple quantitative methods of summative character (assessing multiple isolated intermediate- or end-goal dimensions, e.g., patents, export, number of new networks or firms) are useful ways to evaluate broad innovation programs like these without fixed goals formulated before program launch.

The usefulness of evaluations here deserves particular attention. Evaluations could contribute to learning at the operational, policy, and especially system levels. They could provide insights to a broad set of stakeholders (ultimately, all society) beyond those involved in or responsible for the initiative.

At the operational level, evaluation could be used as a management feedback tool to improve intervention effectiveness, efficiency, and quality (how to improve design, manage, and implement MDEs).

At the policy level, evaluation could be used to *determine the outcome and impacts of policy interventions and the achievement of objectives* (contributing to the design of future programs and critically testing the assumptions prompting the policy intervention).

At the *system level*, evaluations may guide the design and formulation of (sets of) additional intervention policies and programs, answering questions such as when certain interventions are appropriate, which complementary programs should be used and when, and what appropriate policy mix is needed to achieve the desired effects (Amanatidou et al., 2014).

In addition, the evaluations' *timeliness* is critical. For instance, from the operational managers' perspective, information typically will be most useful at the early stages of design and implementation. On the other hand, an evaluation conducted at a too-early stage of operations may not deliver significant information on the results generated or the extent to which objectives have been achieved. From the policymakers' perspective, an evaluation is useful

when it provides this same information at the policy-cycle stage and can be applied effectively (e.g., when modifications to the program or follow-up program can be applied).

A minimum quality level must be achieved to deliver the evaluation's objectives—stakeholders must have confidence in the results' validity (Miles & Cunningham, 2006). We list a set of challenges tied to evaluating MDEs in a way that ensures causal validity. However, non-causal evaluation approaches also are needed to ensure relevance for the stakeholders involved. Thus, we aim to suggest principles for evaluating MDEs that provide a high degree of *validity* and *relevance* to stakeholders (participants in interventions, policymakers, and actors in the field in focus).

Finally, a prerequisite for addressing the needs of multiple and diverse stakeholders is to gain their trust. Stakeholders who are more engaged with and understand the wider evaluation purposes are less inclined to feel "threatened." They will impart more useful and meaningful information. Trust can be gained by demonstrating the evaluation process's utility to stakeholders (Forss et al., 2002; Stufflebeam & Shinkfield, 2007). For evaluations to become beneficial, they must be conducted regularly. Availability of consistent data is critical here. The MDEs thus need to accept the culture of evaluation and start collecting data routinely (Mazzucato, 2021).

We take these insights into account when developing a framework for evaluating MDEs.

Part 4: Recommendations: Principles for Evaluating MDEs

This section outlines our conclusions and recommendations regarding how the MDEs could be evaluated. These recommendations are geared at making the evaluations useful for

(a) MDEs to facilitate their purposeful, ongoing WP adjustments in a way that contributes to their effectiveness in moving towards their vision, and

(b) Vinnova as a policymaker/funding agency in facilitating their purposeful adjustment of innovation funding policy (particularly, the MDE-related WPs).

The term *evaluation* often is used in a rather general and arbitrary way. Evaluation is distinguished from similar practices, such as auditing or reviewing, through its *judging* feature. Evaluations are not just displays of numbers or opinions but include some assessment of the entity being evaluated relative to a predistinguished norm or goal (Pollitt, 2003; Scriven, 1991).

Part 2 highlighted that the MDEs in focus were similar in that all received funding (relatively small amounts relative to other global MDI initiatives) from Vinnova and were instructed to implement eight WPs. The MDEs also were given agency in determining what challenges to focus on, how to design their visions, and how to implement the WPs. As we showed, the MDEs exhibit great differences in these regards, which has logical consequences for designing an evaluation approach useful for all five MDEs.

A straightforward evaluation approach could be to simply *formatively* evaluate *how* the MDEs are developing governance arrangements, *how* they formulated their visions, and *how* they are pursuing their visions. This could then be tied to a *summative* evaluation of the extent to which they reached their vision (i.e., assessing their "performance"). Such an evaluation could potentially result in conclusions about which MDEs have been relatively successful in reaching their vision and at least some propositions about what types of WP implementations could have contributed to that success. This information could lead to the conclusion that some WPs seem less important than others in certain contexts.

It is vital to highlight two aspects tied to the heterogeneity of the MDEs in focus: the wickedness of the problem they seek to address and the scope of the visions articulated. The five MDEs, to some extent, fall into both types of mission challenges the OECD (2021) described and we discussed in the Introduction. This puts specific challenges on policy design and trade-offs because they cannot be assessed uniformly and may require different WPs to succeed (OECD, 2021, p. 35). Because the MDEs vary in the wickedness of their problems and scope of the challenges they are trying to address, their differential prioritization of WPs may be justified. All MDEs have limited resources; hence, it may be rational to focus on selective types of change activity, such as "bottleneck analysis" (discussion follows). Thus, we caution against assessing the MDEs uniformly on *all* WPs or mere "vision attainment." Rather, we argue that an evaluation of the MDEs also must *assess the WPs themselves*, that is, include an evaluation of the policy design of the overall MDE program.

Drawing on the literature in Parts 1 and 2, we outline a set of *tools for formative and summative evaluation* formulated for *the MDE initiatives and the WPs with which they are set to work* (specifically, in the form of contribution/attribution approaches). Because Vinnova (the policymaker) designed the WPs, our framework also includes an element of policy design evaluation. However, our primary focus is on evaluating the MDEs per se. We end with a discussion of how these tools can be combined fruitfully for a sound and useful evaluation.

Evaluation Design: Evaluation Dimensions and Who Should Be Responsible for What

It is unreasonable to assume that one actor could account for or be tasked with evaluating all aspects relevant to MDE performance (Furubo et al., 2002; Janssen, Wesseling et al., 2021). Based on the different formative and summative evaluation steps for MDE policies outlined in Part 3 (Janssen, Bergerk et al., 2021), we here provide our specific suggestions for how the five MDEs can be evaluated and by whom.

To evaluate specific MDEs, we conclude that the formative Steps B and C in Table 4 are of utmost relevance, particularly from a short-term time perspective, and that the MDEs themselves should be involved in such an evaluation. In Table 4, we detail how an evaluation of these steps could be done and the type of data needed and exemplify useful methods for each evaluation step.

Table 4

Responsibility Levels for Evaluating Mission-Driven Environment (MDE) Policies

Evaluation step	Responsibility level	Analytical level
Formative evaluation steps		
A. If and to what extent the *intervention is justified* in terms of system, market, or current development "failure"	Policy actor responsible for policy mix	Mission-driven innovation (MDI) policy (compared to any other or no policy)
B. If and how governance arrangements are purposeful and coherent	Policy actor responsible for MDE policy design	MDE policy design (e.g., work principles and their instructions and implementation)
C. If and how there is a "*match: between the interventions and identified barriers* (weaknesses, bottlenecks, impeding regulations, social norms, etc.).	MDE / potential external expert panel	MDE and their activities
Formative and summative evaluation step		
D. If and how the targeted sociotechnical system/field demonstrates *improved performance*	Policy actor responsible for MDI policy design/external evaluator	System field-level data. such as register and network data on actors' innovative activities and innovative outputs, collaboration activities. and outcomes; ideally compared to a similar field and/or over time (Powell et al., 1996)
Summative evaluation steps		
E. If and how the targeted sociotechnical system/field exhibits *structural changes*	Policy actor responsible for MDI policy design/external evaluator	System field-level data on changes in types of innovations, new firms, and actors active in the system to gauge changes in the system over time (e.g., NESTA's evaluation of mission innovation in chronic diseases; Mateos-Garcia, 2019)
F. If and how there is *measurable impact* on the societal level in mitigating the failure addressed and reaching the intervention's "vision" or "mission"	Policy actor/government	Not intended for evaluation; very hard to evaluate in a valid way; but *assessment* is relevant as input for discussion

Evaluating the Policy Designs' Governance Arrangements

Step B in Table 4 highlights that the policy actor responsible for a specific policy design (e.g., Vinnova) needs to evaluate the governance arrangements of the specific policy. Such an evaluation should assess the policy design's *processes and structures* and implementation (i.e., ways of working and formalized routines/standards/decisions/rules), focusing on whether the design and its implementation by the MDE are:

Purposeful to the overall goal of the policy
Coherent with other processes and structures the MDEs are set to work under[15]

To evaluate Vinnova's five MDEs in health, focusing this part (Step B) of an evaluation on the WPs is a natural analytical level. Such an evaluation should formulate different questions for evaluating the policy design's purposefulness, consistency, and coherence: *Are the formulated WPs necessary and sufficient for MDEs in health? Are some WPs more important than others to achieving expected process outcomes? How do MDEs develop routines and decision rules to operationalize the WPs, and with what results for their progress?*

The policy actor (Vinnova) sets the WPs and requires each MDE to adhere to these. However, each MDE is responsible for operationalizing and adapting the WPs to their specific innovation field and encountered conditions for reaching the vision they articulated. This means we could expect variability in how each MDE understands, operationalizes, and adheres to WPs—among MDEs and over time. Our exploratory study (Part 2) indicated that, so far, the MDEs implemented the eight WPs with different emphasis and prioritization. Thus, an evaluation of the formative stage B should exploit this variability to answer questions of the type exemplified above. Because in-depth data from each MDE is needed, cooperation with the MDEs is necessary for such an evaluation. It could involve each MDE's self-assessment and external evaluations based on documentation/observations of meetings/decisions made/activities performed. Suitable methods for this first evaluation stage could consist of process-tracing methods, comparative case analysis, or similar in-depth qualitative methods. The frequency and extent/volume of certain behaviors also might be quantified based on systematic analyses of protocols or regular observations. Crucial for these to be useful are (a) the WPs, not the MDEs themselves, are the primary analysis level, and (b) actors involved in the MDE are engaged and contribute data access or self-assessment to evaluate the policy design's purposefulness, consistency, and coherence.

Evaluating the Match between Each MDE's Interventions and Identified Barriers

Step C in Table 4 highlights that each MDE must evaluate the "match" between the interventions and initiatives they initiate and the barriers towards reaching the vision they identified. This type of evaluation needs to be (a) conducted on an ongoing basis and (b) handled or coordinated by the MDEs because identifying barriers towards their goal and launching initiatives to address such barriers are, in fact, their raisons d'être.

---

[15] In general, evaluating MDEs' governance arrangements also should assess the governance arrangements' *consistency* across MDEs and over time. As mentioned, we caution against doing so in the current program because the five MDEs are too heterogeneous in their vision statements and wickedness of the problems they seek to address.

Evaluating the "match" between interventions/initiatives and the barriers to reaching the vision can be done several ways. Following, we focus on "bottleneck analysis," a type of formative evaluation with the flexibility to be adapted to various MDEs and necessitating active involvement (if not, as we recommend, leadership or responsibility) of the MDEs themselves.

The *contribution approach* discussed in Part 3 is designed to assess whether an MDE seems to contribute to the elimination/diminishing the power of bottlenecks in a sociotechnical system (Janssen, 2019a). *Ideally, this should focus on the most critical bottlenecks*. The contribution approach also could be used to assess whether the MDE spurs the emergence of new, needed functions in the sociotechnical system (ideally focusing instead on the *key functions*).

This kind of assessment *requires knowledge of what bottlenecks exist or functions are lacking* in the system and their relative importance. Based on this knowledge (ideally stemming from the MDEs themselves), the assessment is done by mapping *if and how the MDEs are investing in activities that seem to reduce bottlenecks*, given the bottlenecks' identified relative importance. Examples of such bottlenecks in health care innovation could be features of the regulatory environment or procurement system complicating the uptake of new innovations, lacking economic incentives in care production to implement new methods, or cultural barriers within the care professions favoring proven solutions over new ones.

Identifying bottlenecks could be the work of cross-disciplinary, multisectorial expert panels. However, as we argue here, this is at least equally and potentially better handled by the MDEs, given that they establish collaboration with central actors in their own environments. For instance, in the context of Vision Zero Cancer, this could be implemented as illustrated in Table 5. In this example, evaluating the contribution approach would focus on assessing whether:

(a) and how MDE actors invest in activities according to the bottleneck's deemed importance and

(b) activities carried out contribute to eliminating/reducing bottleneck(s).[16]

Such assessments would rely on *data about MDE priorities and investments*/activities, and *data on how they prioritized* between these, given their own or external panels' deemed importance of the bottleneck. The assessment could, in principle, be self-monitored and self-guided by the MDEs themselves but also by external evaluators given access to the data described. The top row of Table 5 shows the steps for such an assessment, with the rows below examples from Vision Zero Cancer.

---

[16] Some activities targeting bottlenecks may be possible to link to the WPs. For instance, if 'limited possibilities to work with new partners across sectors' is identified as a bottleneck, then the WP emphasizing a cross-sectorial core team may be relevant. However, we expect the MDEs to perform and facilitate a range of activities that go beyond complying with the WPs outlined by Vinnova.

Table 5

Example of Bottleneck Assessment as an Evaluation Tool

Identified (by/through MDE) bottlenecks/ lack of conditions that stimulate innovation in desired direction	Importance gauged by MDEs/ external experts	MDE investments/ activities focused on reducing/ removing bottleneck	Assess if and how MDE actors invests in activities according to bottleneck's deemed importance	Assess whether activity contributes to eliminating/ reducing bottleneck
Type of assessment:	**Formative**			**Summative**
System-level lack of trustworthy, easily accessible knowledge of cost-effectiveness of prevention/ early detection System-level lack of trustworthy, easily accessible knowledge of how to change operational routines to enable prevention/ early detection		Knowledge development (e.g., commissioned reports, health economic evaluations, predictions, simulations) Financing PhD studies of early detection in primary care, rehabilitation experiments, and study at one clinic	Data on MDE *priorities, investments, activities* Data on *how* MDE prioritized between investments / activities	*Potential assessment tools:* Updated metrics on bottleneck magnitude, counterfactual scenario analyses, follow-up studies after new interventions/activities launched, case comparison with similar interventions in other fields, surveys, external Delphi/ expert panel assessments
Lack of educational content about prevention in current curriculum for relevant professions Lack of interest among young people specializing in lung cancer prevention		Acknowledged but few concrete actions involving health care educational institutions needed Acknowledged but few concrete actions; general mobilization of interest in topic may contribute	Track efforts to make changes in education curriculum or stimulate others to make such changes	Track changes in health care education curriculum and number of specialists trained
Existing regulations impede data sharing for individualized prevention		Acknowledged as a bottleneck but concrete actions; delegated to another MDE	Track direct or indirect activities that help address the bottleneck (e.g., monitoring involvement of regulations, ethics experts, data- sharing experts)	Track whether regulatory changes are discussed, examined, executed
Distrust among actors across sectors (e.g., unidirectional distrust: health care professions and providers distrust industry), which impedes cross-sectorial alignment		Arrange meeting arenas (workshops/ seminars) where actors can explain/showcase what they do and collaborate to create understanding of other's prerequisites		Follow-up surveys gauging cross-sectorial understanding/trust among workshop participants or at participants' workplaces
Current reimbursement system rewards reactive rather than preventive care		Ordering and financing health economic studies demonstrating financial benefits of preventing rather than treating forms of cancer (row 1)	How prioritized are these activities among MDE activities?	Track potential changes in/between regions in reimbursement for treatment vs. preventive care
Disconnects and delays between research and operational/clinical development of health care innovation		Applying/getting funding for a "testbed," allowing implementation/ testing of innovations; additional activities with potential identified	How prioritized are these activities among MDE activities?	Comparative case studies / contrafactual studies comparing development speed of "testbed innovations" to other health care innovations

One component of a bottleneck analysis also can be determining *which pathways* (Molas-Gallart et al., 2021) or *sociotechnical trajectories* are evolving out of an MDE (Janssen, 2019a). Ideally, a bottleneck analysis should be based on the MDEs' self-identified ways of working to develop pathways or trajectories. *Identifying and articulating such pathways* could mean working with visual cause–effect chains to link actors' actions and new interventions to envisioned change and future output. Based on these kinds of pathways, *summative evaluation of progress in the pathway* may be designed subsequently and assessed using the attribution tools described in the next section. This general idea was partially implemented in evaluations of smaller initiatives using an "outcomes harvesting" approach (Railer et al., 2020) and illustrated in pathway approaches (Griniece et al., 2020).

From Formative to Summative Evaluation: An Attribution Approach to Evaluating MDE's Progress

Complete "elimination" of the bottlenecks exemplified in Table 5 may not be possible. What is important is *progress*, which can be measured in numerous ways (summarized in Table 5's rightmost column). Examples include longitudinal metrics on bottleneck magnitude, counterfactual scenario analyses, follow-up studies after new interventions/activities are launched, case comparisons with similar interventions in other fields, surveys, and external Delphi/expert panel assessments (e.g., Ács et al., 2014; Furubo et al., 2002; Johannesson, 1996). Because bottlenecks may change over time, MDE activities may need to be adjusted accordingly. Hence, evaluations ideally should be performed on an ongoing basis.

The attribution approach towards MDE evaluation aims to *assess outcomes* and *the degree to which an intervention reached its end goals* (Molas-Gallart et al., 2021). As noted in Part 3, this implies a "working backwards" approach, where observable changes are reviewed, followed by an analysis of *whether they can be causally linked to an MDE intervention or activity* (Janssen, 2019a). Such analyses provide a basis for accountability and *ongoing learning* (Mazzucato, 2021) because MDEs should work experimentally, trying out different interventions or initiatives (Lindholm-Dahlstrand et al., 2019).

One example would be to assess whether organizations subjected to interventions an MDE initiated are performing better than others. For instance, in the case of malnutrition among the elderly, assessing whether organizations subject to the MDE's interventions exhibit improved outcomes on central indicators compared to similar organizations. These are standard assessment tools in health, nutrition, and geriatric research, as well as in business settings (A–B testing). Hosts of such assessment tools are available from Swedish health assessment institutes and research groups. Important for such comparisons is validating that the impact stemmed from the MDE activities and not another factor. For a causal interpretation, such interventions must be randomized or, in some other way, resemble randomization across accessible organizations so a control group can be identified (Furubo et al., 2012). If the intervention was launched in a way that made randomization unfeasible, a comparison might still be made. The comparison would use naturally occurring variation in other dimensions of the organizations linked to the likelihood of participating in the interventions, such as the programs the 2021 Nobel Laureates in Economics pioneered—so-called "natural experiments" (Imbens, 2003). This is illustrated in Table 6.

Table 6

Example Intervention Assessment as Evaluation Tool

Factor	Treatment group	Control group
Example: malnutrition	Change in nutrition levels among older adults at care units directly subject interventions by the MDE	Change in nutrition levels among older adults at care units NOT subject to MDE interventions by the MDE
Data	Nutrition levels before and after the intervention	Nutrition levels measured at intervals comparable to the treatment group
Analytical tool	Difference-in-difference comparison between treated and control organizations, matched sample comparisons, etc.	

If participating units exposed to the MDE exhibit improved outcomes on central indicators compared to similar organizations, this could indicate improved underlying conditions. However, it also could signal that those improvements manifested only in the direct "treatment group." In contrast, the overall system operated as usual or may even deteriorate "when in particular incumbents manage to obtain support for clinging on to traditional activities rather than giving space to the emergence of new ones" (Janssen, 2019a, p. 22).

Evaluating whether the Sociotechnical System Demonstrates Improved Performance and Structural Changes

In the intermediate and long terms, policy actors launching MDEs would want to evaluate *whether* and *how* the sociotechnical system(s) targeted demonstrate *improved performance* (formative/summative evaluation Step D, Table 4) and whether the system exhibits any *structural changes of the type intended to reach the vision* or in other ways (summative evaluation Step E, Table 4). As noted following, such evaluations can be conducted ex post of the current MDE initiatives because systematic change often takes years to accrue. As such, these types of evaluations should be conducted by the policy actor or external evaluators working on their behalf and not the MDEs themselves. Next, we outline our conclusion from Parts 1–3 above how such ex post evaluations could be conducted:

Evaluating if and how MDEs Lead to Improved System Performance

Whether and *how* the sociotechnical system demonstrates improved performance needs to be based on system-field-level data beyond the actors directly involved in or affected by the MDE. Such data could comprise register and network data on actors' innovative activities and innovative outputs, collaboration activities and outcomes, and so forth, ideally compared to a similar field and over time (Powell et al., 1996). Such broadly available system-level data previously were used to assess innovation policies in Sweden and internationally (Aghion et al., 2021; Grillitsch et al., 2019) and can be readily employed. As noted in Part 3, progress on such indicators alone does not necessarily tell whether the MDE is working towards achieving the intended objectives. However, causal effects are more plausible if the assessment framework's preceding parts also point in this direction.

### Evaluating if and how MDEs Spur Structural Changes

Evaluating *whether* and *how* the sociotechnical system exhibits any structural changes could focus on specific changes in the innovation *direction*, analyzing *whether organizations* in the sociotechnical system are *shifting their R&D investments and innovative activities towards topics prioritized relative to the MDEs' visions*.

For such assessment, text-based algorithms could be deployed relatively easily to assess the type of innovative activities and outputs actors generated in the relevant field(s). For instance, Prochaska and Schiller (2021) examined whether a governmental mission to stimulate innovations related to the bioeconomy (replacing fossil-based with bio-based feedstocks) affected the innovation projects' content. They classified almost 200,000 innovation projects funded by the German federal ministries and used simple text-mining approaches to detect biomass-connected terms and expressions. They calculated the number of bioeconomy-relevant terms in each project title and classified the projects according to the bioeconomy value chain to examine *shifts in the innovation projects' content*. In a setting resembling the MDE-focused setting here, NESTA tentatively assessed progress relative to the U.K. grand challenge, "Use data, artificial intelligence, and innovation to transform the prevention, early diagnosis, and treatment of chronic diseases by 2030.'

NESTA used text algorithms to analyze 37,478 grant applications in the U.K. Research and Innovation's Gateway with information on the project subject, organizations and researchers involved, funding awarded, and outputs such as publications, patents, spinoffs, and technology outputs (software, diagnostic tools, etc.; Mateos-Garcia, 2019). The assessment was based on keyword combinations, such as *artificial intelligence* or *machine learning*, combined with a range of known chronic diseases. They trained a hierarchical topic model to cluster the projects' data and track activity levels and funding in a specific mission field. The authors studied the evolution of the activity levels over time in the mission field, the level of interdisciplinarity in the mission field over time, which and how actors were active in the mission field and their "novelty," the distribution of outcomes in the mission field, and the diversity of technological trajectories in the mission field. Figure 5 showcases that in the United Kingdom, projects classified as belonging to the mission fields focusing on artificial intelligence in chronic disease treatment exhibited a higher-than-average number of new products, technologies, spinouts, publications, and databases compared to the two comparative (traditional) innovation fields of 'AI' or 'Chronic'.

Figure 5

Using Text Algorithms to Assess Output of Mission Fields versus Traditional Innovation Fields
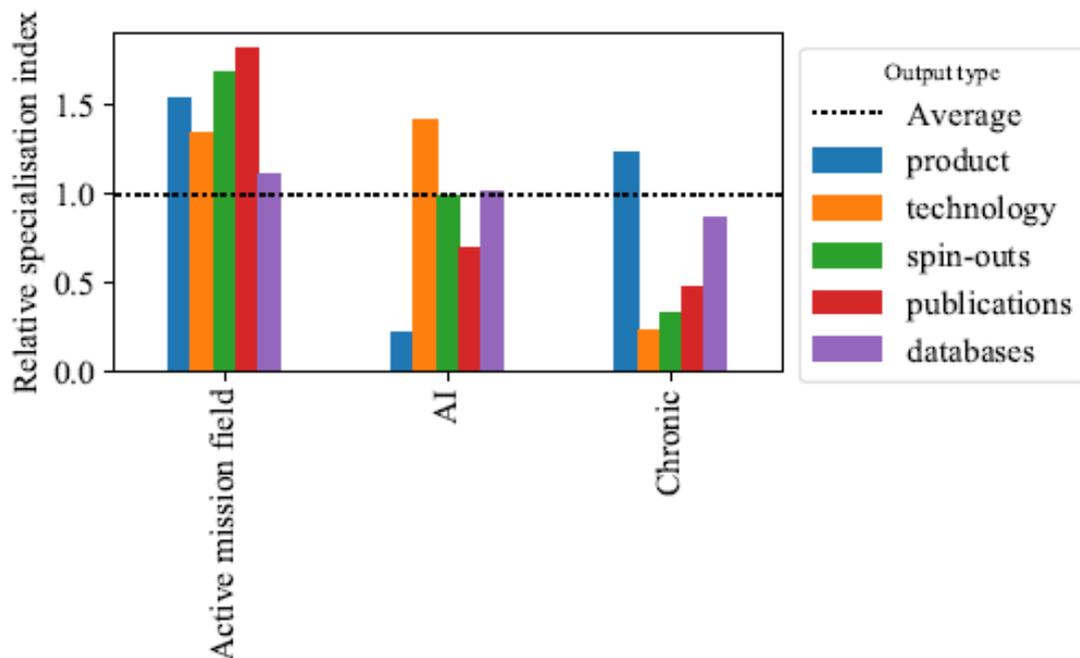
Source: (Mateos-Garcia, 2019)

Other analyses of structural changes spurred by MDE initiatives could probe *whether organizations increase their activity in existing, adjacent, or even new markets or sectors*. This would reflect changes in entrepreneurial experimentation (Lindholm-Dahlstrand et al., 2019) and could be gauged using microdata on firms' and other actors' innovative and market activities over time, within and across sections (Andersson & Klepper, 2013). Analyses of structural changes spurred by MDEs could also seek to examine *whether novel R&D and innovation collaboration patterns are emerging*. Here it is relevant to analyze if recently formed partnerships structurally differ from those formed at the beginning. This would entail assessing not only the number of between-firm collaborations but also whether the collaborations reflect ties spanning sectors that historically have not collaborated. Readily available tools for assigning novelty R&D and innovation collaboration can be drawn from innovation studies and innovation economics, where co-developed technologies, products, patents, and joint R&D projects are public information often assessed using network analysis (Björk & Magnusson, 2009; Choe & Lee, 2017; Gerken & Moehrle, 2012; Monechi et al., 2017). Such tools are readily available from bibliometricians or innovation economists at any major university.

Evaluating whether there is Measurable Impact at the Societal Level

A final ideal evaluation of MDEs would be to assess whether and how there is measurable impact on the societal level in terms of mitigating the failure addressed and reaching the intervention's "vision" or "mission" (summative evaluation Step F, Table 4). As we documented in Part 2, observations from the MDEs suggest they all see significant challenges

in assessing their own societal development contributions towards the vision because of the great complexity of actors and activities in each field.

Given these challenges, we suggest focusing evaluations with a "summative" orientation on intermediate goals, such as changes in Steps D and E as outlined earlier. Long-term "societal impact" represents the vision articulated. By definition, MDIs should be guided by a concrete vision that is easy to articulate and assessed for being fulfilled (Mazzucato, 2021).

If an assessment of long-term societal impact is conducted, we suggest designing it carefully and interpreting it with caution. As discussed previously, attributing actual societal impacts to transformations in the system focus (potentially triggered by an MDE initiative) is fraught with challenges. Societal impact is a type of system-level output in which change output depends on many combinations of system input, processes, or conditions (Levin et al., 2013).

That some interventions, policy initiatives, and interventions similar to MDEs have measurable end goals that compare policy impacts may help such evaluations. In the best of worlds, these evaluations could be conducted by the policy actor, external authorities, or other external evaluators based on the documented material from A–E-type evaluations.

Portfolio Thinking

The MDEs are designed to have radical visions, experiment, and embrace risk. Many, if not most, are bound to be unsuccessful, especially those seeking to address "wicked" problems. Hence, summative evaluations of types E–F should consider the overall effects of Vinnova's MDE programs from a project/portfolio perspective. Risk diversification can be made relative to focus areas (more or less wicked problems), coordinator organizations (sectors, commitment, experience), and impact pathways that these develop (in some settings, multiple MDEs that are geographically and organizationally dispersed may be potential). If ongoing evaluations of the C–D type are conducted, evaluators or Vinnova's program administrators may facilitate learning not only within the MDEs over time but also between the MDEs as part of an overall program.

References

Ács, Z. J., Autio, E., & Szerb, L. (2014). National systems of entrepreneurship: Measurement issues and policy implications. *Research Policy*, *43*(3), 476–494. https://doi.org/http://dx.doi.org/10.1016/j.respol.2013.08.016

Aghion, P., Antonin, C., & Bunel, S. (2021). *The power of creative destruction*. Harvard University Press.

Amanatidou, E., Cunningham, P., Gök, A., & Garefi, I. (2014). Using evaluation research as a means for policy analysis in a "new" mission-oriented policy context. *Minerva*, *52*(4), 419–438. https://doi.org/10.1007/s11024-014-9258-x

Andersson, M., & Klepper, S. (2013). Characteristics and performance of new firms and spinoffs in Sweden. *Industrial and Corporate Change*, *22*(1), 245–280. https://doi.org/10.1093/icc/dts046

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In National Bureau of Economic Research (Ed.), *The rate and direction of inventive activity: Economic and social factors* (pp. 609–626). Princeton University Press.

Aspuru-Guzik, A., & Persson, K. (2018). Materials acceleration platform: Accelerating advanced energy materials discovery by integrating high-throughput methods and artificial intelligence [Innovation Challenge 6]. Mission Innovation. http://nrs.harvard.edu/urn-3:HUL.InstRepos:35164974

Åström, T., Arnold, E., & Olsson, J. (2021). *Metautvärdering av tredje omgången strategiska innovationsprogram efter sex år Vinnova Rapport* VR 2021:10. Faugert & Co Utvärdering/Technopolis Sweden.

Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., & Rickne, A. (2008). Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. *Research Policy*, *37*(3), 407–429. https://doi.org/10.1016/j.respol.2007.12.003

Björk, J., & Magnusson, M. (2009). Where do good innovation ideas come from? Exploring the influence of network connectivity on innovation idea quality. *Journal of Product Innovation Management*, *26*(6), 662–670. https://doi.org/10.1111/j.1540-5885.2009.00691.x

Bloom, N., Van Reenen, J., & Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives*, *33*(3), 163–184.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brown, R. (2021). Mission-oriented or mission adrift? A critical examination of mission-oriented innovation policies. *European Planning Studies*, *29*(4), 739–761.

Choe, H., & Lee, D. H. (2017). The structure and change of the research collaboration network in Korea (2000–2011): Network analysis of joint patents. *Scientometrics, 111*(2), 917–939. https://doi.org/10.1007/s11192-017-2321-2

Coenen, L., Asheim, B., Bugge, M. M., & Herstad, S. J. (2017). Advancing regional innovation systems: What does evolutionary economic geography bring to the policy table? *Environment and Planning C: Politics and Space*, *35*(4)*, 600–620. https://doi.org/10.1177%2F0263774X16646583

Czarnitzki, D., & Fier, A. (2002). Do innovation subsidies crowd out private investment? Evidence from the German service sector. *Applied Economics Quarterly, 48*(1), 1–25.

Douthwaite, B., Alvarez, S., Cook, S. E., Davies, R., George, P., Howell, J., Mackay, R., & Rubiano Mejía, J. E. (2007). Participatory impact pathways analysis: A practical application of program theory in research-for-development. *Canadian Journal of Program Evaluation*, *22*(2), 127–159.

Economic and Societal Impact of Research. (2017). *Towards a mission-oriented research and innovation policy in the European Union: An ESIR memorandum* [European Commission Technical Report]. https://doi.org/10.2777/715942

Edquist, C. (2010). Systems of innovation perspectives and challenges. *African Journal of Science, Technology, Innovation and Development*, *2*(3), 14–45.

Edquist, C., & Zabala-Iturriagagoitia, J. M. (2012). Public procurement for innovation as mission-oriented innovation policy. *Research Policy*, *41*(10), 1757–1769. https://doi.org/https://doi.org/10.1016/j.respol.2012.04.022

Eisenack, K., Villamayor-Tomas, S., Epstein, G., Kimmich, C., Magliocca, N., Manuel-Navarrete, D., Oberlack, C., Roggero, M., & Sietz, D. (2019). Design and quality criteria for archetype analysis. *Ecology and Society*, *24*(3), 6. https://doi.org/10.5751/ES-10855-240306.

Engberg, E., Tingvall, P. G., & Halvarsson, D. (2021). Direct and indirect effects of private-and government-sponsored venture capital. *Empirical Economics*, *60*(2), 701–735. https://doi.org/10.1007/s00181-019-01770-w

European Commission. (2018). Mission-oriented research and innovation: Inventory and characterisation of initiatives. European Commission EC.

Foray, D., Mowery, D. C., & Nelson, R. R. (2012). Public R&D and social challenges: What lessons from mission R&D programs? *Research Policy*, *41*(10), 1697–1702. https://doi.org/https://doi.org/10.1016/j.respol.2012.07.011

Forss, K., Rebien, C. C., & Carlsson, J. (2002). Process use of evaluations: Types of use that precede lessons learned and feedback. *Evaluation*, *8*(1), 29–45. https://doi.org/10.1177%2F1358902002008001515

Furubo, J.-E., Rist, R. C., & Sandahl, R. (2002). *International atlas of evaluation*. Transaction.

Gerken, J. M., & Moehrle, M. G. (2012). A new instrument for technology monitoring: Novelty in patents measured by semantic patent analysis. *Scientometrics*, *91*(3), 645–670. https://doi.org/10.1007/s11192-012-0635-7

Grilli, L., Mazzucato, M., Meoli, M., & Scellato, G. (2018). Sowing the seeds of the future: Policies for financing tomorrow's innovations. *Technological Forecasting and Social Change*, *127*, 1–7. https://doi.org/10.1016/j.techfore.2017.10.021

Grillitsch, M., Hansen, T., Coenen, L., Miörner, J., & Moodysson, J. (2019). Innovation policy for system-wide transformation: The case of strategic innovation programs (SIPs) in Sweden. *Research Policy*, *48*(4), 1048–1061. https://doi.org/10.1016/j.respol.2018.10.004

Griniece, E., Angelis, J., Reid, A., Vignetti, S., Catalano, J., Helman, A., Barberis Rami, M., & Kroll, H. (2020). Guidebook for socio-economic impact assessment of research infrastructures. Ri-Paths.

Hamrefors, S. (2010). Communicative leadership. *Journal of Communication Management*, *14*(2), 141–152. https://doi.org/10.1108/13632541011034592

Hekkert, M. P., Suurs, R. A., Negro, S. O., Kuhlmann, S., & Smits, R. E. (2007). Functions of innovation systems: A new approach for analysing technological change. *Technological Forecasting and Social Change*, *74*(4), 413–432. https://doi.org/10.1016/j.techfore.2006.03.002

Hellquist, A., & Birksjö, D. (2021). Styrning och implementering av innovationspolitik genom samverkan. Stockholm: Tillväxtanalys.

Hertting, N., & Vedung, E. (2012). Purposes and criteria in network governance evaluation: How far does standard evaluation vocabulary takes us? *Evaluation*, *18*(1), 27–46. https://doi.org/10.1177%2F1356389011431021

Howell, S. T. (2017). Financing innovation: Evidence from R&D grants. *American Economic Review*, *107*(4), 1136–1164. https://doi.org/10.1257/aer.20150808

Humphries, J. E. (2021). The causes and consequences of self-employment over the life cycle [Working paper]. Yale University.

Imbens, G. W. (2003). Sensitivity to exogeneity assumptions in program evaluation. *American Economic Review*, *93*(2), 126–132. https://doi.org/10.1257/000282803321946921

Janssen, M. J. (2019a). *Methodology for the IGCI initial impact evaluation* v.17-12-2019. Copernicus Institute of Sustainable Development, Utrecht University.

Janssen, M. J. (2019b). What bangs for your buck? Assessing the design and impact of Dutch transformative policy. *Technological Forecasting and Social Change*, *138*, 78–94. https://doi.org/10.1016/j.techfore.2018.08.011

Janssen, M. J., Bergerk, A., & Wesseling, J. H. (2021). Evaluating systemic innovation and transition programs: Towards a culture of learning. *PLoS One: Sustainability and Transformation*, *1*(3), 1–6. https://econpapers.repec.org/scripts/redir.pf?u=https%3A%2F%2Fdoi.org%2F10.1371%252Fjournal.pstr.0000008;h=repec:plo:pstr00:0000008

Janssen, M. J., Torrens, J., Wesseling, J. H., & Wanzenböck, I. (2021). The promises and premises of mission-oriented innovation policy: A reflection and ways forward. *Science and Public Policy*, *48*(3), 438–444. https://doi.org/10.1093/scipol/scaa072

Janssen, M., Wesseling, J., Torrens, J., Weber, M., Klerkx, L., & Penna, C. (2021). Missions as boundary objects for transformative change: Understanding coordination across policy, research and stakeholder communities. *DRUID Conference 2021, Copernicus Institute of Sustainable Development, Utrecht University*.

Johannesson, M. (1996). *Theory and methods of economic evaluation of health care* (Vol. 4). Springer Science & Business Media.

Kuhlmann, S., & Rip, A. (2018). Next-generation innovation policy and grand challenges. *Science and Public Policy*, *45*(4), 448–454.

Kuittinen, H., Polt, W., & Weber, K. (2018). Mission Europe. A revival of mission-oriented policy in the European Union. In RFTE–Council for Research and Technology Development (Ed.), *Rethinking Europe: Positions on shaping an idea* (pp. 191–207). Austrian Council for Research and Technology Development.

Kvale, S., Brinkmann, S., & Torhell, S.-E. (2014). *Den kvalitativa forskningsintervjun*. Studentlitteratur.

Levin, S., Xepapadeas, T., Crépin, A.-S., Norberg, J., De Zeeuw, A., Folke, C., Hughes, T., Arrow, K., Barrett, S., & Daily, G. (2013). Social-ecological systems as complex adaptive systems: modeling and policy implications. *Environment and Development Economics*, *18*(2), 111–132. https://doi.org/10.1017/S1355770X12000460

Lindholm-Dahlstrand, Å., Andersson, M., & Carlsson, B. (2019). Entrepreneurial experimentation: A key function in systems of innovation. *Small Business Economics*, *53*(3), 591–610. https://doi.org/10.1007/s11187-018-0072-y

Lindman, Å., & Söderholm, P. (2016). Wind energy and green economy in Europe: Measuring policy-induced innovation using patent data. *Applied Energy*, *179*, 1351–1359. https://doi.org/10.1016/j.apenergy.2015.10.128

Magro, E., & Wilson, J. R. (2013). Complex innovation policy systems: Towards an evaluation mix. *Research Policy*, *42*(9), 1647–1656. https://doi.org/10.1016/j.respol.2013.06.005

Mateos-Garcia, J. (2019). Mapping research and innovation missions: With an application to the UK government mission to transform the prevention, diagnosis and treatment of chronic diseases using artificial intelligence. SSRN. https://dx.doi.org/10.2139/ssrn.3483203

Mazzucato, M. (2018). Mission-oriented innovation policies: Challenges and opportunities. *Industrial and Corporate Change*, *27*(5), 803–815.

Mazzucato, M. (2021). Mission economy: A moonshot guide to changing capitalism. Penguin UK.

Mazzucato, M., & Robinson, D. K. (2018). Co-creating and directing Innovation Ecosystems? NASA's changing approach to public-private partnerships in low-earth orbit. *Technological Forecasting and Social Change*, *136*, 166–177. https://doi.org/10.1016/j.techfore.2017.03.034

Miles, I., & Cunningham, P. (2006). *A practical guide to evaluating innovation programs*. European Communitoes, Brussels-Luxembourg.

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

Monechi, B., Ruiz-Serrano, A., Tria, F., & Loreto, V. (2017). Waves of novelties in the expansion into the adjacent possible. *PloS One*, *12*(6), e0179303. https://doi.org/10.1371/journal.pone.0179303

Morgan, K., & Marques, P. (2019). The public animateur: Mission-led innovation and the "smart state" in Europe. *Cambridge Journal of Regions, Economy and Society*, *12*(2), 179–193. https://doi.org/10.1093/cjres/rsz002

Nauwelaers, C., Boekholt, P., Mostert, B., Cunningham, P., Guy, K., Hofer, R., & Rammer, C. (2009). *Policy mixes for R&D in Europe*. European Commission–Directorate-General for Research, Maastricht.

Nelson, R. R. (1977). *The moon and the ghetto*. Norton.

Nelson, R. R. (1993). *National innovation systems: A comparative analysis*. Oxford University Press on Demand.

Nelson, R. R. (2011). The moon and the ghetto revisited. *Science and Public Policy*, *38*(9), 681–690. https://doi.org/10.1093/scipol/38.9.681

Organisation for Economic Co-operation and Development. (2021). The design and implementation of mission-oriented innovation policies: A new systemic policy approach to address societal challenges [OECD Science, Technology and Industry Policy paper No. 100]. https://www.oecd-ilibrary.org/science-and-technology/the-design-and-implementation-of-mission-oriented-innovation-policies_3f6c76a4-en

Pollitt, C. (2003). *The essential public manager*. McGraw-Hill Education.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, *41*(1), 116–145. https://doi.org/10.2307/2393988

Prochaska, L., & Schiller, D. (2021). An evolutionary perspective on the emergence and implementation of mission-oriented innovation policy: The example of the change of the leitmotif from biotechnology to bioeconomy. *Review of Evolutionary Political Economy*, *2*(1), 141–249. https://doi.org/10.1007/s43253-021-00033-8

Pyka, A. (2017). Dedicated innovation systems to support the transformation towards sustainability: Creating income opportunities and employment in the knowledge-based digital bioeconomy. *Journal of Open Innovation: Technology, Market, and Complexity*, *3*(4), 27. https://doi.org/10.1186/s40852-017-0079-7

Railer, J., Stockley, D., Flynn, L., Hastings-Truelove, A., & Hussain, A. (2020). Using outcome harvesting: Assessing the efficacy of CBME implementation. *Journal of Evaluation in Clinical Practice*, *26*(4), 1132–1152. https://doi.org/https://doi.org/10.1111/jep.13359

Ramböll. (2021). Programmet Utmaningsdriven Innovation: Slutrapport. Vinnova.

Raven, R., & Walrave, B. (2020). Overcoming transformational failures through policy mixes in the dynamics of technological innovation systems. *Technological Forecasting and Social Change*, *153*, 119297. https://doi.org/10.1016/j.techfore.2018.05.008

Rodrik, D. (2008). *Industrial Policy for the Twenty-first Century*. Princeton University Press.

Roth, A. E. (2012). *The theory and practice of market design* [Nobel Prize Lecture]. https://www.nobelprize.org/prizes/economic-sciences/2012/roth/lecture/.

Sabel, C. F., & Zeitlin, J. (2012). Experimentalist governance. In D. Levi-Faur (Ed.), *The Oxford handbook of governance* (pp. 2–4). Oxford University Press.

Schot, J., & Steinmueller, W. E. (2018). Three frames for innovation policy: R&D, systems of innovation and transformative change. *Research Policy*, *47*(9), 1554–1567. https://doi.org/10.1016/j.respol.2018.08.011

Scriven, M. (1991). *Evaluation thesaurus*. Sage.

Smutylo, T. (2005). *Outcome mapping: A method for tracking behavioural changes in development programs* [ILAC Brief 7]. https://cgspace.cgiar.org/bitstream/handle/10568/70174/ILAC_Brief07_mapping.pdf?sequence=1&isAllowed=y

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. Jossey-Bass.

Thirtle, C. G., & Ruttan, V. W. (1987). *The role of demand and supply in the generation and diffusion of technical change* (Vol. 21). Harwood Academic.

Tillväxtanalys. (2020). Den tredje generationens innovationspolitik: Kunskapsöversikt och problematisering [Tillväxtanalys rapport 2020:12, Dnr: 2020/45].

Van Poeck, K., Östman, L., & Block, T. (2020). Opening up the black box of learning-by-doing in sustainability transitions. *Environmental Innovation and Societal Transitions*, *34*, 298–310. https://doi.org/10.1016/j.eist.2018.12.006

Vinnova. (2020). *Mission-oriented innovation in Sweden: A practice guide*. www.vinnova.se

Wanzenböck, I., Wesseling, J. H., Frenken, K., Hekkert, M. P., & Weber, K. M. (2020). A framework for mission-oriented innovation policy: Alternative pathways through the problem–solution space. *Science and Public Policy*, *47*(4), 474–489. https://doi.org/10.1093/scipol/scaa027

Warwick, K., & Nolan, A. (2014). Evaluation of industrial policy: methodological issues and policy lessons [OECD Science, Technology, and Industry Policy Paper No. 35]. https://www.oecd-ilibrary.org/science-and-technology/evaluation-of-industrial-policy_5jz181jh0j5k-en

Weber, K. M., & Rohracher, H. (2012). Legitimizing research, technology and innovation policies for transformative change: Combining insights from innovation systems and multi-level perspective in a comprehensive "failures" framework. *Research Policy*, *41*(6), 1037–1047. https://doi.org/10.1016/j.respol.2011.10.015

Wesseling, J. H., & Edquist, C. (2018). Public procurement for innovation to help meet societal challenges: A review and case study. *Science and Public Policy*, *45*(4), 493–502. https://doi.org/10.1093/scipol/scy013

Appendix A: Survey Questions

Uppföljning av visionsdrivna miljöer

Enkät till kärnteamsmedlemmar

1. Den här enkäten berör visionsdrivna innovationsmiljöer inom hälsa som etablerats med finansiering av Vinnova. Vilken miljö har du bäst kännedom om? (Flervalsfråga)
2. Är det möjligt att nå er vision? (Ja/Nej)
3. Eventuell kommentar (Fritext)
4. Är det möjligt att konkret följa upp huruvida ni når eller kommer närmare er vision? (Ja/Nej)
5. Eventuell kommentar (Fritext)
6. Lista de viktigaste aktiviteterna ni genomfört så här längt som du tror varit meningsfulla för att föra er närmare visionen. (Fritext)
7. Vilka förändringar som involverar andra aktörer har ni sett eller hoppas få se som en konsekvens av miljöns arbete? (t ex. innovativ teknik, infrastruktur, affärsmodeller, policyutveckling, kultur och beteenden) (Fritext)
8. Hur anser du att de förändringarna skulle kunna följas upp? (Fritext)
9. Vad tycker du är viktigt att följa upp på kort sikt för att utvärderingen ska bli användbar för er som miljö? (Fritext)
10. Vad är viktigt att följa upp på lång sikt? (Fritext)
11. Vad ser du för utmaningar med att följa upp miljöns resultat enligt ovan? (Fritext)
12. Vad tror du ni klarar av att följa upp själva / var krävs extern hjälp? (Fritext)
13. Flexibla arbetsformer efterfrågas av Vinnova. Hur skulle man kunna följa upp/utvärdera flexibilitet på ett sätt som skulle bli användbart för miljön? (Fritext)
14. Har du tänkt på eller diskuterat några etiska aspekter på miljöns arbete? I så fall vilka, och hur tänker du eller resonerade ni kring dessa? (Fritext)
15. Ser du några risker för oönskade riktningar i miljöns arbete eller resultat som gör det viktigt att följa upp särskilda aspekter? (Fritext)
16. Är det önskvärt att er miljö överlever på längre sikt?

Appendix B: Summary of Mission-Oriented "Work Principles" from the Literature

Mission-Oriented "Work Principles" (1/2)	Mazzucato, M. (2021). *Mission economy. A moonshot guide to changing capitalism*	Janssen (2019). *Methodology for Australia's IGCI initial impact evaluation*	Janssen (2020). *Methodology for Australia's IGCI initial impact evaluation*	Brown (2021). *Mission-oriented or mission adrift?*	Wanzenböck et al. (2020) *Framework for mission-oriented innovation policy* [wicked problems]
Vision & Leadership	*Leadership: Vision and purpose.* Should depart from a problem and provide a direction. "A vision infused with a strong sense of purpose" (p. 60). Should *not* focus on a technology; rather, the problem that innovation can solve). Should be measurable and time bound.	*Leadership competence.* Effective policy execution in implementing agencies, able to get commitment	*Entrepreneurial experimentation*: Entrepreneurs can be either new startups or incumbents diversifying into new economic activities.	MO policy in Scottish National Investment Bank (SNIB) "**Fuzzy" in what missions the bank should target**. The UN's 17 SDGs useful but "too broad" to be actionable (Mazzucato & and Macfarlane 2019, p. 10).	*Complexity* of wicked problems and how they can be addressed
Organization	*Risk-taking and innovation.* Should include experimentation, adapting to new information and circumstances. Tolerance for failures	*Open architecture*: Engage and target forward-looking "domains" and not stick to sectors boundaries	*Knowledge development*: R&D and experience-based learning are fundamental. Knowledge about new technologies, products, or services but also markets, networks, and users	**MO policies in SNIB lack detail** regarding how actors are structured, what types of innovation is supported, products offered, and how potential customers are targeted.	

	Organizational dynamism. Enable actors to "see the whole" and their own part; communication between functions dynamic	***Focus on change***: Diversify, do not just back winners' incremental improvements		**Knowledge exchange**: Joint projects, conferences, etc.	Uncertainty
	Avoid complexity paradox of modern public policy: The more complex issues are, the more compartmentalized policy-making becomes, fragmented into different and competing initiatives and departments (p. 75). ***Dynamic capabilities*** to integrate, build, and reconfigure internal and external competences to address changing environments. Rules of thumb that are adjusted through experimentation	***Information retrieval***: Mechanisms for identifying system-specific problems and solutions	Guiding direction of search: key processes summarizing all activities and events, convincing actors to enter a certain development path or further invest in it. Clear, positive expectation about potential to move in a certain direction.		
Stakeholders & Interaction	Dynamic partnership between public and private actors; beyond use of consultants Partnerships with flexible procurement contracts (direct, not via intermediaries) ***Create competition***: Share risks and rewards of value creation. Citizen organizations should be represented in evaluating.		***Market formation:*** Creation of markets is necessary for them to further develop. Such niches can be found in export markets or created with favorable tax regimes, governmental procurement policies.	MO policies in SNIB **not aligned with innovation ecosystem targeted.** In the Scottish case, a bifurcation between strong R (world-class universities) and weak D (low-productive SMEs)	***Contestation*** of the nature of wicked problems and how they can be addressed

	Collaboration and spillovers across multiple sectors. Serendipity (unexpected benefits) should be embraced, major hurdles negotiated.	*Broad scope of support:* Adjust institutions and structures in the ecosystem	*Resource mobilization:* Sufficient resources are needed for the emerging development paths to function properly. Complementary assets must be developed, such as complementary products, services, and network infrastructure**.**	Selecting missions lack practicable mechanisms to be "open and inclusive," i.e., democratic (see Roth's Nobel lecture, https://www.nobelprize.org/uploads/2018/06/roth-lecture.pdf)	*Complexity* of wicked problems and how they can be addressed
Outcomes	Long-term financial horizons and budgeting that focuses on outcomes (for actors like Vinnova). Ensuring system is moving in direction of mission by achieving intermediate milestones and user engagement Spending to crowd in private financial investment	*Accountability:* Transparency, proactive communication strategy, visible / responsible leadership *Outcome inclusivity:* Structural change, not profits; ensure new knowledge is broadly accessible *Policy adaptation:* Learning about effectiveness, changing policy course	*Legitimation/counteracting resistance:* Create legitimacy, favorable tax regimens, and put the new technology on the political agenda.		
Mission-Oriented "Work Principles" (2/2)	**Hertting & Vedung. (2012).** *Purposes and criteria in network governance evaluation.*	**Morgan & Marques. (2019).***The public animateur: Mission-led innovation and "smart state" in Europe*	**Sabel & Seitlin. (2012).** *Experimental governance*	**Hellquist & Birksjö. (2021). Styrning och implementering av innovationspolitik genom samverkan; Tillväxtanalys. (2020).**	
Vision & Leadership	Work towards fulfilment of clear goals; facilitate accountability		Vision operationalized in broad framework goals and metrics should be provisionally established by central and local units	Collaborations led by: **One actor as primary responsible** for one question but consults with others (consultation)	

				Actors coordinating activities to reach results tied to common interests, although activities performed isolated by the actors (coordination) **Closer collaboration** with actors performing activities together by using shared resources (Hjortstjö, 2005).
Organization	Collaboration can be a means to strengthen the legitimacy of other governance modes.	Adopt **multi-scalar policy, respecting principle of subsidiarity** to avoid higher level authorities encroaching in a heavy-handed fashion on lower-level decision-making Promote culture of collaboration and trust (stronger feedback mechanisms) Avoid rules and regulations that are ungraspable by beneficiaries and authorities involved. Emphasize experimentation and allow failures; need for feedback mechanisms" (p. 185). **Dismantle barriers that hinder learning (e.g., silo structures or high staff turnover);** maintain some stability that allows continuing monitoring and evaluation to feed back into the system	**Local units should be given broad autonomy** and discretion to pursue these goals in their own way.	**Balance collaboration** (network governance) and hierarchical forms of governance (not least in the environment) Allow **questioning of current norms and institutions**. Double loop learning (Argyris, 2002) Build **trust in** not only participants but **the process** and decision order; important for controversial and conflictful issues Allow experimentation**, niche experiments protected from established actors' resistance;** important to allow the operative testing of measures and create a minimal experience; at strategic and operational level.

Stakeholders & Interaction	*Democratic participation:* Are various relevant societal groups represented in the collaboration? Be mindful of what communicative ideals are upheld, e.g., **deliberative or agonistic ideal for democratic** dialogue	Organizations that have the technical skills and social capital to engage with a variety of agents and to maintain cohesion	Goals, metrics, decision-making proceduresshould be revised by widening circle of actors in response to problems and possibilities revealed by the peer-review process	Interorganizational collaboration (*samverkan*) clarifies different sectors' perspectives, which may reduce institutional uncertainty in innovation. Collaborations, however, contribute to a minor extent to cross-sectorial agreement and prioritization. The arguments often become more of '*partsinlagor*'. Industry associations (*Branschorganisationer*) provide legitimacy because they represent many firms but can only communicate consensus. **Conflicts are hidden**. **Risk that collaborations often involve established incumbents**, which often leads to small, incremental changes rather than radical innovation. **Collaboration may reduce possibilities for accountability and transparency;** already strong actors increase their influence at the expense of weaker groups.
Outcomes	Have (in)direct costs tied to governance of the collaboration been justified relative to outcomes? Institutionalization. Does the collaboration lead to persistent and durable results? Synergies. Did the collaboration lead to the actors furthering their ambitions more than if they had acted on their own? Does collaboration strengthen trust among actors—a fundamental prerequisites for democratic society.			Van Poeck et al. (2020). Balance democratic and pragmatic values when deciding scope if societal change is desired. Actors are legitimate participants in decisions based on (a) **being affected directly by a decision** (production/consumption of an innovation) (b) having a formal mandate to represent others (being "elected") (c) expert knowledge (d) having a **legitimate right to represent others** based on similarities in lived experience/opinions