



Applying Theory of Change in research program planning: Lessons from CGIAR

Brian M. Belcher^{a,*}, Enrico Bonaiuti^b, Graham Thiele^c

^a Sustainability Research Effectiveness Program, College of Interdisciplinary Studies, Royal Roads University, Victoria V9B 5Y2, Canada

^b International Center for Agricultural Research in Dry Areas (ICARDA), Beirut, Lebanon & International Potato Center (CIP), Lima, Peru, Via Marcello Malpighi 4, Firenze 50139, Italy

^c Independent consultant, Caparus 400, Santa Cruz de la Sierra, Bolivia

ARTICLE INFO

Keywords:

Theory of Change
Research planning
Research evaluation
Research-for-development

ABSTRACT

Theory of Change (ToC) is widely used as a tool to support strategic planning, monitoring, and evaluation in many fields, especially for social and environmental programs. However, there is still limited documented experience with the application and use of ToC in a research context. CGIAR, a global network of 15 centers conducting international research-for-development, included a standardized ToC approach in a recent round of developing 32 large research Initiatives. This unique experience offers an ideal opportunity to learn from organization-scale ToC implementation and use. The paper provides an overview of research-for-development challenges and ToC concepts and a brief history of ToC use in CGIAR. We describe the application of ToC in this recent case and then assess strengths and weaknesses of the process and the ToCs developed as part of the Initiative proposals. CGIAR made important advances in standardizing ToC concepts and terminology, tools, and guidance, and in integrating ToC into annual reporting and evaluation. Nevertheless, many of the ToCs were insufficiently clear and specific, with substantial scope for further improvement. This is due in part to the rushed and decentralized proposal development process, undertaken during pandemic restrictions, but also reflects different mental-models of research-for-development processes and gaps in understanding and capacity. Recommendations to improve development and use of ToC include capacity development in conceptualizing research impact pathways, ensuring that research design teams have a dedicated M&E specialist paying particular attention to ToCs, improved ToC templates, and better accountability for ToC development and use over the life of a program.

1. Introduction

Theory of Change (ToC) has become popular as a tool to support strategic planning, monitoring, and evaluation in philanthropy, international development, social enterprise, and other fields that aim to contribute to social change (Brest, 2010; James, 2011; Vogel, 2012). Building on that experience, many research funders now require ToCs in proposals, especially for challenge-driven, impact-oriented research. This is intended to encourage context analysis and planning beyond research outputs to articulate how the research is expected to lead to higher-level outcomes and impacts (Hessels et al., 2023). While there is great potential in this approach, there is still much to learn about how to design and use ToC in a research context.

A ToC is a model of a change process, presented as a diagram and/or

a narrative description, which positions a program or project within a wider analysis of how change comes about. It is essentially a map of a sequence of events leading to a desired change. ToC is part of a larger set of program theory approaches, which includes logical frameworks (logframes), results chains, and realist matrices that aim to facilitate critical thinking about change processes, make causal relationships explicit, and improve intentionality in design and implementation (Prinsen and Nijhof, 2015; Better Evaluation, n.d.). As such, ToC has great potential for use in research approaches that involve active and deliberate engagement with stakeholders and/or other system actors, with multiple impact pathways operating in complex systems, known as transdisciplinary research (TDR), sustainability science, research-for-development (R4D), and change-oriented research more generally (Vogel, 2012; Oberlack et al., 2019; Belcher and Hughes,

* Corresponding author.

E-mail addresses: brian.belcher@royalroads.ca (B.M. Belcher), E.Bonaiuti@cgiar.org (E. Bonaiuti), grahamthiele@hotmail.com (G. Thiele).

<https://doi.org/10.1016/j.envsci.2024.103850>

Received 27 February 2024; Received in revised form 17 July 2024; Accepted 23 July 2024

Available online 27 July 2024

1462-9011/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

2021). A ToC helps model and manage the inherent complexity and provides a reference for on-going reflection on the change process (James, 2011; Oberlack et al., 2019). Ideally, a ToC is developed in collaboration with stakeholders and partners, thereby facilitating more inclusive priority-setting and improved problem diagnosis, research design, and shared vision.

With increased experience in applying ToC in different contexts, there is a growing literature documenting that experience and the lessons learned (James, 2011; Vogel, 2012; Stein and Valters, 2012; van Es et al., 2015; Maini et al., 2018). Lam (2020) reviewed the ToC literature to identify common concerns and suggested a need for more attention to context and complexity, incorporating theory and engaging stakeholders in the process. However, there is still limited experience or documented learning from ToC use in research applications (Deutsch et al., 2021). This paper aims to help fill that gap.

A recent round of research planning and design in CGIAR,¹ a large international agricultural R4D consortium, provides an excellent case to learn from practical experience with ToC in a research context. CGIAR invited multiple proposals for large (up to \$30 M) three-year research ‘Initiatives’ from its member centers. Each proposal was required to have a well-articulated ToC at the program (Initiative) and sub-program (Work Package (WP)) scales. Those ToCs were also expected to link to a higher-level thematic (Action Area) ToC. Initiative and WP ToCs were assessed as part of the proposal evaluation and approval process. Funded Initiatives are expected to report against their ToCs annually.

Although there was some previous experience with using ToC in CGIAR, this standardized and system-level integration into research planning and evaluation is novel. The purpose of this paper is to report and analyze lessons from this effort to promote and embed ToC thinking into R4D program design and implementation at the organizational scale (i.e., CGIAR). It starts with a brief discussion of the role of research in development and an overview of current ToC theory. We describe the case study, with emphasis on the ways ToC use was supported. Our analysis focuses on the quality of Initiative and WP ToCs in terms of their presentation (ToC diagrams and narratives), key elements and evaluability. We conclude with a presentation of lessons learned and recommendations for improvements to ToC support in a research context.

2. Background and context

2.1. Research-for-development

Scientific research is expected to both create knowledge and contribute to human well-being. This dual R4D role is explicit in CGIAR’s mission “to deliver science and innovation that advance transformation of food, land, and water systems in a climate crisis” (CGIAR, n.d.a., p.8). Obviously, to contribute to social, economic, and environmental changes means that research results need to be used. Some types of new knowledge can be embedded in technological innovations, such as improved crop varieties, that are ready to be adopted and used by farmers and other actors. However, even mature technological innovations require social processes for their dissemination, uptake, and use. Furthermore, increasingly complex ‘wicked’ problems cannot be solved with purely technical solutions. They require values-based decision-making, negotiations, and trade-offs (Pielke, 2007; Sarewitz, 2016). To be effective, research strategy, planning, design, and implementation must understand and respond to the needs of intended research users. Scientific knowledge is necessary but insufficient; it must be made useful through a series of transformations performed by different actors in the R4D system (Joly et al., 2015).

Those system actors can be divided into two main categories. First order users help translate and transform knowledge by adapting or packaging scientific outputs to support their use or application

(Bozeman, 2003). This might include knowledge brokering, adapting or developing new technologies for a particular context, or applying knowledge to policy-making (Meyer, 2010). Second order users are those who use knowledge, perhaps packaged as a technology, without seeking to add or create new uses (Bozeman, 2003). Bozeman (2003) notes that an actor can move from being a second order to a first order user when they actively support further use of knowledge, such as lobbying for policy or practice change.

While many CGIAR outputs are ultimately aimed at second order users, often called ‘end users’ or ‘beneficiaries’ in CGIAR parlance, there has been a long history of collaboration with national agricultural systems in research and knowledge mobilization (Byerlee and Lynam, 2020). More recently, CGIAR has increased emphasis on partnerships with other system actors in a variety of knowledge generation and application processes, both upstream (e.g., identifying and defining research needs) and downstream (e.g., knowledge translation, mobilization, scaling), with increased emphasis on knowledge co-generation. This requires different ways of planning, designing, and evaluating the work, which is what ToC is intended to help facilitate. This is manifest for example in the development and use of a CGIAR Quality of Research for Development Framework, an expanded set of research evaluation principles and criteria that go beyond assessing scientific credibility to also include considerations of social relevance, legitimacy, and positioning for use (ISDC, 2020; Meinke et al., 2023). Likewise, CGIAR’s Common Results Reporting includes many indicators of intermediate outcomes, which recognize the actions of first order users.

In other words, there is (intended to be) much more explicit and deliberate attention to the systematic work needed to link research products and processes to other actors and processes to contribute to impact (Woolley and Molas-Gallart, 2023). This is part of what Jaganathan et al. (2023) call the “science of actionable knowledge” (p.175), going beyond naive assumptions that good science will lead directly to good decision-making. It recognizes that the connections between scientific knowledge, decision-making, resulting actions, and their outcomes and impacts are not straightforward. This especially true for complex issues with high stakes, risks, and uncertainties that are inherent in the broad range of challenges addressed by CGIAR. For science to contribute effectively to solving these problems, it is critical to understand how it is used to influence decisions and actions that will enhance social and environmental well-being; that is, how science is made actionable (Meinke et al., 2006; Kirchhoff et al., 2013; Meadow and Owen, 2021).

2.2. ToC in R4D

CGIAR defines ToC as an “explicit, testable model of how and why change is expected to happen along an impact pathway² in a particular context. A basic R4D ToC identifies the context and key actors in a system and specifies the causal pathways and mechanisms by which the research aims to contribute to outcomes and impacts” (CGIAR, 2021, p.4). This definition acknowledges that research alone cannot directly cause mission-level impact³ in complex systems. Research can generate new knowledge and understanding, as well as develop technologies, institutional arrangements, policy options, and other innovations to help solve problems and take advantage of opportunities, but it must be

² An impact pathway is a sub-component of a ToC. It is the sequence of outputs, outcomes, and the relevant assumptions and causal logic that explain a particular high-level outcome within a ToC (CGIAR, 2021, p.8).

³ CGIAR (2021) defines impact as “[a] durable change in the condition of people and their environment brought about by a chain of events or change in how a system functions to which research, innovations and related activities have contributed. It refers to a change of state (e.g. nutritional status; farm productivity; household wealth; gender representation in land rights) or flow (e.g. average annual income; CO2 emissions)” (p.7).

¹ See <https://www.cgiar.org/how-we-work/>

utilized through a variety of social processes. A research project or program can also make substantial and important contributions beyond knowledge creation by supporting capacity development, networking, advocacy, and collective action to help translate and mobilize knowledge to action. A ToC explicitly articulates a change process, from activities and outputs through to desired outcomes and impacts.

The ToC development process is intended to support and document a program strategy. It encourages and facilitates research design teams, research managers, and researchers themselves to analyze the system(s) they aim to influence and design their programs accordingly. This means identifying the other actors and processes that operate in the system and imagining what is needed for the desired changes to transpire. Ideally, this kind of strategic analysis will guide the planning of a program's research activities and other supporting activities (e.g., networking, capacity-building, facilitating negotiated decisions) to effectively inform, support, and otherwise influence those other system actors and processes. It goes beyond simply assuming (or hoping) that outputs will be used, to identify key research users and develop strategies to engage them. In other words, it is a vision and plan of how the proposed work will contribute to the process of change and the intended impacts in context.

Ideally, key partners and stakeholders will be involved in the ToC development process to contribute their knowledge and understanding of the context, identify needs and opportunities, and develop a shared vision. Such engagement is expected to help improve research relevance, enrich analysis with broader perspectives, and foster co-ownership of the research. Practically, it can be difficult and time-consuming to identify and engage appropriate and willing partners (Apgar et al., 2017).

When used for research design, a ToC provides a framework and guide for strategic planning. During implementation, a ToC provides progress markers (indicators) that can be used for monitoring, learning, and adaptive management. After a project is completed, the ToC can be used as a framework for evaluation, providing a basis to define indicators and measures of change that can be tested empirically (Belcher et al., 2020). A good ToC should provide a clear and concise presentation of a program's strategy and plan, so it also serves a communication function.

A ToC is typically presented as a diagram and/or a narrative description. A ToC diagram takes the form of a flow chart, with boxes to illustrate the key steps (activities, outputs, outcomes, and impacts) in the change process, and arrows to show the causal linkages between steps. The ToC narrative provides a more complete description of the context, assumptions, and causal logic, and offers space for detail and nuance. It explains what the program will do, who it will work with, what it will produce (outputs), and how and why (theoretically) it is expected to contribute to the intended outcomes and, ultimately, impacts. The narrative presents the ToC in story form to communicate the key ideas, which may help identify gaps or flaws in the logic. Some readers may more easily digest one form over another. Using both a diagram and narrative description together captures and communicates the key ideas as completely as possible.

In more scientific terms, a ToC is a set of hypotheses. If the intermediate steps in the model are well specified, they can serve as indicators of progress. In the early stages of a program, we can monitor and assess whether the planned outputs have been produced. Over time, we would also expect to see evidence that outcomes (i.e., observable changes in behaviour or actions of other system actors) have been realized. Over longer periods, higher-level outcomes and mission-level impacts should be realized if: i) the underlying theory/assumptions are sound; ii) the program has been implemented as planned; and iii) the context has not changed significantly from what was anticipated in the ToC.

Change processes are complex and difficult (probably impossible) to predict entirely. Therefore, it is unlikely that any ToC will be fully realized. If a ToC is used as a tool for monitoring progress during

program implementation, unrealized or unexpected outcomes can signal weaknesses and errors in the theoretical assumptions, shortcomings in program delivery, and/or changes in the context. Any and all of these kinds of lessons can be used to guide adaptive management of the program. After program completion, the ToC can be used as an evaluation framework for systematic learning about whether and how the research has contributed to outcomes and impacts (Belcher et al., 2020). This is critically important in a R4D context as we seek better understanding of how to design and use research as a tool for change.

2.3. ToC in CGIAR

Program theoretical approaches have been used for many years in CGIAR in the form of logframes. While similar to a ToC, logframes have been criticized for being simplistic, linear, and inflexible (Chambers and Pettit, 2004). The two most important differences from ToC are that logframes typically do not explain the causal logic leading from one step to the next or specify the key actors in the change process.

A version of a more flexible and inclusive theory-based approach, called Participatory Impact Pathway Analysis, was used in several CGIAR programs beginning in the mid-2000s (Douthwaite et al., 2003, 2007; Douthwaite and Ashby, 2005; Douthwaite, n.d.). As CGIAR increased attention to results, there was a greater need for research programs to project and explain how their research would contribute to development impacts (Mayne and Johnson, 2015). In 2011–2012, Consortium-wide CGIAR Research Program (CRP) proposals were required to include basic ToCs. In practice, most of those ToCs were presented as simple flow diagrams, with limited explanation of the underlying theory. CGIAR encouraged further development of ToCs with stakeholders as part of an increased emphasis on Results Based Management around 2014–2015. Although this effort subsequently lost momentum, some CRPs invested in staff training and more detailed ToC development, including work to engage stakeholders in ToC development (Hambly et al., 2022).

The second phase of CRP funding in 2015–2016 required more detailed ToCs in all proposals, but there was still considerable variation in the approach, presentation, and use of ToC. CRPs made different investments and used different templates for their ToCs. Some developed integrated ToCs linking different program levels; some used ToC in management and evaluation of their research projects (Belcher et al., 2019; Ramirez and Belcher, 2019; 2020) and programs (Delahais and Toulemonde, 2017; Halimanjaya et al., 2018; Omore et al., 2019; Coccia et al., 2022; Hambly et al., 2022). Overall, though, there was limited use of ToCs to track progress and adapt programming.

A series of independent evaluative reviews of the 12 CRPs in 2020 included an evaluation question on ToC (Holderness et al., 2021). They concluded that ToC use had facilitated shared thinking about change in a larger context and was useful for priority-setting, assessing the contribution of scientific outputs, seeking and justifying funding, mapping trajectory to impact, and reporting. The greatest value of ToC was in the process to develop them along with their underlying assumptions, which enhanced collaboration. However, the reviews also found:

- Limited use of ToC for routine management, internal learning, and resource allocation;
- The ToC concept was mainstreamed in some CRPs, with much scope for improvement; and
- ToCs were mainly developed because they were required for obtaining CGIAR funding, less for use.

Following CGIAR's 2030 Research and Innovation Strategy (CGIAR, n.d.a.), a new research portfolio was organized as One CGIAR, a coordinated set of Initiatives (programs), each with several Work Packages (sub-programs). The strategy identifies three main pathways: science-based innovation; targeted capacity development; and policy influence and explicitly appreciates that CGIAR operates within broader

innovation systems along with many other actors and processes. The Initiatives are organized within three Action Areas: Systems Transformation, Resilient Agrifood Systems, and Genetic Innovation.

A total of 32 Initiative proposals were prepared and submitted in two rounds. Each proposal included an Initiative-level ToC diagram and narrative, as well as detailed ToC diagrams and narratives for four to six WPs per Initiative. CGIAR made a substantial investment to facilitate and incentivize the integration of ToC, and theory-based research planning more generally. Initiatives were expected to have clear strategies for impact. As part of this, ToCs were to reflect a strategic analysis of the system that the research aims to influence. The overall aim was to shift to a more demand-driven R4D approach. For some teams, this marked a substantial departure from past practice. It required new ways of thinking and working, with more attention to, and anticipation of, the needs and expectations of intended research users (Belcher and Hughes, 2021; Meinke et al., 2023).

We review the ToCs in terms of the quality of presentation (diagrams and narratives), the specification of key elements, and the evaluability of individual ToCs and the overall set.

3. Methods

3.1. One CGIAR case description

One CGIAR proposal development and approval was done in stages, with initial guidance for the Initiatives concept provided early in 2021. Proponents were invited to submit pre-proposals outlining preliminary ideas for Initiatives. Feedback on those ideas was used to guide the development of concept notes that were assessed against key design principles and criteria, including their ToCs. Full proposals, submitted September 28th, 2021, were to include a plan of results and budget, with annual reporting against results anticipated in the ToC.

A guidance document provided direction on ToC design and expectations, with specific definitions of key concepts, including: outputs (“the knowledge, technical or institutional advances produced by CGIAR research, as well as engagement and capacity development activities”); outcomes (“change in knowledge, skills, attitudes and/or relationships (KASR), which manifests as a change in behavior in particular system actors, to which research outputs and related activities have contributed”), and; impacts (“durable change in the condition of people and their environment brought about by a chain of events or change in how a system functions to which research, innovations and related activities have contributed”) (CGIAR, 2021, p.21–23). ToCs are framed within spheres of control (i.e. under the direct control of the Initiative, up to outputs), influence (i.e. actions of other system actors = outcomes) that can be influenced by initiative activities and outputs), and interest (i.e. outcomes and impacts that can only be influenced indirectly, through the actions of other system actors). The guidance distinguished theoretical and context assumptions, indicating that context assumptions should focus on key factors that could affect outcomes. Theoretical assumptions should explain why system actors are expected to respond to precursor outputs and outcomes as hypothesized in the ToC. This requires understanding the motivations and circumstances of those other actors; the ToC process was intended to encourage strategic analysis and learning from experience. The ToCs modeled change processes up to impacts (sphere of interest), but the Initiatives were funded for just three years in the first phase. To deal with the inevitable time-lags between research-based interventions and impacts, the concept of end-of-initiative outcomes was introduced. End-of-initiative outcomes are defined as outcomes that are challenging but realistic to be realized by end of initiative and are the highest-level of results for which the Initiatives are responsible. Nevertheless, the full ToC, including longer-term, higher-level results, needs monitoring with appropriate indicators and focused impact assessments (Belcher and Hughes, 2021).

Online seminars presented key ideas and expectations and provided an opportunity for initiative design teams to ask questions and raise

concerns. The CGIAR System Organization supported ToC resource people, including the first author, to provide guidance and feedback. Some design teams drew on other external or internal ToC experts. Some, but not all, teams took the opportunity to hold team meetings with ToC resource people. With tight deadlines and geographically dispersed teams, none were able to organize dedicated ToC workshops with their larger teams during the proposal development phase. ToC development was typically led by the design team leaders working in an iterative manner with some team members and sharing drafts for feedback. Most teams shared draft ToCs with resource people and used the feedback provided to revise ToCs, sometimes over several rounds. It is safe to say that none of the Initiatives was developed in a genuinely participatory mode with stakeholders. Teams with long-standing partnerships were better placed to consult bilaterally and/or anticipate and incorporate partner and other stakeholder interests.

A set of facilitating questions (Text Box 1) was provided to stimulate thinking about the context in which the Initiatives were proposing to work. The questions start with broad analysis of the development challenge and the high-level changes the Initiative or WP aims to contribute to and the key actors and processes that need to be engaged. They then focus on the research and other activities required to stimulate and support the change process, next identifying the necessary outputs and the expected outcomes of the work. It was recommended that teams consider those questions and use the responses to draft a ToC diagram and narrative.

A detailed proposal template included a section for a one-page ToC diagram using a standard format along with a 500-word narrative ToC description for the full Initiative, and a diagram and 1000-word narrative for each WP. A draft version (July 23, 2021) of the template specified a 500-word presentation of the research questions, methods, and key intended outputs, followed by a 500-word ToC description explaining the causal logic and underlying assumptions. However, the final template emphasized activities and contextual assumptions, with less attention on explanation of the ToC. The final template also suggested sub-headings, including some aspects not normally included in a ToC, which resulted in incoherent and less focused narratives.

It was intended that future evaluations would consider ToC use as an assessment criterion. Annual reporting would focus on results anticipated in the ToCs, and evaluations would consider whether and how ToCs have been adjusted based on learning and/or changing contexts. Also, it was intended that stage-gating reviews at the end of the first phase of Initiatives would consider ToC use as a criterion in decisions to continue for a second phase, but this is yet to be seen.

3.2. Evaluation methods

As per the guidance provided to the design teams, the process of developing a ToC, and the ToCs themselves, are intended to: i) support system-level strategic planning based on analyses of context and understanding of relevant actors and processes with which the Initiative will interact; ii) build appreciation and understanding of the diverse roles required in change processes; iii) encourage critical thinking, integration, collective visioning, and co-ownership among team members and partners; and iv) support transparency, communication, and accountability to results.

We reviewed the full set of Initiative and Work Package ToCs in the versions submitted as part of the proposals (see CGIAR, n.d.b.) against these criteria. Author 1 copied all ToC narratives into a table for review, highlighting content and adding comments on strengths and weaknesses of each ToC. A second table abstracted each outcome statement by Initiative, with columns to analyze the actor and action indicated in each outcome statement, and a column for comments and observations on strengths and weaknesses of those statements. The tables were shared with authors 2 and 3 to review and provide their input by e-mail and in teleconference calls. All authors were involved in supporting the proposal/ToC development process in some way. Our analysis is informed

Text Box 1

Facilitating questions for WP ToC development.

Facilitating Questions for Work Package ToC Development

1. What is the main problem(s) or development challenge the WP aims to address? (Refer to Initiative and Action Area Challenges and refine as needed. Consider differentiated needs and opportunities to address disadvantaged groups by gender and social class.)
2. What work has been already done to analyze and help address the problem, including in previous CGIAR projects? (This is to help establish the background and rationale for the proposed work.)
3. What impacts and Action Area outcomes will the WP contribute to? (Refer to Impact Areas and Action Area outcomes)
4. What are the most important processes (ongoing or new) that will be needed to realize the intended impacts? (This is to identify strategically-relevant processes external to the Initiative that the WP will need to engage with; e.g. national policy reform; trade association initiatives; international conventions or commitments.)
5. Who are the most important system actors in these processes, including other CGIAR WPs/Initiatives, and what do they need to do to realize the intended impacts?
6. What will motivate those actors to take the needed action(s)? (This is to anticipate and analyze what is needed to inform, support, or otherwise influence actors external to the Initiative.)
7. What research is needed to stimulate or support this change process, and why is it important?
8. What will be the main outputs of the WP's research?
9. What other supporting activities (e.g. capacity development; networking; multi-stakeholder process facilitation; technical support) will the WP undertake and why are they important for realizing intended outcomes?
10. Which partners will the WP work with (including other CGIAR WPs/Initiatives)?
11. What Outcomes of the WP's work will be realized by the end-of-Initiative (in other words, what system actors will do what differently in whole or in part because of the WP's work)? Ideally, these should be stated in concrete, measurable terms (try to avoid relative descriptors like 'improved' or 'better').
12. What next-level outcomes (i.e. beyond and resulting from end-of-Initiative outcomes and precursors to the Action Area outcomes and impacts) are anticipated, recognizing that those outcomes will not be realized within the Initiative time frame.

(Source: CGIAR, 2021, Appendix B)

by this experience.

We consider three key aspects of the ToCs

1. ToC Presentation: Are the ToC diagrams and narratives clear and complete? Do they satisfy the intended purpose?
2. ToC Element Analysis: Are individual ToC elements (output, outcome, and assumption statements) adequately framed and specified sufficiently to satisfy their intended purpose?
3. ToC Evaluability Assessment: Are the ToCs in their current form articulated sufficiently to guide program implementation and evaluation, and how they could be improved?

We adapted Davies' (2012) criteria for assessing the evaluability of ToCs, in combination with the ToC guidance provided by CGIAR (2021), to develop the following assessment criteria for a R4D ToC.

An individual ToC should be:

- Clear. The ToC should be clear and easy to understand for a knowledgeable person unfamiliar with the program. This means that both the diagram and narrative must be clear individually and consistent with one another.
- Explained. The ToC should include sufficient explanation of the causal processes. That is, the underlying assumptions about how each step(s) in a change process is expected to lead to the subsequent step(s) should be explicitly articulated.
- Plausible. The causal logic should be believable in context, with no leaps of logic or other logical fallacies. Plausibility can be enhanced by providing evidence supporting the sequence of events/theory in the ToC.
- Complete. The ToC should include the complete chain of events connecting the intervention (research activities and outputs) to intended outcomes to impacts, with relevant actors specified.
- Integrated. The elements of the ToC should be referred to in, and consistent with, other program documents.

- Inclusive and representative. The ToC should reflect the interests and values of key partners and other system actors. This reflects that the ToC development process was participatory and/or used other means to consider and include stakeholder interests. Ideally, stakeholders have ownership.
- Testable. The ToC should be specified sufficiently, with clear and measurable indicators of results, so it is possible to determine whether and to what degree hypothesized changes (results) have been realized. In other words, the ToC is falsifiable.

Collectively, the set of ToCs should be:

- Comparable. The ToCs should be presented using a standard format for diagrams and narratives to facilitate review and comparison. This means that ToC elements must be defined and framed in a common way.
- Coherent. The nested ToCs (WP, Initiative, AA) should be internally consistent and logically connected.

The focus of the analysis is on the design and specification of ToCs, aiming to identify common strengths and weaknesses across the set. As such, we reviewed only ToC diagrams and narratives; we did not delve deeply into individual proposals to assess the correspondence of ToCs with other proposal components.

4. Results and Discussion

4.1. ToC presentation

4.1.1. ToC diagrams

Initiative design teams were provided with a standard template for Initiative and WP ToC diagrams.

The Initiative ToC diagram template (not shown) represents the combined outcomes of the individual Initiative WPs. The diagram shows

collective WP outcomes and end-of-Initiative outcomes (with some overlap between WPs) and subsequent Action Area Outcomes, Impact Areas, and SDGs

A typical WP ToC diagram is presented as Fig. 1. The WP template provided space for outputs (green boxes), partners/stakeholders (orange figures), outcomes (grey hexagon) and end-of-Initiative outcomes (blue hexagon), with arrows to represent causal relationships, and numbered icons (orange heptagram) as place-holders for assumptions. The assumptions themselves are to be documented separately. The activities necessary to produce the outputs are not shown but are left implicit to keep the diagram as simple as possible.

The prescribed standard format resulted in visually similar ToC diagrams across Initiatives. However, each Initiative (and sometimes individual WPs within the same Initiative) used its own approach to define and present outputs, outcomes, and assumptions. So, despite similarities in design, the content of the ToC diagrams is highly variable. Details of inconsistencies in content are discussed below.

Nevertheless, the diagrams provide a reasonable summary the of aims and intent of each Initiative/WP at a glance.

4.1.2. Narratives

The ToC narratives provide space for a more complete and nuanced explanation of the key system actors and processes, and how the Initiative/WP aims to interact with and influence those actors and processes. No format was specified for the ToC narratives so they are highly diverse. Some begin with a problem statement, some with a purpose statement, and some with background on previous research. Some then provide various levels of detail on some or all stages in the ToC, but most focus on what the Initiative or WP will do or what its thematic focus is, reiterating information on rationale, aims, intentions, and descriptions of research activities and outputs, information that should be provided elsewhere in the proposal. In most cases, the ToC narrative adds little value to the ToC diagram, and it is difficult to compare the various narratives.

Many of the ToC narratives refer to collaborative partnership, co-

generation of knowledge, and similar ideas about contributing to change through engagement with other system actors. While this is a step in the right direction, narratives remained general in most cases. As a next step, ToCs should provide more detail on who the partners are, where, and how many. It is important to explain if a particular partnership is already established and mature or if it needs to be built and fostered, and how that will be done. While many ToCs mention partnerships, few provide any contextual analysis that would indicate what other actors are working on related issues, what other related processes are underway or planned, and how the Initiative/WP intends to interact with those processes. In short, there is a general lack of attention in the ToC narratives to other actors and processes that the Initiative/WP needs to interact with to effectively contribute to change.

As with the ToC diagrams, the narratives reflect a predominantly supply-driven approach. Many express a general sentiment that they will produce outputs that will be used by others, with limited theoretical explanation of why those other actors will want to use the outputs and how that use will contribute to higher-level changes.

4.2. ToC elements

4.2.1. Outputs

The outputs presented in the WP ToC diagrams should represent substantial scientific products (including innovations) or supporting contributions generated by the main research and other activities of the WP. Some Initiatives and WPs did this well, providing clear, well-specified output statements (e.g., “Germplasm of selected advanced lines available for NARS release/registration process” or “Implementation guide to exploit re-use of urban waste in food production”). However, there are several common errors in output formulation that undermine presentation of the intended output and its role in the change process (Table 1). Many outputs are stated in vague and general terms, making it difficult to understand what exactly would be produced and therefore how it would be used. Many use generic framing that could apply to almost any research endeavor, using terms like ‘improved

WP3: AMR

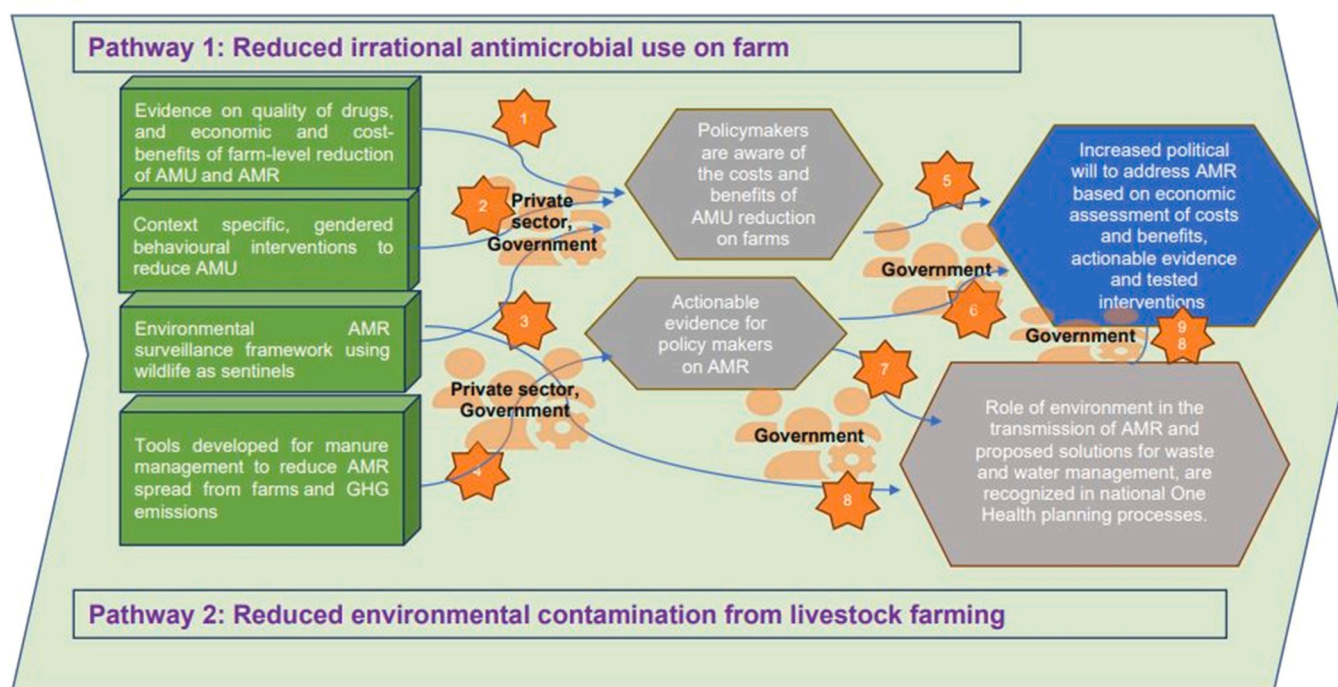


Fig. 1. Example WP ToC Diagram – Protecting human health through a One Health approach Initiative. WP3 Antimicrobial Resistance (AMR). Note: outputs are in green boxes, outcomes in grey, and end-of-initiative outcomes in blue, assumptions are numbered.

Table 1
Examples of common errors in ToC formulation found in Initiative proposals.

Result Level	Type of Error	Example	Comment	Suggestion
Output	Insufficient specification	Working papers or reports analyzing specific aspects of policy coherence in specific sectors or countries	Suggests that the research focus is unclear; the need for specificity is acknowledged.	Indicate what aspects of policy coherence will be addressed in what countries.
Output	Insufficient specification	A collection of tools that have been used or suggested for groundwater governance, including guidance on the conditions under which each could be suitable	Reflects a lack of certainty about what will be produced.	Indicate the most important tools that will be tested.
Output	Redundant information	Decision tool for policy makers and/or project implementers [...]	Intended users should be identified in outcome statements.	Use the space to specify the output.
Output	Insufficient attention to process outputs	No example. Most proposals include activities to strengthen capacity, build relationships, and co-generate knowledge, etc. but these are not reflected in the ToC.	Critical and surprising gap.	Indicate and incorporate key process outputs in ToC.
Output	Confused with outcome	Policies and Institution Landscape Analysis framework applied to food, land and water policies	Application in policy decisions can only be done by responsible authorities beyond the project sphere of control; this is outcome	Revise output statement to specify the Landscape Analysis Framework.
Outcome	No actor identified	Local digital agrifood systems are strengthened through the initiative facilitated open collaborative environment that leads to more than five impact driven use cases that promote inclusive and sustainable impacts	Unclear who will act or what they will do to strengthen systems.	Indicate who will act as specifically as possible.
Outcome	Multiple actors indicated	Governments, NGO's, CSO's, and/or private sector actors in at least three LMC's use [Initiative] learning and guidance to better understand how social innovations and organizational strategies, as well as government and private sector policies, can effectively increase the voice and agency of women in AFS governance and their resilience to CC	Each kind of actor will require different kinds of information and support; confounding them makes it impossible to define appropriate strategy.	Identify the most important actors in the change process and specify separate outcome statements for each.
Outcome	Poorly defined actions	NARES and SME's have increased responsibility and participation in processes and decision making	Difficult to determine whether the outcome has been realized.	Reframe 'increased responsibility and participation' as observable actions; separate outcome statement for the two actor groups (NARES and SMES).
Outcome	Change stated in relative terms	Improved operational clarity, effectiveness, and communication in breeding networks	Baseline data required to measure/ assess any relative change.	Reframe as observable action or specify metric and ensure baseline data are available.
Outcome	Excessive information	Public-private sector, NARES, and civil society actors across sub national agricultural innovation systems in 4 LAC countries use [Initiative] learning knowledge management and evidence to better accelerate on- farm uptake of SCT innovations by making them more gender responsive, production- friendly, and context-specific reaching, at least 200,000 farmers (2022–2024)	The actual outcome is obscured in the long presentation which combines multiple changes.	Reframe as observable action per actor group.
Outcome	Targets included in outcome statement	Bundled [Initiative] climate services will reach at least 300,000 vulnerable farmers, at least 30 % of whom are women, in six focal countries by 2024	The statement does not indicate who will do what; focuses on reach target.	Reframe what actor(s) will take what action(s) to reach vulnerable farmers. Present target separately.
Theoretical Assumption	Inward looking	Initiatives will be effectively coordinated, appropriately targeted and strongly supportive of cereal seed system actors	Focuses on program management (sphere of control), not causal logic.	Reframe to explain why seed system actors can be expected to use support to be provided.
Theoretical Assumption	Weak connection from cause to effect	[...] the empowerment of women and other disadvantaged is seen as a crucial step to achieve long lasting, positive and peaceful development outcomes [...]	Implies that there will be a change of attitude or priorities without explaining how such a change will be realized.	Reframe to explain why it is reasonable to expect the actor to be influenced to prioritize women's empowerment.
Theoretical Assumption	Assertion in place of theoretical assumption	Bundled solutions based on climate-smart and conservation-agriculture approaches help to bridge the production yield gaps and minimize their vulnerability to climatic and other risks [...]	This assumes that the Initiative will produce useful products that will be used, without explaining why.	Reframe to explain how and why the specified actor(s) will be motivated to adopt the intended outputs.
Theoretical Assumption	Wishful thinking in place of theoretical assumption	Key risks and assumptions are that scaling partners will have a strong enough interest to partner with [the Initiative] on the development and testing of value chain interventions (1). This may be a risk as [Initiative] does not intend to finance the scaling process	This is a statement of hope rather than an assumption. The authors recognize that there is not a strategy to help realize the outcome.	Revise the project design or link with other projects to improve the linkage to scaling partners.

evidence base' or 'more robust data'. If instead the output statement specified what evidence base or data and how they will be improved, it would be much more useful to guide implementation and to evaluate achievements. Many output statements include indications of who might use the output or how. This is useful information in a ToC, but should be recorded in an outcome statement, not the output. Some output

statements try to incorporate targets (e.g. how many users will be reached) and others try to signal the process by which the output is to be produced, sacrificing clarity. Overall, the frequency of these errors suggests that proponents lacked experience and understanding about how to formulate a ToC, but also that the research ideas were not fully developed, making it difficult to know exactly what would be produced.

A separate problem is an overall lack of attention to process-oriented outputs. The One CGIAR model explicitly recognizes the need to support capacity-building, relationships, and knowledge translation activities, but relatively few of the ToC outputs relate to social process contributions. This is a significant and puzzling gap as so many of the Initiative proposals include references to partnerships, capacity-building, knowledge co-generation, and other process-oriented elements in their overall strategies.

4.2.2. Outcomes

Very few outcome statements in the set fully meet the expectations stated in the guidelines. Ideally, an outcome should specify an observable action or behavior by a particular actor (e.g., government agency X in country Y) or actor group (e.g., type of government agency in countries of region Z). In this way, outcomes can be used as indicators of progress along the ToC. However, some outcome statements fail to identify any actor and many refer to multiple different actors (see examples in Table 1). This would be appropriate if all the actors listed could be expected to respond in a similar way to the same precursor outputs and outcomes. Most refer to quite different actors that could not be expected to respond in the same way, which undermines the logic of the ToC. Likewise, the actions/behaviors are often poorly defined, stated in relative terms, or not stated at all. This makes it challenging to assess whether an outcome has been realized, and hinders understanding of the logic of subsequent changes modelled in the ToC. Many outcome statements also include long qualifying statements that tend to reduce clarity. Many include targets, which would be better presented elsewhere.

4.2.3. Theoretical assumptions

Theoretical assumptions are indicated in the WP ToC diagrams as numbered lines connecting outputs (and outcomes) to outcomes, end-of-initiative outcomes, and beyond. Some provide a legend in the diagram to present the assumption statements, which makes reading the diagram easier. Others present the assumptions in a separate table or only in the ToC narrative. According to the guidance, theoretical assumptions should provide the theoretical reason that a particular actor is expected to be influenced by precursor results as hypothesized in the outcome statement. In other words, what would motivate a certain actor to behave as predicted? Further downstream, why should the resulting actions of Actor X and Actor Y lead to subsequent actions by Actor Z, and why should that contribute to the AA outcomes and/or impacts targeted by the Initiative?

Ideally, theoretical assumptions in this kind of formulation should reflect a demand analysis; what information, capabilities, technology, incentives, or other support is needed by downstream actors to enable or encourage them to take the actions needed to realize higher-level outcomes and impacts? Relatively few of the current assumptions use an explicitly demand-oriented framing, and those that do tend to be general (see examples in Table 1). Instead, many restate what the Initiative will produce with the implication that the outputs are needed/wanted by users, or simply assert that the outputs will drive downstream actions. Some assumptions identify a need, but make only a loose connection to how the Initiative’s outputs will meet that need. Some mention profitability or other advantages that could appeal to intended users, but without analysis of those users’ needs or wants. Some assumptions are inward-looking, focusing more on program delivery (sphere of control) than on causal logic.

4.3. ToC evaluability assessment

A brief assessment of the overall set of Initiative ToCs against the evaluability criteria is shown in Table 2.

The combined effect of the common errors in ToC formulation and presentation outlined above resulted in a set of ToCs that provide a general indication of the aims and intentions of the research Initiatives,

Table 2
Summary of Initiative ToC assessment.

Criterion	Comments on Current ToCs	Recommendations for Improvement
Clear	ToC diagrams and narratives give only a general sense of the aims of the Initiatives/WPs and how they will be realized.	Specify outcome statements by key actor and observable action.
Logical	Imprecise outputs, poorly formulated outcomes and simplistic causal logic leave many questions about who will do what differently and why.	Provide clear theoretical assumptions to explain why actors will be motivated to act as hypothesized in each Outcome statement.
Plausible	There is typically insufficient explanation of the causal logic to evaluate it.	With well-specified outcomes and assumptions, it will be easier to evaluate plausibility.
Complete	The nested ToCs provide reasonable linkage from project activities all the way to mission-level impacts.	Maintain use of nested ToCs to provide theoretical linkage from activities to impacts.
Integrated	Outcomes to be reported against annually; evaluation to consider use and adaptation of ToC. Some inconsistency between proposals and ToCs (e.g. insufficient attention to process-oriented outputs in ToCs)	Maintain reporting against ToC outcomes and evaluation of ToC use and adaptation; nominate managers with responsibility for ToC oversight.
Inclusive and representative	Little indication that whole teams were included in ToC development, let alone partners and stakeholders. Some Initiatives anticipated partner needs and expectations based on past experience.	Increase attention to and inclusion of stakeholder perspectives.
Testable	Outcome statements poorly formulated, not empirically testable.	Well-specified outcomes and assumptions will be testable.
Comparable	Many different ways of defining and presenting ToC components across the set, in diagrams and narratives.	Increase efforts to standardize terminology, concepts and presentation of ToCs; nominate managers with responsibility for ToC oversight.
Coherent	Initiatives in some AAs are well-coordinated and collectively include a full range of scientific and supporting activities needed for transformational change. Other AAs show less internal coherence.	Continue efforts to develop coherent AA-scale ToCs; nominate managers with responsibility for ToC oversight.

but insufficient specification to be truly useful in program management or evaluation.

5. Conclusions

CGIAR invested in a major effort to incorporate ToC in research planning, design, and evaluation. This included developing standard terminology, guidance, templates for ToC diagrams and narratives, and providing technical support. CGIAR ToCs are built around the concept of spheres of control, influence, and interest, which helps distinguish outputs (sphere of control) from outcomes (sphere of influence) (Montague, 2000; CGIAR, 2021). The concept of End-of-Initiative outcomes facilitates setting realistic results targets within a longer-term ToC. Proposal assessment considered the quality of ToCs, while mid-term and end-of-term evaluation will consider how ToCs have been used and adapted. Annual reporting is organized around the outputs, outcomes, and impacts specified in Initiative and WP ToCs, and collated in a publicly-accessible dashboard (see CGIAR, 2023). Currently, the

results dashboard presents these results at an aggregate level, out of context of their specific ToCs. The same data can be used in context by Initiatives to show progress along the ToC.

Some design teams used ToC integrally, as part of the research strategy and planning. These teams invested more time and effort to develop their ToC diagrams. Teams that used the facilitating questions developed more holistic ToCs. Individualized support and feedback encouraged more coherent ToC development. Some teams have subsequently invested effort to further refine, improve, and update their ToCs (to be considered in future evaluations by CGIAR).

Nevertheless, many of the ToCs reviewed were insufficiently clear and specific, with substantial scope for improvement. The overall proposal development period was short and rushed. Most design teams and their larger research teams were newly created. They faced time pressures and had limited opportunities to meet in-person with their decentralized teams (e.g., pandemic restrictions, time zones). For many Initiatives, research locations (even to the country level) were not yet decided, and partners had not yet been engaged, so detailed contextual analysis could not be done. Moreover, the ToCs were just one part of a demanding process to design large integrated research programs.

There was and still is a general lack of ToC experience and literacy in many teams. Although guidance and technical support was offered, there was also countervailing and confusing information. The ToC narrative section in the proposal template caused problems because it asked for information that is not typically included in a ToC, and the organization was inconsistent with the ToC Guidance. For this reason, with the rushed process, many/most teams did not fully use the ToC Guidance or facilitating questions. In many cases, the ToC was derived from the (draft) research plan and not used, as intended, as a foundation for strategic research design and planning. More and better support is needed for ToC design and use in a R4D context. Proposal templates for ToC diagrams and narratives should provide appropriate structure and clear instructions to encourage good quality and comparable ToC documentation. There is still a need for capacity-building for team leaders and team members to plan and facilitate ToC development with broad team (and, ideally, partner) representation. This kind of capacity-building is probably best done in a workshop setting, where skilled facilitators can guide participants through a process to develop a ToC (Deutsch et al., 2021). It can be supported with refined ToC guidance, facilitating questions, and other resources, along with ongoing support from resource people and external experts. CGIAR and relevant design teams could draw on the ToC evaluability criteria presented above as a checklist as Initiative ToCs are prepared, improved, and reviewed. This could also help build a set of well-formulated ToCs as examples going forward.

Research planning and evaluation discussions are replete with ambiguous terminology and concepts. Although CGIAR ToC Guidance provided definitions for key terms, some teams inevitably fell back to their own familiar definitions of terms like output, outcome, impact, and assumptions. This ambiguous use of terms and concepts leads to incoherence and incomparability within and between ToCs (and the related research strategies and plans). This explains some of the inconsistencies and weaknesses observed in the current set of ToCs: poor formulation of outputs; poor definition of outcomes; and weak assumptions. For example, there is still a commonly understood concept of assumptions as used in logframes as ‘factors beyond the control of the program’. Some of the ToCs framed assumptions in this way rather than as explanations of the causal reasoning. Similarly, it is common in proposals to refer to outputs in terms of publications or other communication media, rather than in terms of the knowledge content produced, as intended in a ToC. Furthermore, there is chronic confusion in defining outcomes and impacts (Belcher and Palenberg, 2018).

The various practical challenges to ToC development notwithstanding, a key observation is that there is still a prevalent supply-driven approach to research design in CGIAR. Despite concerted efforts to develop stronger and broader partnerships, promote knowledge co-

generation, and strategically situate R4D within broader development processes, many of the current ToCs are still built around simple assumptions that good, well-executed science will be taken up and used in positive ways. As noted by Woltering et al. (2019), project level ToCs in research organizations often use linear cause-effect chains even though they rely on societal transformation of knowledge with multiple feedback loops. This was reflected in the lack of attention to process-oriented outputs in the ToCs, despite abundant discussion of the importance of partnerships and co-generation in the proposals. ToC puts emphasis on understanding the needs of intended research users and applies science as appropriate to address those needs, while recognizing and, to some extent, supporting, other actors and process to facilitate change.

Based on this case we can propose some recommendations to improve development and use of ToC. First, develop capacity in conceptualizing research impact pathways which focus on the needs of intended research users; second improve templates for new program proposals, especially for ToC narratives; third, use the ToC evaluability assessment criteria presented above as a checklist for ToC design teams to review and improve proposals; fourth, build appropriate skills for ToC development, use, review, and updating; fifth, ensure that teams have dedicated monitoring and evaluation specialists to guide ToC development and enable effective integration with monitoring and reporting; and finally, include ToC accountability in the terms of reference of research leaders.

Funding

This work was supported by CGIAR Portfolio Performance Unit, CIP, and the Royal Roads University Ashoka Chair in Sustainability Research Effectiveness.

CRediT authorship contribution statement

Brian Belcher: Writing – original draft, Project administration, Methodology, Conceptualization. **Enrico Bonaiuti:** Writing – review & editing, Validation, Conceptualization. **Graham Thiele:** Writing – review & editing, Validation, Conceptualization.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Brian Belcher reports financial support was provided by CGIAR. All Authors reports a relationship with CGIAR that includes: consulting or advisory, employment, and travel reimbursement. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Submissions for the CGIAR 2022–2024 Investment Prospectus are available at: <https://bit.ly/3SWBW1B>

Acknowledgements

We thank Marco Ferroni, Pascale Sabbagh, David Spielman, Scarlett Crawford, Jim Hammond and Lennart Woltering for excellent feedback on an earlier version of the paper, and Rachel Claus for support with formatting and submission.

References

- Apgar, J.M., Allen, W., Albert, J., Douthwaite, B., Paz Ybarnegaray, R., Lunda, J., 2017. Getting beneath the surface in program planning, monitoring and evaluation: Learning from use of participatory action research and theory of change in the

- CGIAR Research Program on Aquatic Agricultural Systems. *Action Res.* 15 (1), 15–34. <https://doi.org/10.1177/1476750316673879>.
- Belcher, B., Palenberg, M., 2018. Outcomes and impacts of development interventions: toward conceptual clarity. *Am. J. Eval.* 39 (4), 478–495. <https://doi.org/10.1177/1098214018765698>.
- Belcher, B.M., Hughes, K., 2021. Understanding and evaluating the impact of integrated research-oriented research programmes: concepts and considerations. *Res. Eval.* 30 (2), 154–168. <https://doi.org/10.1093/reseval/rvaa024>.
- Belcher, B.M., Claus, R., Davel, R., Ramirez, L.F., 2019. Linking transdisciplinary research characteristics and quality to effectiveness: a comparative analysis of five research-for-development projects. *Environ. Sci. Policy* 101, 192–203. <https://doi.org/10.1016/j.envsci.2019.08.013>.
- Belcher, B.M., Davel, R., Claus, R., 2020. A refined method for theory-based evaluation of the societal impacts of research. *MethodsX* 7, 100788. <https://doi.org/10.1016/j.mex.2020.100788>.
- Better Evaluation. (n.d.). Develop programme theory / theory of change. Retrieved July 17, 2024 from <https://www.betterevaluation.org/frameworks-guides/rainbow-framework/define/develop-programme-theory-theory-change>.
- Bozeman, B. (2003). Public Value Mapping of Science Outcomes: Theory and Method. A Monograph of the Public Value Mapping Project of the Center for Science, Policy and Outcomes. Centre for Science, Policy and Outcomes and School of Public Policy, Georgia Tech. Retrieved July 17, 2024 from (<https://cspo.org/legacy/libraries/110215F3AQ.lib.BozemanPVMKnowle.pdf>).
- Brest, P., 2010. The power of theories of change. *Stanf. Soc. Innov. Rev.* 8 (2), 47–51. <https://doi.org/10.48558/NPZS-C374>.
- Byerlee, D., Lynam, J.K., 2020. The development of the international center model for agricultural research: a prehistory of the CGIAR. *World Dev.* 135, 105080.
- CGIAR. (2021). Theory of Change (ToC) Design Guidance Document: Proposal Stage. Retrieved July 17, 2024 from <https://docs.google.com/document/d/1Uok7uspauk5P2oLz6vTDKSPvD5b4hh-/edit#heading=h.gjdgxs>.
- CGIAR. (2023). CGIAR Results Dashboard. Retrieved January 16, 2024 from (<https://www.cgiar.org/food-security-impact/new-results-dashboard/>).
- CGIAR. (n.d.a.). CGIAR 2030 Research and Innovation Strategy. Montpellier, France: CGIAR System Organization. Retrieved January 16, 2024 from <https://cgspace.cgiar.org/bitstream/handle/10568/110918/OneCGIAR-Strategy.pdf>.
- CGIAR (n.d.b.). Submissions for the CGIAR 2022–2024 Investment Prospectus. Retrieved July 17, 2024 from <https://www.cgiar.org/how-we-work/governance/system-council/initial-set-of-submissions-for-the-cgiar-2022-2024-investment-prospectus/>.
- Chambers, R., Pettit, J., 2004. Shifting Power to Make a Difference. In: Grove, L., Hinton, R. (Eds.), *Changing Power and Relationships in International Development*. Earthscan, London, Canada, pp. 137–162.
- Coccia, F., Rouge, J.-C., Claus, R., Davel, R., Hughes, K., Pinto, D., Gotor, E., Jones, S., Thuerer, R., Rachman, R., Wakesho, G., Nijbroek, R., Saigenji, Y., Mausch, K., Woldeyohanes, T., Kimaiyo, J., Kozicka, M., Wiederkkehr Guerra, G., Belcher, B., Meybeck, A., Gitz, V., 2022. Assessing the Impact of 10 years of FTA Research: a synthesis of the 5 integrated impact studies. FTA Working Paper #14. Center for International Forestry Research, Bogor, Indonesia. Retrieved January 16, 2024 from <https://doi.org/10.17528/cifor/008514>.
- Davies, R. (2012). Criteria for assessing the evaluability of Theories of Change. Rick on the Road. Retrieved October 18, 2023 from <http://mandenews.blogspot.co.uk/2012/04/criteria-for-assessing-evaluability-of.html>.
- Delahais, T., Toulemonde, J., 2017. Making rigorous causal claims in a real-life context: has research contributed to sustainable forest management? *Evaluation* 23 (4), 370–388. <https://doi.org/10.1177/1356389017733211>.
- Deutsch, L., Belcher, B., Claus, R., Hoffmann, S., 2021. Leading inter-and transdisciplinary research: Lessons from applying theories of change to a strategic research program. *Environ. Sci. Policy* 120, 29–41. <https://doi.org/10.1016/j.envsci.2021.02.009>.
- Douthwaite, B. (n.d.). PIPA usage. Retrieved January 16, 2024 from <http://boru.pbworks.com/w/page/13774911/PIPA%20Usage>.
- Douthwaite, B., & Ashby, J. (2005). Innovation histories: A method for learning from experience. ILAC Brief 5. Institutional Learning and Change Initiative, CGIAR. Retrieved January 16, 2024 from <https://hdl.handle.net/10568/70176>.
- Douthwaite, B., Kubly, T., van de Fliet, E., Schulz, S., 2003. Impact pathway evaluation: an approach for achieving and attributing impact in complex systems. *Agric. Syst.* 78 (2), 243–265. [https://doi.org/10.1016/S0308-521X\(03\)00128-8](https://doi.org/10.1016/S0308-521X(03)00128-8).
- Douthwaite, B., Alvarez, S., Cook, S., Davies, R., George, Pamela, Howell, J., Mackay, R., Rubiano, J., 2007. Participatory impact pathways analysis: a practical application of program theory in research-for-development. *Can. J. Program Eval.* 22 (2), 127–159. <https://doi.org/10.3138/cjpe.22.007>.
- van Es, M., Guijt, I., Vogel, I. (2015) Hivos ToC Guidelines Theory of change thinking in practice: A stepwise approach. The Hague, Netherlands: Hivos. Retrieved June 10, 2024 from <https://hivos.org/assets/2020/10/hivos.toc.guidelines.pdf>.
- Halimanjaya, A., Belcher, B., Suryadarma, D., 2018. Getting forest science to policy discourse: a theory-based outcome assessment of a global research programme. *Int. For. Rev.* 20 (4), 469–487. <https://doi.org/10.1505/146554818825240638>.
- Hambly, H., Friedmann, M., Proietti, M., Polar, V., Fernandes, S., Thiele, G., 2022. Innovation models to deliver value at scale: The RTB Program. In: Thiele, G., Friedmann, M., Campos, H., Polar, V., Bentley, J. (Eds.), *Root, Tuber and Banana Food System Innovations: Value Creation for Inclusive Outcomes*. Springer International Publishing, Cham, pp. 29–69. https://doi.org/10.1007/978-3-030-92022-7_2.
- Hessels, L., van Rijn, F., de Jong, S., Reijmerink, W., van Houwelingen, J., McIntosh, A., Jan Ellen, G., & Koers, G. (2023, October 9). How to make Theories of Change deliver societal impact. LSE Impact Blog. Retrieved January 16, 2024 from (<https://blogs.lse.ac.uk/impactofsocialsciences/2023/10/09/how-to-make-theories-of-change-deliver-societal-impact/>).
- Holderness, M., Howard, J., Jouini, I., Templeton, D., Iglesias, C., Molden, D., & Maxted, N. (2021). Synthesis of Learning from a Decade of CGIAR Research Programs. Retrieved January 16, 2024 from https://iaes.cgiar.org/sites/default/files/pdf/June21_2021%20Synthesis%20Report_Final%20Updated%2002_2022.pdf.
- Independent Science for Development Council (ISDC). (2020). Technical Note: Quality of Research for Development in the CGIAR Context. Rome, Italy: ISDC. Retrieved January 16, 2024 from <https://iaes.cgiar.org/isdc/publications/quality-research-development-cgiar-context-0>.
- Jagannathan, K., Emmanuel, G., Arnott, J., Mach, K.J., Bamzai-Dodson, A., Goodrich, K., Meyer, R., Neff, M., Dana Sjöstrom, K., Timm, K., Turnhout, E., Wong-Parodi, G., Bednarek, A., Meadow, A., Dewulf, A., Kirchhoff, C., Moss, R., Nichols, L., Oldach, E., Lemos, M., Klenk, N., 2023. A research agenda for the science of actionable knowledge: Drawing from a review of the most misguided to the most enlightened claims in the science-policy interface literature. *Environ. Sci. Policy* 144, 174–186. <https://doi.org/10.1016/j.envsci.2023.03.004>.
- James, C. (2011). Theory of Change Review. A report commissioned by Comic Relief, Comic Relief. Retrieved January 16, 2024 from <http://mande.co.uk/2012/uncategorized/comic-relief-theory-of-change-review/>.
- Joly, P.-B., Gaunand, A., Colinet, L., Larédo, P., Lemarié, S., Matt, M., 2015. ASIRPA: a comprehensive theory-based approach to assessing the societal impacts of a research organization. *Res. Eval.* 24, 440–453. <https://doi.org/10.1093/reseval/rvv015>.
- Kirchhoff, C.J., Carmen Lemos, M., Dessai, S., 2013. Actionable knowledge for environmental decision making: broadening the usability of climate science. *Annu. Rev. Environ. Resour.* 38, 393–414. <https://doi.org/10.1146/annurev-environ-022112-112828>.
- Lam, S., 2020. Toward learning from change pathways: reviewing theory of change and its discontents. *Can. J. Program Eval.* 35 (2), 188–203. <https://doi.org/10.3138/cjpe.69535>.
- Maini, R., Mounier-Jack, S., Borghi, J., 2018. How to and how not to develop a theory of change to evaluate a complex intervention: reflections on an experience in the Democratic Republic of Congo. *BMJ Glob. Health* 3 (1), e000617. <https://doi.org/10.1136/bmjgh-2017-000617>.
- Mayne, J., Johnson, N., 2015. Using theories of change in the CGIAR research program on agriculture for nutrition and health. *Evaluation* 21 (4), 407–428. <https://doi.org/10.1177/1356389015605198>.
- Meadow, A.M., & Owen, G. (2021). Planning and Evaluating the Societal Impacts of Climate Change Research Projects: A guidebook for natural and physical scientists looking to make a difference. Retrieved January 16, 2024 from https://swcasc.arizona.edu/sites/default/files/2022-03/Meadow-Owen_Societal-Impacts_Guidebook.pdf.
- Meinke, H., Nelson, R., Kocic, P., Stone, R., Selvaraju, R., Baethgen, W., 2006. Actionable climate knowledge: from analysis to synthesis. *Clim. Res.* 33 (1), 101–110. <https://doi.org/10.3354/cr033101>.
- Meinke, H., Ash, A., Barrett, C.B., Grove Smith, A., Graff Zavin, J.S., Abera, F., Garcia, M., Just, D.R., Obokoh, N.H., Kadiyala, S., Negra, C., Torrance, L., Beaudreault, A.R., Boulanger, P., 2023. Evolution of the One CGIAR's research and innovation portfolio to 2030: approaches, tools, and insights after the reform. *NPJ Sustain. Agric.* 1 (6), 1–8. <https://doi.org/10.1038/s44264-023-00005-x>.
- Meyer, M., 2010. The rise of the knowledge broker. *Sci. Commun.* 32 (1), 118–127. <https://doi.org/10.1177/1075547009359797>.
- Montague, S. (2000). Circles of influence: An approach to structured, succinct strategy. Retrieved January 9, 2024, from www.pmn.net/wp-content/uploads/Circles-of-Influence.pdf.
- Oberlack, C., Breu, T., Giger, M., Harari, N., Herweg, K., Mathez-Stiefel, S.-L., Messerli, P., Moser, S., Ott, C., Providoli, I., Tribaldos, T., Zimmermann, A., Schneider, F., 2019. Theories of change in sustainability science: understanding how change happens. *GAIA-Ecol. Perspect. Sci. Soc.* 28 (2), 106–111. <https://doi.org/10.14512/gaia.28.2.8>.
- Omoro, A., Kidoido, M., Twine, E., Kurwijila, L., O'Flynn, M., Githinji, J., 2019. Using “theory of change” to improve agricultural research: recent experience from Tanzania. *Dev. Pract.* 29 (7), 898–911. <https://doi.org/10.1080/09614524.2019.1641182>.
- Pielke, R.A., 2007. *The Honest Broker: Making Sense of Science in Policy and Politics*. Cambridge University Press, Cambridge, UK.
- Prinsen, G., Nijhof, S., 2015. Between logframes and theory of change: reviewing debates and a practical experience. *Dev. Pract.* 25 (2), 234–246. <https://doi.org/10.1080/09614524.2015.1003532>.
- Ramirez, L.F., Belcher, B.M., 2019. Stakeholder perceptions of scientific knowledge in policy processes: a Peruvian case-study of forestry policy development. *Sci. Public Policy* 46 (4), 504–517. <https://doi.org/10.1093/scipol/scz003>.
- Ramirez, L.F., Belcher, B.M., 2020. Crossing the science-policy interface: lessons from a research project on Brazil nut management in Peru. *For. Policy Econ.* 114, 101789. <https://doi.org/10.1016/j.forpol.2018.07.018>.
- Sarewitz, D., 2016. Saving science. *N. Atlantis* 49 (37), 4–40 (<https://www.thenewatlantis.com/publications/saving-science>).
- Stein, D., & Valters, C. (2012). Understanding theory of change in international development. Justice and Security Research Programme. London School of Economics and Political Science: London, UK. Retrieved June 10, 2024 from https://www.theoryofchange.org/wp-content/uploads/toc_library/scz/pdf/UNDERSTANDINGTHEORYOFChangeSteinValtersPN.pdf.

- Vogel, I. (2012). Review of the Use of 'theory of change' in International Development, Review Report, Department for International Development. Retrieved January 16, 2024 from https://www.theoryofchange.org/pdf/DFID_ToC_Review_VogelV7.pdf.
- Woltering, L., Fehlenberg, K., Gerard, B., Ubels, J., Cooley, L., 2019. Scaling from reaching many to sustainable systems change at scale: a critical shift in mindset. *Agric. Syst.* 176, 102652 <https://doi.org/10.1016/j.agsy.2019.102652>.
- Woolley, R., Molas-Gallart, J., 2023. Research impact seen from the user side. *Res. Eval.* 32 (3), 591–602. <https://doi.org/10.1093/reseval/rvad027>.